Towards a Framework for Socially Influencing Systems: Meta-analysis of Four PLS-SEM Based Studies

Agnis Stibe(✉)

MIT Media Lab, Cambridge, MA, USA
agnis@mit.edu

Abstract. People continuously experience various types of engagement through social media, mobile interaction, location-based applications, and other technologically advanced environments. Often, integral parts of such socio-technical contexts often are information systems designed to change behaviors and attitudes of their users by leveraging powers of social influence, further defined as socially influencing systems (SIS). Drawing upon socio-psychological theories, this paper initially reviews and presents a typology of relevant social influence aspects. Following that, it analyzes four partial least squares structural equation modeling (PLS-SEM) based empirical studies to examine the interconnectedness of their social influence aspects. As a result, the analysis provides grounds for seminal steps towards the development and advancement of a framework for designing and evaluating socially influencing systems. The main findings can also deepen understanding of how to effectively harness social influence for enhanced user engagement in socio-technical environments and guide persuasive engineering of future socially influencing systems.

Keywords: Socially influencing systems · Framework · Persuasive technology

1 Introduction

The dynamic evolution of social media, mobile connectivity, and a digital economy is continuously reshaping how businesses approach and engage customers [1]. Rapidly growing connectedness not only provides new methods for organizations to retain existing relationships with consumers, but also opens new ways to enrich customer engagement experiences and foster innovation [21]. Along the way, businesses and customers tend to naturally follow new market trends and steadily develop an understanding of the spectrum of opportunities provided by emerging technologies. People seamlessly acquire new habits of interaction and consumption behavior, which then set their expectations about how products and services should be designed [34].

Customers increasingly demand products and services that better match their needs and individual preferences [29]. Therefore, businesses face a need to continuously understand the individual and evolving expectations of their customers [26]. Thus, organizations stand to benefit from systems that are designed to reach customers more proactively and provide convenient ways for interaction [32].

T. MacTavish and S. Basapur (Eds.): PERSUASIVE 2015, LNCS 9072, pp. 172–183, 2015.
DOI: 10.1007/978-3-319-20306-5_16

The Internet has become increasingly mobile and social over the last decade [1]. Social media has rapidly expanded and businesses tend to use social media more often for the development of customer relationships. Simultaneously, these advancements exert various effects on everyday life by changing human behavior in both virtual and physical spaces. For example, it has become common for people to use social media through mobile devices [10]. The combination of socially dynamic and technologically advanced contexts has gradually introduced unique modes for businesses to engage customers almost instantly. Such socio-technical spaces often comprise information systems designed to change behavior and attitudes of their users by leveraging powers of social influence, further described as socially influencing systems [38].

To enrich an understanding of how to effectively harness social influence for enhanced user engagement through socio-technical environments, this paper continues with the following sections. In the next, a social cognitive perspective on human behavior is described. Then, the paper introduces a concept of socially influencing systems and reviews a typology of relevant social influence aspects. Thereafter, it provides a meta-analysis of four empirical studies based on partial least squares structural equation modeling (PLS-SEM) methodology and reports seminal steps towards a framework structure for designing and evaluating socially influencing systems. Finally, implications of the main outcomes are discussed and conclusions drawn.

2 Social Cognitive Foundation

According to Ryan and Deci [36], whether people become proactive and engaged depends largely on the social environments in which they develop and function. Bandura [5] has extended this perspective by suggesting that human self-development, adaptation, and change are embedded in social systems. In such systems, according to social cognitive theory [4], personal, behavioral, and environmental factors all interact continuously, perpetually influencing each other and determining the effect of each. There is an endless dynamic interplay between people, their behavior, and the environments where their behavior occurs.

The described triadic reciprocal determinism unfolds multiple angles for studying behavioral change, including environmental and personal change. Human behavior alters environmental conditions and, in turn, is changed by the same conditions that it creates [4]. Along the same vein, social cognitive theory suggests exploring how ambient environments maintain aspects of social persuasion.

Theories of persuasion often aim at describing either influences on or changes in behavior and attitudes on individual, group, and societal levels [33]. According to Fogg [19], persuasive technologies can be designed as social actors and are therefore capable of social influence even in the absence of other people in an immediate physical space. When properly designed, such persuasive technologies can become very effective for inducing behavioral and attitudinal changes in novel socio-technical contexts. Exploring the ability of persuasive technologies and systems to engage users is an essential direction for future research [9].

3 Socially Influencing Systems (SIS)

Social influence, as a substantive phenomenon, has a longstanding history in psychology [12], providing insights on various forms of potential influences on human behavior by the actual, imagined, or implied presence of other people [35]. Along its history, social influence has often been associated with compliance, identification, internalization, obedience, and persuasion, although at the same time kept distinct from conformity, power, and authority. Recent research on social influence mainly has been addressing either minority influence in group settings, dynamic social impact theory, social influence in expectation states theory, or persuasion [8], the latter being broadly defined as change in behavior or attitudes due to information received from others [13-14].

Human beings can experience social influence not only from others in physical proximity around them, but likewise through information systems that are engineered to serve such purpose. Information systems can exert social influence through their design and user interfaces when augmented with relevant social influence aspects, such as *social learning*, *social comparison*, *normative influence*, *social facilitation*, *cooperation*, *competition*, and *recognition* [38]. An information system becomes socially influencing when it has been enriched with social influence aspects to facilitate changes in behavior and attitudes of its users (Fig. 1).

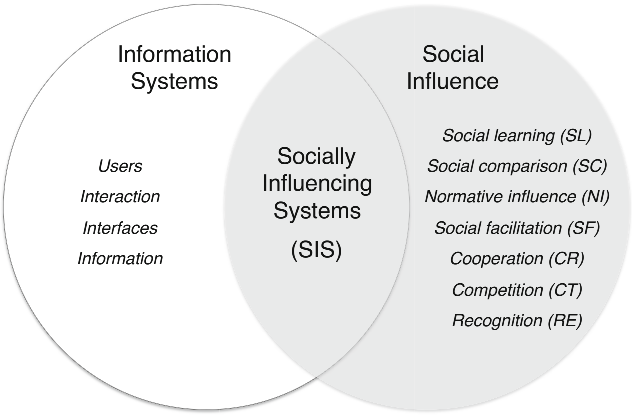


Fig. 1. Socially Influencing Systems

Besides a general comprehension of information system development and software engineering, designers of *socially influencing systems* are required to acquire or maintain a decent level of understanding about human behavior and social psychology. The foundation of the theory and research on socially influencing systems is underpinned by a list of fundamental theories originating from affined areas of social and cognitive psychology. The following are primary theories that are used in this research: social cognitive theory [4], social comparison theory [17], focus theory of normative conduct [13], social facilitation theory [20], cooperation theory [2], competition theory [15], and taxonomy of intrinsic motivators [25].

The listed theories suggest multiple sources of reference for the seven aforementioned social influence aspects that have the capacity to alter behavior and attitudes of users of socially influencing systems. Table 1 summarizes the descriptions, implementation examples, and relevant references of each aspect, while Fig. 2 provides more graphical representation of their sub-dimensions that are discussed and *highlighted* in the following sub-sections.

Table 1. Social influence aspects

Aspect	Description	Implementation example
Social learning (SL) [3-5]	Learning new behavior by observing how other people perform them.	Enabling users to see how others are using a system.
Social comparison (SC) [17], [40-42]	Comparing a behavior of an individual with behavior of others.	Names of active users grow larger as compared to passive users.
Normative influence (NI) [13], [16], [23]	People tend to follow norms and experience peer pressure.	Presenting normative statements or how a majority of others behave.
Social facilitation (SF) [20], [44]	Influence on an individual when surrounded or watched by others.	Displaying how many others are using a system at the same time.
Cooperation (CR) [2], [25], [27]	Activity aimed at achieving a common goal or working together.	Exposing results of cooperative efforts through a system.
Competition (CT) [15], [25], [28]	Endeavoring to gain what others are striving to gain at the same time.	Demonstrating a list of users who are ordered based on their performance.
Recognition (RE) [6], [25], [37]	Value that a person derives from gaining acceptance and approval from others.	Receiving a special title that is displayed to everybody through a system.

3.1 Social Learning

People learn from others by observing their behavior in social contexts [3]. This implies that the information from one individual to another can be transferred through imitation, teaching, and spoken or written language. According to Bandura [4], social learning is ubiquitous and potent, because it allows people to avoid the costs of individual learning. Accordingly, new behavioral patterns can be obtained through observational learning, for example, to share knowledge [11].

3.2 Social Comparison

When individuals use information about other people to evaluate themselves, they engage in social comparison [17]. Specifically, social comparison is described as the process of thinking about others in relation to the self [42]. This process affects motivation, because people tend to look for self-enhancement when comparing themselves *downward* with others who are worse off [40], or individuals look *upward* for self-improvement when seeking a positive example for comparison [41]. In any case, social comparison affects human attitude and behavior [31].

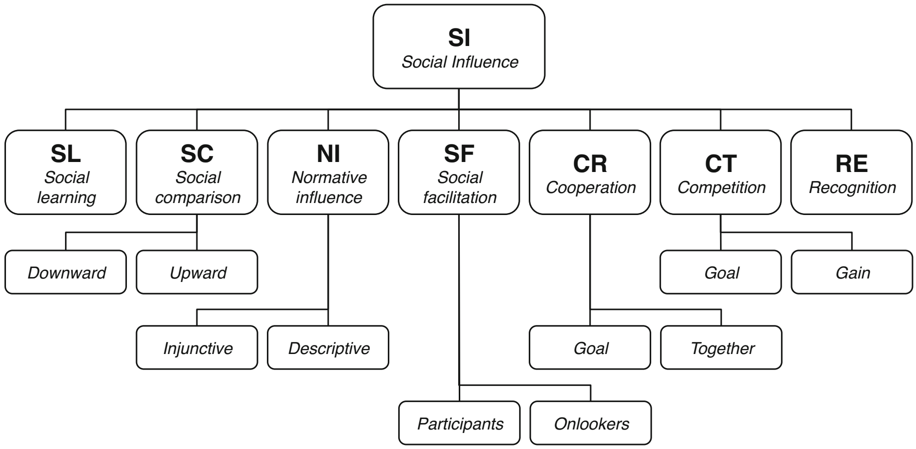


Fig. 2. Structure of the social influence aspects and their sub-dimensions

3.3 Normative Influence

Influence of other people also leads individuals conforming in order to be liked and accepted [16]. Such human action is guided by perceptions of the popularity of certain behavior, that is, by social norms. Research emphasizes that both injunctive and descriptive norms are effective in altering behavior and attitudes of people [23]. *Injunctive* norms inform individuals about what ought to be done, whereas *descriptive* norms refer to what most people actually do [13]. Thus, normative influence affects a wide range of behaviors, e.g. blog usage [22].

3.4 Social Facilitation

The mere or imagined presence of other people in social situations creates an atmosphere of evaluation, which enhances the speed and accuracy of well-practiced tasks, but reduces the performance of less familiar tasks [43]. These social facilitation effects occur in the presence of either passive *onlookers*, or people who are active *participants*, or both [20]. As a result, these effects influence human behavior [44].

3.5 Cooperation, Competition, and Recognition

Interpersonal factors of cooperation, competition, and recognition provide intrinsic motivation that would not be present in the absence of other people [25]. Competition and cooperation are directed toward the same social end by at least two persons [27]. On a social level, people cooperate when they are striving to achieve the same *goal* or are they are working *together*, but compete when they are trying to achieve the same *goal* that is scarce or are seeking to *gain* what others are endeavoring to gain at the same time [28].

Combining the scores of independent tasks performed by different people can encourage cooperation, but providing some salient metric for individuals to compare their performances can promote competition [25]. Meanwhile, recognition can be experienced after competing or cooperating with others [37] or can simply be enjoyed when gaining acceptance and approval from others [6]. The three motivating factors influence various behaviors, including learning [25] and the use of podcasts for generating a sense of community [18].

4 Meta-analysis

To enrich an understanding of how to effectively harness the previously reviewed social influence aspects for enhanced user engagement through socio-technical environments, this section presents a meta-analysis of four empirical studies conducted using partial least squares structural equation modeling (PLS-SEM) methodology.

The four studies (Table 2) were selected based on the shared methodological approach, context, and equal granularity in the exploration of three to seven of the social influence aspects [38]. To the best of accessible knowledge, there were no other comparable studies for inclusion into the present meta-analysis of four socially influencing systems.

Table 2. The list of analyzed studies [38]

Study	Description	Examined aspects
I	Empirical study involving 37 users of a socially influencing system designed for feedback collection through situated displays integrated with Twitter	SL, SF, CR, CT, RE
II	Empirical study involving 69 users of a socially influencing system designed for feedback collection through situated displays integrated with Twitter	SL, SC, NI
III	Empirical study involving 101 participants and a socially influencing system designed for collaborative engagement through situated displays integrated with Twitter	SL, SF, CR
IV	Empirical study involving 77 users of a socially influencing system designed for feedback collection through situated displays integrated with Twitter	SL, SC, NI, SF, CR, CT, RE

The analysis was performed in several consecutive steps. First, the structural models from all studies were reviewed in terms of the present social influence aspects and directed paths (arrows) interconnecting them (Table 3). Second, all seven aspects were mapped out into a single model and all arrows from original models were drawn into the new model. Third, if there were several arrows connecting a pair of aspects, i.e. same directed path originated from several studies (Table 3), only one arrow was kept to represent all of them.

Table 3. Summary of the directed paths from all studies

Directed path between aspects		From study
Social facilitation → Social learning	SF → SL	III
Social facilitation → Social comparison	SF → SC	IV
Social facilitation → Cooperation	SF → CR	I
Social facilitation → Competition	SF → CT	I
Social learning → Social comparison	SL → SC	II
Social learning → Normative influence	SL → NI	II
Social learning → Cooperation	SL → CR	I, III, IV
Social comparison → Normative influence	SC → NI	II
Social comparison → Competition	SC → CT	IV
Competition → Recognition	CT → RE	IV
Recognition → Cooperation	RE → CR	I, IV
Cooperation → Normative influence	CR → NI	IV

Fourth, to obtain deeper understanding of particular interaction effects between the social influence aspects, each arrow was reviewed separately before its inclusion in the final framework structure of the meta-analysis (Fig. 3).

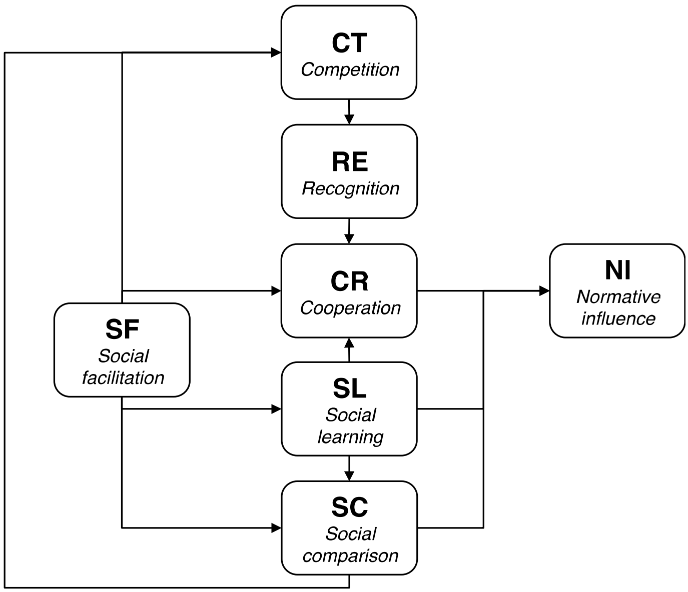


Fig. 3. Framework structure of the social influence aspects

In further analysis, the social facilitation [20] aspect was found to be the only aspect that has no inbound arrows from other social influence aspects in the original structural models. At the same time, the social facilitation aspect directly affected the social learning (study III), social comparison (study IV), competition, and cooperation (both in study I) aspects. The social learning [3-4] aspect was found as one of the most central aspects in all of the studies - having strong direct effects on the cooperation (studies I, III, and IV), normative influence (study II), and social comparison (study II) aspects. The social comparison [17] aspect was found to have strong direct effects on the normative influence (study II) and competition (study IV) aspects.

The competition [15] aspect was found to have strong direct effects on the recognition aspect (study IV). The recognition [6] aspect, in turn, was found to have a strong direct effect on the cooperation aspect (studies I and IV). But, the cooperation [2] aspect was found to have a strong direct effect on the normative influence aspect (study IV). Finally, the normative influence [13] aspect was found to be the only aspect that has no outbound arrows to other social influence aspects.

5 Towards a Framework for Socially Influencing Systems

The reviewed social influence aspects and the output of the meta-analysis provide grounds for making seminal steps towards the development of a solid framework for designing and evaluating socially influencing systems. The achieved results can help deepen understanding of how social influence aspects are affecting each other, and therefore advancing further theory development with regards to the role of each aspect in explaining and predicting how influential an envisioned socially influencing system will be.

5.1 How to Read the Framework

Designers of future socially influencing systems will find the structural framework (Fig. 3) beneficial with some additional guidance. Therefore, this section provides basic instruction for interpreting the presented interconnectedness of the social influence aspects. For example, the social facilitation aspect does not have any inbound arrows but has four outbound arrows directed to other aspects. This implies that social influence effects may commence as soon as other people are present [20]. In the presence of others [44], people can begin to learn from others [3-5], to compare themselves with others [17], and to cooperate [2] or to compete [15] with them.

Following the same logic, the arrow from the social comparison to the competition aspect implies that when people are able to compare themselves with others [17] they are likely to be prompted to compete [15] with those who are better than them, which also might create a sense of social norms [16]. The arrow from the competition to the recognition aspect explains that people who are ranked higher naturally receive some kind of public recognition [25] as others can see how well they have performed. Meanwhile, those who receive public recognition [6] can become more motivated to keep up their excellent performance, which means that they would continue contributing to a collective goal in a cooperative context [28].

Social learning has always played an important role in the evolution of mankind [3-5], because it helps kids to learn quickly just by observing what adults do, to put it simply. The performed meta-analysis reveals that the vast majority of arrows going out from the social learning aspect bear very strong, large, and highly significant effects on succeeding aspects. The framework presents that in a social context people can learn how to compare themselves to others [17], cooperate [27], and read or create an understanding about social norms [16]. Besides, the more people cooperate the more likely they will experience cooperation as a norm for the particular occasion.

In the framework, the arrows are not meant to be completely isolated, i.e. if there is an aspect which has both inbound and outbound arrows, then there is a high likelihood that the aspect also plays a mediating role. For example, besides the direct effect of the social learning aspect on the normative influence aspect, the reviewed studies [38] also reveal significant mediating effects of both the social comparison and cooperation aspects on the relationship.

5.2 Implications for Further Research

The results of the meta-analysis reveal the strength and prominence of social influence aspects in designing socially influencing systems for user engagement. While the world gets increasingly interconnected, such systems could help build novel sociotechnical environments for active participation and contribution rather than for passive consumption [30].

The paper reviewed four socially influencing systems (Table 2) that are potentially applicable and useful for engaging people in a wide range of contexts, including business and education. According to earlier studies, socially influencing systems could enable organizations to facilitate innovative collaborations with customers [32], designing novel models for better anticipation of market changes [34], effective responses to customer needs, and catalyzing innovations [21]. In education, these systems could positively impact student learning and engagement [7].

Further research should focus on testing the current framework and expanding research into other potential social influence aspects. Such studies would contribute to a richer and more elaborate understanding of various social influence aspects and their effects when software features them in user interfaces. The next steps should also include a deeper analysis of how social influence aspects can explain the perceived persuasiveness of socially influencing systems and predict user involvement, participation, and engagement with such systems [38].

Another direction for further research would be to study the design of particular implementations of social influence aspects. The number of different designs for a single social influence aspect is limitless. Thus, this research direction would reveal new design patterns that might have increased potential to shape user behavior and attitude. These designs then should be applied and tested in various contexts to find their best fit.

6 Conclusions

The presented meta-analysis is a highly relevant and timely research effort, because it advances the methodology for engineering future socially influential systems [24]. Along these lines, the present paper provides both researchers and practitioners with richer insights on how social influence aspects can help them to build socio-technical environments aimed at facilitating behavioral and attitudinal changes.

Drawing upon a list of fundamental socio-psychological theories, such as social cognitive theory [4], social comparison theory [17], focus theory of normative conduct [13], social facilitation theory [20], cooperation theory [2], competition theory [15], and taxonomy of intrinsic motivators [25], this paper explored a list of seven social influence aspects and their interconnectedness. In achieving that, four empirical studies based on partial least squares structural equation modeling (PLS-SEM) approach were methodologically analyzed.

Main contributions of the meta-analysis include the reviewed background of social influence aspects and the originated framework structure for designing and evaluating socially influencing systems. These contributions supplement the existing body of knowledge and can be instrumental for scholars focusing on research related to persuasive engineering of socially influencing systems for behavior change.

For the future, socially influencing systems can open up new seamless and natural channels for businesses to engage with customers. These channels can potentially play a significant role in advancing customer relationships, as they enable immediate interaction at the place and time where customers acquire new experiences about certain product or services.

Presented research includes a limited number of empirical studies that were available for the current review. Thus, scholars are encouraged to conduct similar studies in order to extend the meta-analysis [38] and overall understanding of the role of social aspects in the typology of computer-supported influence [39].

References

1. Appleford, S., Bottum, J.R., Thatcher, J.B.: Understanding the Social Web: Towards Defining an Interdisciplinary Research Agenda for Information Systems. ACM SIGMIS Database 45(1), 29–37 (2014)
2. Axelrod, R.: On Six Advances in Cooperation Theory. Analyse & Kritik 22(1), 130–151 (2000)
3. Bandura, A.: Social Learning Theory. Prentice Hall, Englewood Cliffs (1977)
4. Bandura, A.: Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice Hall, Englewood Cliffs (1986)
5. Bandura, A.: Social cognitive theory of mass communication. Media Psychology 3(3), 265–299 (2001)
6. Baumeister, R.F.: The self. In: Gilbert, D.T., Fiske, S.T., Lindzey, G. (eds.) The Handbook of Social Psychology, pp. 680–740. McGraw–Hill, New York (1998)

7. Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., Javier Sese, F.: Using Clickers in Class. The Role of Interactivity, Active Collaborative Learning and Engagement in Learning Performance. Computers Science Education 13(2), 137–172 (2012)
8. Cacioppo, J.T., Petty, R.E., Stoltenberg, C.D.: Processes of Social Influence: The Elaboration Likelihood Model of Persuasion. In: Kendall, P.C. (ed.) Advances in Cognitive-Behavioral Research and Therapy, pp. 215–274. Academic Press, San Diego (1985)
9. Chatterjee, S., Price, A.: Healthy Living with Persuasive Technologies: Framework, Issues, and Challenges. Journal of the American Medical Informatics Association 16(2), 171–178 (2009)
10. Cheng, Y., Liang, J., Leung, L.: Social Network Service Use on Mobile Devices: An Examination of Gratifications, Civic Attitudes and Civic Engagement in China. New Media & Society (2014)
11. Chiu, C.M., Hsu, M.H., Wang, E.T.: Understanding Knowledge Sharing in Virtual Communities: An Integration of Social Capital and Social Cognitive Theories. Decision Support Systems 42(3), 1872–1888 (2006)
12. Cialdini, R.B.: Influence: The Psychology of Persuasion. HarperCollins e-books (2009)
13. Cialdini, R.B., Kallgren, C.A., Reno, R.R.: A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. Advances in Experimental Social Psychology 24(20), 1–243 (1991)
14. Crano, W.D., Prislin, R.: Attitudes and Persuasion. Annual Review of Psychology 57, 345–374 (2006)
15. Deutsch, M.: A Theory of Cooperation-Competition and Beyond. In: Handbook of Theories of Social Psychology, vol. 2, p. 275 (2011)
16. Deutsch, M., Gerard, H.B.: A Study of Normative and Informational Social Influences upon Individual Judgment. The Journal of Abnormal and Social Psychology 51(3), 629 (1955)
17. Festinger, L.: A Theory of Social Comparison Processes. Human Relations 7(2), 117–140 (1954)
18. Firpo, D., Kasemvilas, S., Ractham, P., Zhang, X.: Generating a Sense of Community in a Graduate Educational Setting through Persuasive Technology. In: 4th International Conference on Persuasive Technology, p. 41 (2009)
19. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann, San Francisco (2003)
20. Guerin, B., Innes, J.: Social Facilitation. Cambridge University Press, Cambridge (2009)
21. von Hippel, E.: Democratizing Innovation: The Evolving Phenomenon of User Innovation. International Journal of Innovation Science 1(1), 29–40 (2009)
22. Hsu, C.L., Lin, J.C.C.: Acceptance of Blog Usage: The Roles of Technology Acceptance, Social Influence and Knowledge Sharing Motivation. Information & Management 45(1), 65–74 (2008)
23. Lapinski, M.K., Rimal, R.N.: An Explication of Social Norms. Communication Theory 15(2), 127–147 (2005)
24. Loock, C.M., Staake, T., Landwehr, J.: Green IS Design and Energy Conservation: An Empirical Investigation of Social Normative Feedback. In: International Conference on Information Systems, p. 10 (2011)
25. Malone, T.W., Lepper, M.: Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning. In: Snow, R.E., Farr, M.J. (eds.) Aptitude, Learning and Instruction: III. Conative and Affective Process Analyses, pp. 223–253. Erlbaum, Hillsdale (1987)
26. Mangold, W.G., Faulds, D.J.: Social Media: The New Hybrid Element of the Promotion Mix. Business Horizons 52(4), 357–365 (2009)

27. May, M.A., Doob, L.W.: Cooperation and Competition. Social Science Research Council Bulletin, 125 (1937)
28. Mead, M.: Cooperation and Competition among Primitive Peoples. McGraw-Hill, New York (1937)
29. Moeller, S., Ciuchita, R., Mahr, D., Odekerken-Schröder, G., Fassnacht, M.: Uncovering Collaborative Value Creation Patterns and Establishing Corresponding Customer Roles. Journal of Service Research 16(4), 471–487 (2013)
30. Mumford, E.: A Socio-Technical Approach to Systems Design. Requirements Engineering 5(2), 125–133 (2000)
31. Mumm, J., Mutlu, B.: Designing Motivational Agents: The Role of Praise, Social Comparison, and Embodiment in Computer Feedback. Computers in Human Behavior 27(5), 1643–1650 (2011)
32. Nambisan, S., Baron, R.A.: Virtual Customer Environments: Testing a Model of Voluntary Participation in Value Co-creation Activities. Journal of Product Innovation Management 26(4), 388–406 (2009)
33. O'Keefe, D.J.: Theories of Persuasion. In: Nabi, R., Oliver, M.B. (eds.) Handbook of Media Processes and Effects. Sage Publications, Thousand Oaks (2009)
34. Prahalad, C.K., Ramaswamy, V.: The New Frontier of Experience Innovation. MIT Sloan Management Review 44(4), 12–18 (2003)
35. Rashotte, L.: Social Influence. The Blackwell Encyclopedia of Social Psychology 9, 562–563 (2007)
36. Ryan, R.M., Deci, E.L.: Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. American Psychologist 55(1), 68 (2000)
37. Schoenau-Fog, H.: Teaching Serious Issues through Player Engagement in an Interactive Experiential Learning Scenario. Eludamos, Journal for Computer Game Culture 6(1), 53–70 (2012)
38. Stibe, A.: Socially Influencing Systems: Persuading People to Engage with Publicly Displayed Twitter-based Systems. Acta Universitatis Ouluensis (2014)
39. Stibe, A.: Advancing Typology of Computer-Supported Influence: Moderation Effects in Socially Influencing Systems. In: MacTavish, T., Basapur, S. (eds.) Persuasive Technology. LNCS, vol. 9072, pp. 251–262. Springer, Heidelberg (2015)
40. Wills, T.A.: Downward Comparison Principles in Social Psychology. Psychological Bulletin 90(2), 245 (1981)
41. Wilson, S.R., Benner, L.A.: The Effects of Self-Esteem and Situation upon Comparison Choices During Ability Evaluation. Sociometry, 381–397 (1971)
42. Wood, J.V.: What is Social Comparison and How Should We Study It? Personality and Social Psychology Bulletin 22(5), 520–537 (1996)
43. Yerkes, R.M., Dodson, J.D.: The Relation of Strength of Stimulus to Rapidity of Habit-Formation. Journal of Comparative Neurology and Psychology 18(5), 459–482 (1908)
44. Zajonc, R.B.: Social Facilitation. Science 149, 269–274 (1965)

Advancing Typology of Computer-Supported Influence: Moderation Effects in Socially Influencing Systems

Agnis Stibe(✉)

MIT Media Lab, Cambridge, MA, USA
agnis@mit.edu

Abstract. Persuasive technologies are commonly engineered to change behavior and attitudes of users through persuasion and social influence without using coercion and deception. While earlier research has been extensively focused on exploring the concept of persuasion, the present theory-refining study aims to explain the role of social influence and its distinctive characteristics in the field of persuasive technology. Based on a list of notable differences, this study outlines how both persuasion and social influence can be best supported through computing systems and introduces a notion of computer-moderated influence, thus extending the influence typology. The novel type of influence tends to be more salient for socially influencing systems, which informs designers to be mindful when engineering such technologies. The study provides sharper conceptual representation of key terms in persuasive engineering, drafts a structured approach for better understanding of the influence typology, and presents how computers can be moderators of social influence.

Keywords: Influence typology · Computer-moderated · Persuasive technology · Computer-mediated · Computer-human · Socially influencing systems

1 Introduction

Persuasive technologies are commonly engineered to change behavior and attitudes of users through persuasion and social influence without using coercion and deception [9]. Both persuasion [22] and social influence [11] have been studied as concepts in behavioral, cognitive, and social psychology for long time. Evidently, they both exert capacity to alter human attitude and behavior, but each of them employs specific attributes to achieve that through face-to-face communication and presence in the physical world [3], [21].

While computers are becoming ubiquitous as tools, media, and social actors, it is necessary to clarify how the concepts of persuasion and social influence can be engineered in computing systems [9], [29]. More importantly, before designing such persuasive systems, scholars and practitioners should be aware of how each concept can be operationalized and what consequences each design component can bear [28].

According to Fogg [9], people can respond socially to computer products, which opens the door for social influence aspects [29] to exert their powers of motivating and persuading users. Thus, computers can be perceived as social entities or actors

T. MacTavish and S. Basapur (Eds.): PERSUASIVE 2015, LNCS 9072, pp. 253–264, 2015.
DOI: 10.1007/978-3-319-20306-5_23

that influence people on their own [28]. This happens when people make inferences about social presence [26] in persuasive technology through social cues, such as physical (face, eyes, body, movement), psychological (personality, similarity, feelings), language (spoken language, praise, language recognition), social dynamics (dialogues, reciprocity), and social roles (authority, doctor, teacher). Although it broadens understanding of computers as social actors, such discussion is focused on perceiving computers as individual entities with human-like characteristics rather than means for computer-supported influence that originates from other users [29].

To address this gap, the present research aims at clarifying the role of social influence [17], [25], [27], [33] and its distinctive qualities in the field of persuasive technology (Section 2). Based on these differences, this paper outlines how the concepts of persuasion and social influence can be best supported through computing systems (Section 3). Further, the notion of interpersonal computer-moderated influence is introduced, its place among other relevant types of persuasion is defined, and its specific role in the realm of persuasive technology is explained (Section 4). Lastly, the paper discusses implications of this research for scholars and designers of socially influencing systems (Section 5), and provides final conclusions (Section 6).

2 Socio-Psychological Foundation

Concepts of persuasion and social influence are often used interchangeably when describing a phenomenon of behavioral or attitudinal change that is caused by other people. Although persuasion and social influence can achieve the same goal of shaping human attitude and behavior, research in social psychology (Table 1) demonstrates that both concepts have notable differences in character and encompass distinct properties [14], [16].

According to Wood [31], *persuasion* typically includes detailed argumentation that is presented to people in a context with only minimal social interaction (e.g. one-to-one or one-to-many verbal persuasion), whereas *social influence* is usually enabled and facilitated by more complex social settings (e.g. many-to-one or many-to-many social contexts). O'Keefe [21] has argued that persuasion mainly relies on and is built upon reasoning and argument to shift attitudes and behavior of individuals towards a desired agenda, but social influence is commonly driven by the behavior and actions of surrounding people.

An additional perspective by Cialdini [3] has proposed that persuasion works by appealing to a set of deeply rooted human drives and needs, such as liking, reciprocity, consistency, authority, and scarcity. At the same time, humans look for social proof as a source of influence, and rely on the people around them for cues on how to think, feel, and act. In earlier work, Cialdini together with Goldstein [4] also claimed that social influence is a psychological phenomenon that often occurs in direct response to overt social forces. Finally, in recent collaborative work with Guadagno and Ewell [11], Cialdini specified that social influence refers to the changing of attitudes, beliefs, or behavior of an individual because of real or imagined external pressure.

Table 1. Persuasion and social influence in social psychology literature

Reference	Persuasion	Social Influence
Cialdini [3]	Works by appealing to a set of deeply rooted human drives and needs, such as liking, reciprocity, consistency, authority, and scarcity.	Humans look for social proof, therefor rely on the people around them for cues on how to think, feel, and act.
Guadagno et al. [11]		Refers to the changing of attitudes, beliefs, or behavior of an individual because of real or imagined external pressure.
O'Keefe [21]	Mainly relies on and is built upon reasoning and argument to shift attitudes and behavior of individuals towards a desired agenda.	Commonly driven by the behavior and actions of surrounding people.
Petty and Cacioppo [22]	Two basic routes to persuasion. One is based on the thoughtful consideration of arguments central to the issue, whereas the other is based on peripheral cues.	
Rashotte [23]	Focuses merely on written or spoken messages sent from source to recipient.	Defined as change in thoughts, feelings, attitudes, or behavior of an individual that results from interaction with another individual or a group.
Wood [31]	Typically includes detailed argumentation that is presented to people in a context with only minimal social interaction.	Usually enabled and facilitated by complex social settings.

Petty and Cacioppo [22] have argued that there are two basic routes to persuasion. One route is based on the thoughtful consideration of arguments central to the issue, whereas the other is based on peripheral cues in the persuasion situation. Rashotte [23] has defined social influence as change in thoughts, feelings, attitudes, or behavior of an individual that results from interaction with another individual or a group.

2.1 Persuasion

Persuasion is broadly defined as the action of causing someone to do something through reasoning or argument [21-22]. According to Rashotte [23], current research on persuasion focuses merely on written or spoken messages sent from source to recipient. This research is based on the assumption that people process messages carefully whenever they have motivation and ability to do so. Modern persuasion research is mainly dominated by studies employing either the elaboration likelihood model (ELM) [22] or heuristic-systemic models (HSM) [8].

2.2 Social Influence

Social influence is broadly defined as the capacity to have an effect on the behavior of someone in a social context. In general, social influence is naturally and instantly present in most social contexts of everyday life. According to earlier research [27], the study of social influence is central to social psychology and essential to understand group dynamics and intergroup relations. Historically, the research on social influence covers a broad range of topics, from persuasion and attitude change [31], to compliance and conformity [4], to collective action and social change [18]. Social influence is the process by which people really change their behavior depending on interaction with others who are perceived to be similar, desirable, or expert [23].

2.3 Understanding Distinctive Characteristics

Earlier discussion on persuasion and social influence creates an understanding that both paradigms are present in settings with two or more people that lead to behavioral or attitudinal changes in one or many of them. However, it is also important to clarify the distinctive characteristics of the two paradigms so that researchers and designers would be able to implement them in a proper way and study their effects on human behavior in a rigorous manner. Social psychology research on persuasion and social influence (Table 1) suggest numerous aspects that differentiate the two, therefore further discussion focuses only on the four main distinctive characteristics that are categorized in Table 2, i.e. the *origin*, the *driver*, the *impact*, and the *direction*.

Table 2. Distinctive characteristics of persuasion and social influence

	Persuasion	**Social Influence**
Origin	Intention or agenda	Presence of other people
Driver	Reasoning or argument	Behavior of surrounding people
Impact	Controlled and guided	Unpredictable and ambient
Direction	Push	Pull

Origin. Persuasion generally originates either from an *intention* to change an attitude and behavior of an individual or from a broader agenda of shaping what crowds of people think and do. In contrast, social influence effects occur and persist in the *presence* of other people around an individual.

Driver. According to earlier definitions [21], persuasion mainly relies on and is built upon *reasoning* and argument to shift attitudes and behavior of individuals towards a desired agenda, whereas social influence is commonly driven by the behavior and *actions* of surrounding people.

Impact. For persuasion to exert a desired impact on an individual through consistent reasoning and argumentation, it has to be performed in a *controlled* and guided manner. But, social influence primarily depends on the presence of other people and their behavior in a given social environment, therefore making its impact *unpredictable* and reliant on a particular context.

Direction. Prior research demonstrates that persuasion by definition operates as *push* mechanism that communicates an intended agenda with supportive arguments through guided approach, i.e. a persuader intentionally attempts to shapes the behavior and attitudes of receivers. Whereas in case of social influence, an individual is rather *picking* up an influence from a particular social context, i.e. individuals acquire sense of influence from surrounding people and their behavior.

3 Influence Typology

Computer-supported influence holds considerable promise as a topic of research [10]. Prior research in the realm of persuasive technology [9] has distinguished three relevant types of persuasion [13], i.e. interpersonal persuasion, computer-mediated persuasion, and human-computer persuasion. To advance this research area, the aforementioned types have been adjusted and are further discussed as: interpersonal *face-to-face (FTF)* influence, interpersonal *computer-mediated (CME)* influence, and *computer-human (CHU)* influence, respectively.

Based on the distinctive characteristics of persuasion and social influence (Table 2) and the ways in which both can be supported through computing systems, this paper outlines the existence of another type, namely interpersonal *computer-moderated (CMO)* influence, and explains its place and role within the realm of computer-supported influence (Fig. 1). More elaborate comparison of the four types of influence is provided in Table 3.

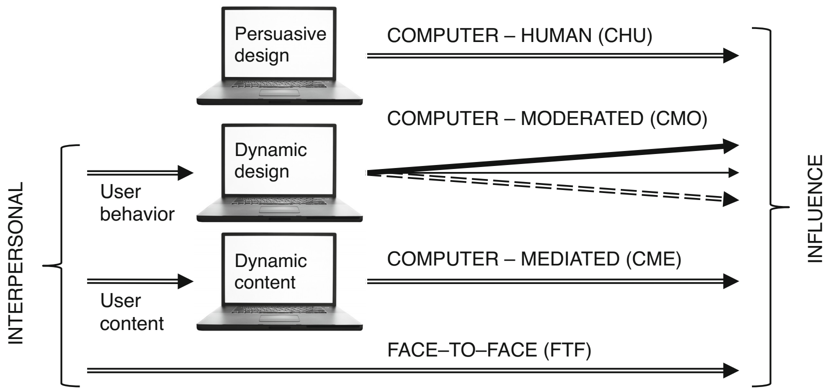


Fig. 1. Influence typology

3.1 Interpersonal Face-to-Face (FTF) Influence

According to Wilson [30], interpersonal influence can take place during an interaction of two or more people, involving verbal and non-verbal forms of behavior, personal feedback, and coherence of behavior. Further, this type of influence is termed as face-to-face (FTF) [20] to distinguish it from computer-supported influence (Fig. 1).

3.2 Interpersonal Computer-Mediated (CME) Influence

Interpersonal influence can also take place through various computing technologies, such as emails, mobile messaging, video chats, etc. In this case, the chosen technology serves as a mediator of interpersonal influence without any additional agenda to affect its users. Therefore, this type of influence is termed as interpersonal computer-mediated (CME) influence (Fig. 1) [11-12] and it can be well operationalized through fixed content (FC) and dynamic content (DC) components that are further explained in Section 4.1 and presented in Table 4.

Prior research also exposes that scholars have been active in studying interpersonal CME persuasion [15], [24], and its comparison to interpersonal FTF persuasion for many years. For example, Di Blasio and Milani [7] found that computer-mediated discussion could possibly activate the central route of persuasion [22] more easily than face-to-face interaction. This knowledge can be instrumental to explore more granular differences between the two types of influence.

Table 3. Comparing the four types of influence

	Interpersonal			
	Face-to-face (FTF)	**Computer-mediated (CME)**	**Computer-moderated (CMO)**	**Computer-human (CHU)**
Origin	Human	User	User behavior	Designer
Description	People can influence each other in the physical world.	Users can influence each other through computers.	Computers can amplify, decrease, or reverse influence based on the presence (or absence) of other users and their behavior.	Computers can influence users when designed to do so.

3.3 Computer-Human (CHU) Influence

Computer-human (CHU) influence is very different from both types of interpersonal influence previously discussed, i.e. FTF and CME, because it is based on the notion that computers can be designed to perform the role of social actors [9], and thus they can have capacity to influence users independently of interpersonal relationships with other users (Fig. 1). For this reason, the CHU influence can be better operationalized through the fixed design (FD) component, which is described in Section 4.1 and presented in Table 4.

Earlier research provides different collections of techniques and principles that can be useful for designing and evaluating persuasive technologies. For instance, the behavior change technique taxonomy contains 93 hierarchically clustered techniques to build an international consensus for the reporting of behavior change interventions [19]. According to Fogg [9], there are various theory-driven persuasive principles that can be incorporated into the design of computers to improve their persuasiveness. Many principles and techniques by definition and design can support the CHU influence, but not all of them. Those principles that are primarily dependent on behavior of other users rather fall under the interpersonal computer-moderated (CMO) influence, as discussed further in the next section.

3.4 Interpersonal Computer-Moderated (CMO) Influence

Interpersonal computer-moderated (CMO) influence is distinct from all other types of influence described above with its unique characteristic of being able to amplify, decrease, or reverse the persuasion effect through computing technology depending on the presence (or absence) of other users and their behavior.

Interpersonal CMO and CME influences differ, because the latter serves as a mediator without affecting interpersonal influence, while the effects of the former can fluctuate depending on the actual behavior of other concurrent users. In other words, the role of a computer in the interpersonal CME influence is mainly to mediate interpersonal persuasion, whereas in case of the interpersonal CMO influence, the role of a computer is to facilitate the effects of social influence though the dynamic design (DD) component, further described in Section 4.1 and exhibited in Table 4.

As it can be observed from Fig. 1, interpersonal CMO influence also substantially differs from CHU influence, because the latter is not supposed to receive any input from other users, thus the CHU influence is solely based on the intentions that its designers have preset in the interfaces of the computing technology (the FD component). Another way to better understand the nature of the interpersonal CMO influence and how it differs from the interpersonal CME influence is to think about the common attributes of moderation and mediation in social psychological research [2].

4 Computer-Supported Influence

Computing technologies increasingly penetrate various aspects of everyday life. This advancement continuously expands ways of how people can be reached, thus experience persuasion or social influence through human-computer interaction [10] and computer-mediated communication [10].

According to Wilson [30], communication via computer is intrinsically less suitable for persuading as compared to face-to-face interaction, because of deficiencies to transmit non-verbal cues and limited number of utilizable strategies. At the same time, computers can be designed to play the role of a social actor [9], which means that they are capable not only to mediate persuasive communication but also support persuasion and social influence through intentionally designed computer software and interfaces [13].

4.1 Components

Before designing persuasive technologies [9] and socially influencing systems [29], it is very important to understand the main components of how computers can support influence (Table 4). The two main components of computing systems, which are directly exposed to users through interfaces, are *content* (e.g., texts, photos, sounds, videos) and *design* (e.g., layout, navigation, colors, features). Both components can be operationalized either as *fixed* or *dynamic*.

Table 4. Components of computer-supported influence

	Content	**Design**
Fixed	(FC) Preset by developers and owners Supports CHU influence	(FD) Preset by designers Supports CHU influence
Dynamic	(DC) Generated by users Supports interpersonal CME influence	(DD) Evolving through user behavior Supports interpersonal CMO influence

Historically, computer systems were often built with fixed design that was preset by designers and fixed content that was predefined by system developers and owners. With the overall technological advancement, computer systems are becoming more social and dynamic by both allowing users to contribute own content and displaying their interactions with the systems.

Through clearer understanding of the four components, the designers of computer systems become better equipped with ways of how both persuasion and social influence can be operationalized more effectively.

Based on the distinctive characteristics of persuasion and social influence (Table 1), the likelihood of support for both concepts was assessed and reported for each component in Table 4. That is, if an intention is to persuade users through computing systems, then the *fixed content (FC)*, *fixed designed (FD)*, and *dynamic content (DC)* components are suitable in achieving that. However, if an aim is to leverage social influence through socially influencing systems [29], then the DC and *dynamic design (DD)* components can yield favorable results.

In this case, if a system is implemented with fixed design and fixed content so that users can see only outcomes of their own actions, then the chances for social influence to play a role in the given context are very limited. Of course, the fixed components can contain preset messages conveying social influence aspects, e.g. social normative statements [5], but their effects can decline over time, as they do not change. Nevertheless, dynamic content and dynamic design expand user interaction and enable them to see what others are doing. In that way, both dynamic components open up multiple ways for social influence to occur and affect users.

4.2 Operationalization

The concepts of computer-supported influence can be operationalized in many ways depending on a given context and intended behavior change. To give an example, imagine a situation where a person is concerned about his health conditions and has decided to exercise more by jogging each morning. As part of this plan, the person installs a mobile application intentionally designed to help achieve the target behavior change. First, the mobile app enables a jogger to check weather conditions, and secondly, enables users to see how many others are jogging at that moment (Fig. 2). The counter of joggers is an operationalization of the DD (dynamic design) component from Table 4, as it purely depends on the behavior of other users.

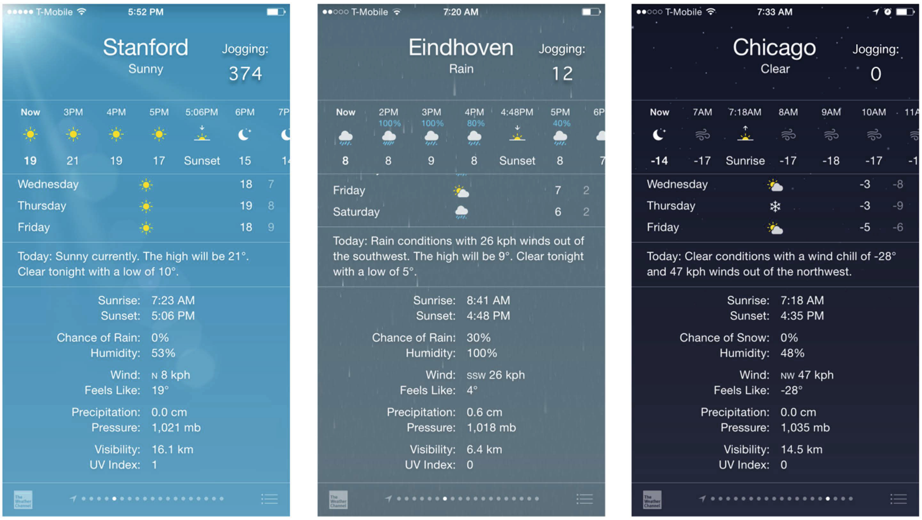


Fig. 2. Example of an interpersonal computer-moderated influence: on the left, nice weather in Stanford and 374 are jogging; in the middle, bad weather in Eindhoven and only twelve are jogging; on the right, cold and windy weather in Chicago and no one is jogging outside

If the weather conditions are great in Stanford and the counter shows that 374 people are currently jogging outside (Fig. 2, left), then the user would experience increased motivation to go out and exercise together with others. Instead, the heavy rain and the comparative low number of joggers on the streets and in the parks of Eindhoven (Fig. 2, middle) would most likely to decrease the motivation of a user to step outside. Now, imagine a situation when there is one extremely cold and windy morning in Chicago (Fig. 2, right). The alarm clock rings, the user opens the application and notices the bad weather conditions, which naturally affects the motivation for jogging that morning. Now what? The user looks at the number of others jogging at that exact moment. Quite simply, a zero joggers in the picture would discourage the individual from jogging that morning, a small number would make the user hesitant, but a large number of other joggers would still give an extra boost to the motivation.

This example demonstrates how the behavior of other users can increase, decrease, or reverse the persuasive effect of a mobile application that is complemented with the design principle of social facilitation [32], which represents social influence. In a similar manner, it can be easily illustrated how other social influence design principles would end up having the same pattern. The competition principle [6] that is implemented as a top, for example, would amplify its persuasive potential only as long as an individual has competitive position among other users. Whenever the individual falls behind the competition, this principle naturally loses its capacity to influence.

5 Discussion

This theory-refining research highlights the importance and necessity to continue studying various facets of persuasion and social influence in the realm of persuasive technology. The substance of this paper demonstrates that both concepts maintain distinctive qualities, and therefore their nature has to be better understood before making an attempt to design and implement them in computing technologies.

The contribution of this paper is fourfold. First, the paper provides a comparison of persuasion and social influence that clarifies the nature of both concepts in a structured manner. Second, the paper outlines four components of how computers can support persuasion [21-22] and social influence [11], [23]. Third, the influence typology is presented and, forth, extended with an introduction of interpersonal computer-moderated (CMO) influence.

Overall, the outcome of this research effort demonstrates that there are various ways that persuasion and social influence can be facilitated through computing technologies, but a positive effect is not always guaranteed. In the case of the interpersonal CMO influence, intended effects can be amplified, decreased, or reversed depending on the presence (or absence) of other users and their behavior.

5.1 Implications for Designers

The designers of persuasive technologies should be very careful when designing interpersonal computer-moderated (CMO) influence, which is mainly about implementing aspects of social influence. In order to avoid possibly negative effects of the interpersonal CMO influence, the designers of persuasive technologies can and oftentimes should incorporate specific rules and triggers to control for the likelihood of unwanted effects occurring. If such control mechanisms were in place, another kind of an implementation could be deployed as long as necessary. For example, when the number of joggers on the streets of Chicago (Fig. 2) drops below twenty, instead of reporting a low number, the mobile app can show an average number of joggers at that time of day which is aggregated over the last month or over ten other days with similar weather conditions.

5.2 Future Research

This research provides additional evidence that the theoretical work on persuasive technologies and socially influencing systems has potential for further research initiatives. In the next steps, each aspect of social influence has to be further studied separately and rigorously in line with related theories from social psychology. Then these aspects need to be designed, implemented, and tested to assess thresholds of when interpersonal computer-moderated (CMO) influence begins to shift its effect from amplifying to decreasing and from decreasing to reversing. Conducting such studies is highly important, as they would contribute to more detailed understanding of how and when socially influencing systems [29] are gaining, losing, or reversing their capacity to affect user involvement, participation, and engagement [28].

6 Conclusions

The present study explained the role of social influence and its distinctive characteristics in the field of persuasive technology. Based on the unique differences between persuasion and social influence, this paper described ways of how both concepts can be best supported through computing systems.

The study introduced the notion of interpersonal computer-moderated (CMO) influence and defined its place within the influence typology. Compared to the other types, the CMO influence firmly relies on four distinguishing characteristics of social influence, namely origin, driver, impact, and direction. By definition, the CMO influence can amplify, decrease, or reverse an intended effect on users, therefore designers of socially influencing systems [28] should be mindful when engineering them.

To summarize, this research outlined a sharper conceptual representation of the key terms in persuasive engineering, drafted a structured approach for better understanding of the influence typology, and presented how computers can be moderators of social influence. Consequently, future research attempts can be directed towards formalizing and operationalizing the influence typology, and advancing the methodology for socially influencing systems [29].

References

1. Angst, C.M., Agarwal, R.: Adoption of Electronic Health Records in the Presence of Privacy Concerns: the Elaboration Likelihood Model and Individual Persuasion. MIS Quarterly 33(2), 339–370 (2009)
2. Baron, R.M., Kenny, D.A.: The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. Journal of Personality and Social Psychology 51(6), 1173 (1986)
3. Cialdini, R.B.: Influence: The Science of Persuasion. HarperCollins Publishers Inc., New York (2009)
4. Cialdini, R.B., Goldstein, N.J.: Social Influence: Compliance and Conformity. Annu. Rev. Psychol. 55, 591–621 (2004)
5. Cialdini, R.B., Kallgren, C.A., Reno, R.R.: A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. Advances in Experimental Social Psychology 24(20), 1–243 (1991)
6. Deutsch, M., Gerard, H.B.: A Study of Normative and Informational Social Influences upon Individual Judgment. Journal of Abnormal and Social Psychology 51(3), 629 (1955)
7. Di Blasio, P., Milani, L.: Computer-Mediated Communication and Persuasion: Peripheral vs. Central Route to Opinion Shift. Computers in Human Behavior 24(3), 798–815 (2008)
8. Eagly, A.H., Chaiken, S.: The Psychology of Attitudes. Harcourt, New York (1993)
9. Fogg, B.J.: Persuasive Technology: Using Computers To Change What We Think And Do. Morgan Kaufmann, San Francisco (2003)
10. Gass, R.H., Seiter, J.S.: Persuasion: Social Influence, and Compliance Gaining, 5th edn. Pearson/Allyn & Bacon, Boston (2013)
11. Guadagno, R.E., Ewell, P.J., Cialdini, R.B.: Influence. In: Cooper, C.L. (ed.) Wiley Encyclopedia of Management, pp. 3–5. John Wiley & Sons (2014)

12. Guadagno, R.E., Muscanell, N.L., Rice, L.M., Roberts, N.: Social Influence Online: The Impact of Social Validation and Likability on Compliance. Psychology of Popular Media Culture 2(1), 51 (2013)
13. Harjumaa, M., Oinas-Kukkonen, H.: Persuasion Theories and IT Design. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 311–314. Springer, Heidelberg (2007)
14. Haslam, S.A., McGarty, C., Turner, J.C.: Salient Group Memberships and Persuasion: The Role of Social Identity in the Validation of Beliefs (1996)
15. Hong, S., Park, H.S.: Computer-Mediated Persuasion in Online Reviews: Statistical Versus Narrative Evidence. Computers in Human Behavior 28(3), 906–919 (2012)
16. Hovland, C.I., Janis, I.L., Kelley, H.H.: Communication and Persuasion. Psychological Studies of Opinion Change (1953)
17. Kiesler, S., Siegel, J., McGuire, T.W.: Social Psychological Aspects of Computer-Mediated Communication. American Psychologist 39(10), 1123 (1984)
18. Lewin, K.: Group Decision and Social Change. In: Newcomb, T.M., Hartley, E.L. (eds.) Readings in Social Psychology, pp. 330–344. Holt, Rinehart, and Winston, NY (1947)
19. Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M.P., Cane, J., Wood, C.E.: The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. Annals of Behavioral Medicine 46(1), 81–95 (2013)
20. O'Keefe, B.J., Shepherd, G.J.: The Pursuit of Multiple Objectives in Face-to-Face Persuasive Interactions: Effects of Construct Differentiation on Message Organization. Communications Monographs 54(4), 396–419 (1987)
21. O'Keefe, D.J.: Persuasion: Theory and Research. Sage, Newbury (1990)
22. Petty, R.E., Cacioppo, J.T.: The Elaboration Likelihood Model of Persuasion. Advances in Experimental Social Psychology 19, 123–205 (1986)
23. Rashotte, L.: Social influence. The Blackwell Encyclopedia of Social Psychology 9, 562–563 (2007)
24. Sassenberg, K., Boos, M., Rabung, S.: Attitude Change in Face-to-Face and Computer-Mediated Communication: Private Self-Awareness as Mediator and Moderator. European Journal of Social Psychology 35(3), 361–374 (2005)
25. Sassenberg, K., Ionas, K.I.: Attitude Change and Social Influence. In: Oxford Handbook of Internet Psychology, 273 (2007)
26. Short, J.A., Williams, E., Christie, B.: The social psychology of telecommunications. Wiley, London (1976)
27. Smith, J.R., Louis, W.R., Schultz, P.W.: Introduction Social influence in action. Group Processes & Intergroup Relations 14(5), 599–603 (2011)
28. Stibe, A.: Socially Influencing Systems: Persuading People to Engage with Publicly Displayed Twitter-based Systems. Acta Universitatis Ouluensis (2014)
29. Stibe, A.: Towards a Framework for Socially Influencing Systems: Meta-Analysis of Four PLS-SEM Based Studies. In: MacTavish, T., Basapur, S. (eds.) Persuasive Technology. LNCS, vol. 9072, pp. 171–182. Springer, Heidelberg (2015)
30. Wilson, E.V.: Perceived Effectiveness of Interpersonal Persuasion Strategies in Computer-Mediated Communication. Computers in Human Behavior 19(5), 537–552 (2003)
31. Wood, W.: Attitude Change: Persuasion and Social Influence. Annual Review of Psychology 51(1), 539–570 (2000)
32. Zajonc, R.B.: Social facilitation. Science 149, 269–274 (1965)
33. Zimbardo, P.G., Leippe, M.R.: The Psychology of Attitude Change and Social Influence. Mcgraw-Hill Book Company (1991)

What Makes You Bike? Exploring Persuasive Strategies to Encourage Low-Energy Mobility

Matthias Wunsch[1,3(✉)], Agnis Stibe[2], Alexandra Millonig[1], Stefan Seer[1], Chengzhen Dai[2], Katja Schechtner[2], and Ryan C.C. Chin[2]

[1]Austrian Institute of Technology, Vienna, Austria
{Matthias.Wunsch.fl,Alexandra.Millonig,Stefan.Seer}@ait.ac.at
[2]MIT Media Lab, Cambridge, MA, USA
{agnis,chengdai,katjas,rchin}@mit.edu
[3]Human Computer Interaction, Vienna University of Technology, Vienna, Austria

Abstract. This paper explores three persuasive strategies and their capacity to encourage biking as a low-energy mode of transportation. The strategies were designed based on: (I) triggering messages that harness social influence to facilitate more frequent biking, (II) a virtual bike tutorial to increase biker's self-efficacy for urban biking, and (III) an arranged bike ride to help less experienced bikers overcome initial barriers towards biking. The potential of these strategies was examined based on self-reported trip data from 44 participants over a period of four weeks, questionnaires, and qualitative interviews. Strategy I showed a significant increase of 13.5 percentage points in share of biking during the intervention, strategy II indicated an increase of perceived self-efficacy for non-routine bikers, and strategy III provided participants with a positive experience of urban biking. The explored strategies contribute to further research on the design and implementation of persuasive technologies in the field of mobility.

Keywords: Low-energy mobility · Persuasion · Biking · Cycling · Behavior change · Transportation · Sustainability · Socially influencing systems

1 Introduction

Cities around the world are growing at an unprecedented pace, creating a manifold of new opportunities to meet and exchange ideas and goods. At the same time, however, they generate more traffic. Creating a transport system that supports high-quality life in urban areas requires shifting from high-energy modes of transportation, such as private cars or even public transport, to sustainable low-energy urban mobility, such as walking and biking [21]. Doing so reduces emissions of greenhouse gases, provides health benefits, and enhances the quality of urban life. However, an adoption of new modes of transportation requires a substantial behavior change [11]. Beyond hard policy measures, persuasive strategies embedded in technologies can be useful in facilitating such behavioral change [8],[12],[19]. The aim of this work is to explore such strategies and suitable technologies to promote sustainable low-energy mobility.

T. MacTavish and S. Basapur (Eds.): PERSUASIVE 2015, LNCS 9072, pp. 53–64, 2015.
DOI: 10.1007/978-3-319-20306-5_5

A promising low-energy mode for urban mobility is biking, as it is easily accessible, fast, low-cost, and uses less space than most other modes of transportation. Although previous work has covered mode choices and bike use, little attention has been paid to a change of choice from high-energy modes to biking as a low-energy mode and how this can be supported by persuasive technologies. Heinen et al. [14] classified five groups of determinants for bike commuting. Amongst them are psychological factors; these are attitudes, perceived social norms and habits, which can be at the center of persuasive strategies. Gatterslaben & Appleton [13] applied the transtheoretical model of behavior change [18] to bike commuting. Their findings suggest that different strategies are needed depending on current attitudes and behavior of individuals. Froehlich et al. [9] developed a mobile phone application that semi-automatically sensed and revealed information about transportation behavior. In combination with a personal ambient display, the app engaged users with the goal of increasing green transportation choices (e.g. walking, biking, public transport). Although some statements from qualitative interviews indicated the willingness for such change, no evaluation of actual change in mobility behavior was conducted. A similar but more recent study by Gabrielli and Maimome [10] examined the effect of a mobile app on supporting eco transport choices by citizens of an urban area. The transport choices and habits of the participants were influenced with several persuasion strategies and an overall increase of sustainable transport choices of 14%, as well as a higher environmental awareness among participants, was observed. However, the study design did not include a control group to better attribute behavior change to the experimental intervention. Even more recently, Flüchter et al. [7] found a positive impact of social normative feedback on e-bike commuting.

In accordance with the literature [14], a preliminary survey conducted at the beginning of this research showed that safety concerns are one of the main barriers for adapting biking as a regular mode of transportation. Therefore, the strategies in this study were designed with a focus on perceived safety of biking. All study participants were given access to bikes in order to prevent issues with bike availability and to concentrate research on motivational aspects.

The research question tackled in this paper is: What types of persuasive strategies can lead to a modal shift towards low-energy mobility by increasing bike use? Three different strategies were designed and evaluated in a pretest-posttest control group experimental design.

Section 2 presents the developed persuasive strategies. Section 3 describes the data collection and data analysis. Results are shown in section 4 and discussed in section 5. The paper ends with a conclusion and an outlook towards future research in section 6.

2 Deployed Persuasive Strategies

We designed and developed three strategies for this study.

Strategy I: Frequent Biking Challenge
In this strategy, the following principles of persuasion (see also [6], [8], [20]) have been combined: triggering, recognition, competition, cooperation, and comparison. The overall hypothesis is that this strategy increases bike use.

Triggering. Participants received emails (Fig. 1) between 3 to 5 times a week, providing them with information about their performance in the challenge and acting as a trigger for biking [8]. Emails were chosen as they are likely to be regularly read as opposed to a webpage or a mobile app providing the same information. They were sent in the evening to influence mobility planning for the next day. The regular email updates also contained a set of notifications tailored to each participant, such as daily weather forecasts and entertaining elements. The purpose of these notifications was to keep the sent emails useful and engaging for the participants. Additionally, the emails provided motivational facts about biking and suggestions on when to use a bike.

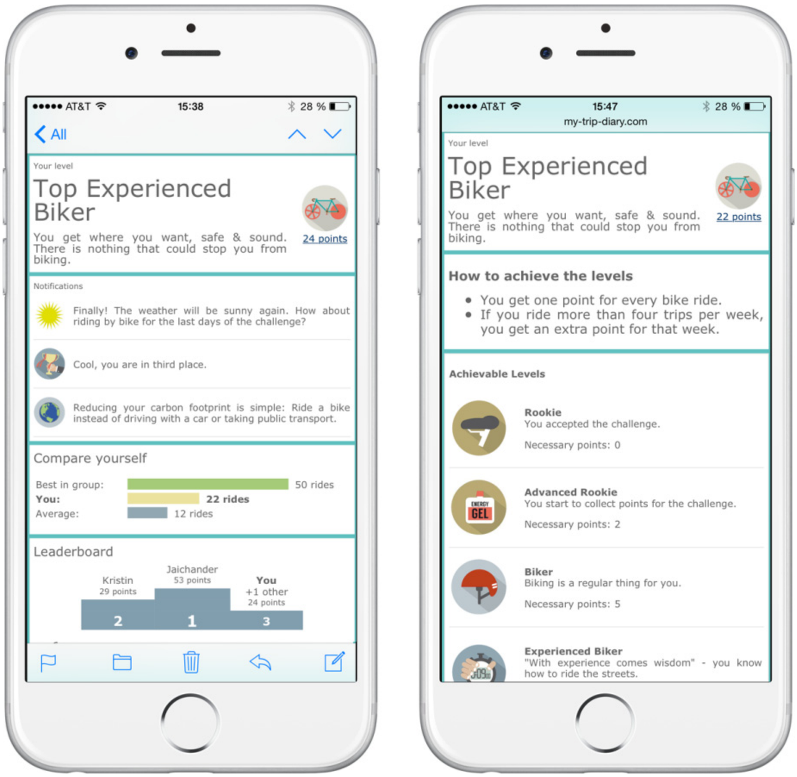


Fig. 1. Left: Regularly sent email updates: Notifications, comparison chart and leaderboard. Right: Explanation of the point scheme and achievable levels within the Frequent Biking Challenge.

Recognition. Based on the number of reported bike trips, participants received points and were awarded different statuses depending on the total number of points. These status levels had titles, were visualized with images and had an exploratory slogan. For example, participants achieving 5 points were recognized with the status "Experienced Biker" and the slogan: "With experience comes wisdom. You know how to ride the streets." Such recognition typically increases enjoyment [4] and influences future behavior [15].

Competition. The email updates furthermore included a leaderboard, showing one's own rank based on the achieved points in comparison to the other participants of the group. It was visualized with a podium for places 1, 2 and 3, followed by a list of the other ranks. Such salient metrics for people to observe their performances among other participants typically promotes competition, which consequently influences their thoughts and behavior [15].

Cooperation. At start, a collective goal (achieving 100 points collectively) was included in the email to facilitate cooperation among participants [15]. This was visualized with a bar graph that showed the sum of points from all participants and how much more were needed to reach the collective goal. The collective goal was reached in the second week of the challenge. Four days later it was replaced with the "compare yourself" comparison chart.

Social Comparison. The "compare yourself" design element allowed participants to compare their number of their bike rides to the average of bike rides and the best participant within the group. This possibly influences motivation as people tend to look for self-enhancement [22] and self-improvement [6],[23].

Strategy II: Virtual Bike Tutorial

The concept of perceived self-efficacy "is concerned with judgments of how well one can execute courses of action required to deal with prospective situations" [3]. Prior studies, such as Chittaro [5], used a persuasive game to increase the perceived self-efficacy[1] of passengers in the situation of an aircraft accident. In this study, the concept of perceived self-efficacy was used in relation to perceived risk and safety, thereby assessing how users perceive their control over their own safety in a biking context. The related assumption is that an increased self-efficacy towards biking will help to overcome safety barriers and hence encourage more biking.

Participants were provided with a short video tutorial on safe urban biking. The safety related information is based on safety guidelines from city officials from New York City, Boston and Vienna. The core concept of the training session is based on the content of a city biking school program. An expert-interview with an experienced biking instructor was conducted in order to gain knowledge on how biking in the city can be taught most effectively to novice bikers.

After the tutorial, a participant should experience the effects of different biking-related decisions in an interactive video training session. The procedure started with a first-person-view video where the participant saw a typical biking scene. The video was then stopped and the participants had to decide on how to continue the ride. (Fig. 2) The consequences of each possible decision were shown in a subsequent video. Different real-life scenarios (e.g. conflict with pedestrian) were tested and partici-

[1] Chittaro referred to it as "safety locus of control". See Ajzen [1] for a discussion on the difference between these concepts.

pants could learn about the consequences of their decisions. An increase of perceived self-efficacy due to that intervention was expected. To measure that, a self-efficacy in biking questionnaire[2] had to be completed by the participant before and after the video training session. The same questionnaire was also included in the survey at the end of the experimental period.

Fig. 2. Screenshot of the Virtual Bike Tutorial showing a conflict situation with a pedestrian

Strategy III: Bike Buddy Program

The start of a new physical exercise is often supported by an experienced person such that guidance and training is provided. In order to apply this kind of learning to the biking context, participants received a one time "bike buddy experience". The hypothesis in this regard is that for novice bikers, the experience of biking in an urban environment will change the perceived safety and risk of doing so. It was expected that this would lead to more positive attitudes towards biking and an overall increase of biking within the participants.

Bike buddies were recruited out of the potential participants for this study who were regular bikers and comfortable biking with new bikers. Bike buddies and participants were matched based on where they live and what routes they usually take. The bike buddies furthermore received instructions for the ride, covering safety aspects and clarifying the goal of showing the participant a safe and enjoyable biking route. They therefore were asked to find a safe and easy route for the planned bike ride and preferable inspect this route prior to the ride. They were also requested to set up a meeting point (ideally at the participant's home) for conducting the ride.

Several persuasive principles were implemented [6],[8]. ***Authority***, by having the bike buddy as a guide for the bike ride. ***Reduction***, by reducing the effort of the user to find a safe route (complex behavior) in the city to a simple behavior (follow the bike buddy). ***Tunneling***, by guiding participants along the route and allowing them to

---

[2] Items were adapted from a self-efficacy in driving questionnaire. [2]

experience the potential benefits of biking. Finally, ***tailoring***, by providing tailored information and personalized support to the user.

3 Data Collection and Analysis

The experiment took place in Cambridge/Boston, Massachusetts area over the period of four weeks in October 2014. A sample of 44 participants continually reported their trip data on a daily basis.

3.1 Sample

Study participants were recruited primarily through mailing lists at the Massachusetts Institute of Technology (MIT). The ideal participants were non-routine bikers (biking not more than three times a week). 55 participants met that requirement and were randomly assigned to one of three experimental groups or the control group. Typical route distance was not included as selection criteria, but as potential participants knew that they would join a biking related study it is likely that people with longer routine routes were less prone to join.

Participants were primarily part of the MIT community. Students made up a large portion of the sample. Therefore, the study sample is not representative of a broader population. Furthermore, the sample most likely exhibits self-selection bias. The process by which participants were recruited encourages those who want to bike, but do not have the means to do so, to join.

44 participants reported their trip data continuously over the period of four weeks. Group sizes were n=12 for (I) Frequent Biking Challenge, n=11 for (II) Virtual Bike Tutorial, n=11 for (III) Bike Buddy Program and n=10 for the control group. Out of all participants, 33 had no access to a bike and were provided with a one-month local bike sharing scheme subscription. 24 participants were provided with a helmet. As prior research shows that there are significant gender differences regarding utilitarian bike use [14], the sample should be balanced in terms of gender. The 44 participants that continually reported their trip data consisted of 22 women and 22 men.

3.2 Data Collection

Participants reported their trips on a daily basis. The collected mobility data included trip purpose and used mode(s). Participants were provided with a web-application that sent the data to a webserver with a relational database. To get continuous trip data, participants were automatically reminded via email in case they forgot to input their trips for the day. The trip diary included a calendar to navigate through the days, a help section, and a statistics graphic where users could see the amount of reported trips and how they were distributed among different modes of transportation. A settings section allowed the users to set a time for the daily reminders, put in custom trip purposes and to set their time zone.

Online questionnaires were used to measure perceived risk and perceived safety in biking at the beginning and end of the experimental period. Open questions were also included at the final questionnaire to ask for perceived behavior change. Interviews were conducted in order to gain further insight on the effect of the strategies. Six participants and six bike buddies from the Bike Buddy Program, two participants of each the Frequent Biking Challenge, the Virtual Bike Tutorial and the control group agreed to be interviewed after the experimental period.

3.3 Data Analysis

Analysis of Quantitative Trip Data. Based on the self-reported mobility data the modal split between modes was computed per person per day. To correct for bad weather, all days with precipitation above average were excluded from the analysis.[3] As can be seen in (1), the difference between the daily bike share of each participant of an experimental group $y_{g,d}$ and the mean of daily bike share within the control group $\bar{y}_{c,d}$ was computed for each day. The sum of these daily differences was divided by the number of days N_{pre} before or N_{post} after the (start of the) intervention.

$$z_{g,pre} = \frac{1}{N_{pre}} \sum_{d=1}^{N_{pre}} \left(y_{g,d} - \bar{y}_{c,d} \right) \; , \; z_{g,post} = \frac{1}{N_{post}} \sum_{d=1}^{N_{post}} \left(y_{g,d} - \bar{y}_{c,d} \right) \tag{1}$$

As can be seen in (1), the result is a value for average bike-share above control per participant before the intervention $z_{g,pre}$ and after the (start of the) intervention $z_{g,post}$. This approach provides a per-day correction of data which is more accurate than just comparing uncorrected per participant pre- and post-intervention mean values between experimental and control group.[4] Based on the computed values a one-sided paired sample t-test[5] was used to test the hypothesis of an increase in bike-share above the control group value.

As an indicator for the dependence between the share of biking and the share of high-energy modes Pearson r correlations have been computed at per participant level. To test for a difference in the means of perceived risk and perceived safety scores, a paired sample t-test was conducted.

Analysis of Qualitative Interview and Questionnaire Data. Qualitative content analysis [16] was used to analyze data obtained through ex-post-interviews and open question surveys. Category application was carried out in a deductive way, with aspects of analysis based on existing theoretical and empirical work.

4 Results

The effect of the presented strategies on actual bike use has to be viewed in light of many other factors with influence in this regard. The analysis of the gathered qualita-

---

[3] Weather data from NOAA [0] was used for that. Average precipitation was 4.8 mm.

[4] For that reason, the common method for pre post control designs of ANCOVA (analysis of covariance) was not applied.

[5] Shapiro-Wilk tests were conducted prior to all t-tests to check for normal distribution.

tive data showed that good biking infrastructure, such as protected bike lanes, makes biking more attractive and is perceived as safe. Knowing a route with good cycling infrastructure or otherwise comfortable interaction with motorized traffic helped participants to bike. Travel distance and difference in travel time compared to other possible modes of transportation played another crucial role. The analysis of the interviews suggests, that participants who could gain significant time savings by taking a bike instead of walking or using public transportation were more motivated to bike. Financial aspects like the upfront costs of buying a bike or the cost of using a bike sharing scheme were also taken into consideration, especially by the financially constrained study participants. For actual day to day bike use, situational factors such as weather, having to wear elegant clothing or having a lot to carry was reported as influential for the decision on whether or not to bike.

The reported mobility data showed an overall increase of bike trips that was mainly fueled by the participants that were provided with access to bikes. Participants with positive experiences while trying out biking subsequently considered buying a bike in the future. The study-participation and especially the use of the trip diary raised general awareness of biking and the experimental period was described as a time of personal reflection on mobility. One participant mentioned: “I also now consider what form or transportation I take before I take it because the trip diary made me consider the different forms of transportation.” Another one reported that she wanted to show that she is able to bike more. A self-monitoring effect was also reported by other participants. One told that he was regularly checking the provided statistics in the trip diary out of curiosity to see his personal statistics. But thinking actively about possible modes of transportation made participants also more aware of problems associated with urban biking: “I watched people get checked by car doors all the time and other bikers not obey lights or pedestrians”.

As for perceived safety and risk, the hypothesis has been that for non-routine bikers the experience of biking in an urban environment will lead to an increase in perceived safety and decrease in perceived risk of doing so. However, comparing the scores of these two variables for the beginning and end of the experimental period did not show a change.

4.1 Mode Shifts

The analysis on an individual level provides an overview on how the share of modes shifted. A change in mobility patterns towards more biking could be rooted in a decrease in use of high-energy modes (car and public transportation), but could also stem from a decrease in walking. The former is of special interest for this research. Pearson r correlations have been calculated as a basic indicator for the dependence of mode share over the four weeks in which the mobility data was recorded. These correlations show a statistically significant ($p<.05$) negative dependency of bike use and use of high-energy modes for at least 16 out of 44 (36%) participants, ranging from $r=-.97$ to $r=-.40$. As can be expected there are also statistically significant ($p<.05$) negative correlations between bike use and walking for 13 out of 44 (30%) participants, ranging from $r=-.94$ to $r=-.41$.

4.2 Strategy I: Frequent Biking Challenge

As shown in Fig 3, an increase in bike share occurred. It rose from 2% in week one to 15% in week two to 33% in week three. Week four showed a decrease to 23%. The change from pre- to during-intervention values of bike share was 13.5 percentage points[6] above the control group at statistically significant levels (p=0.03). Results based on interviews show that the constant reminders of possible mobility choices and the trip diary, which separated mobility patterns into several smaller parts, helped a participant "break the prospect of biking to/from down into achievable goals for myself, e.g. bike from home to the train station or from the train to work […]" The daily reminders were described as interesting and funny. The interviews also inform that cooperation and competition could have been more effective when involving social ties to other participants in the group.

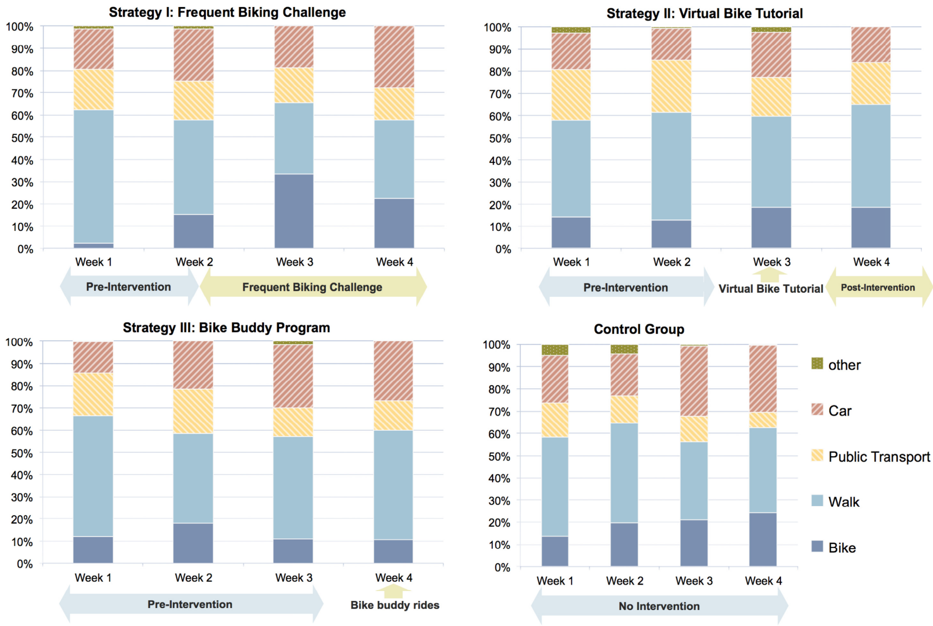


Fig. 3. Modal split during the experimental period

4.3 Strategy II: Virtual Bike Tutorial

The participants in the self-efficacy group conducted the tutorial in weeks two and three of this study. The results demonstrate an increase of biking share within all trips after the intervention from about 14% in week one and two to 19% in weeks three and four. When compared to control group shares, the change in bike use is not statistically significant for this strategy.

---

6 These values refer to percentage points within the modal split.

Self-efficacy of participants that reported lower levels at the beginning of the intervention showed a slight increase. (Fig. 4) However, on average, no clear rise in perceived biking self-efficacy emerged. In line with that, the conducted interviews suggest that the tutorial content was more suited for people without prior biking experience whereas regular bikers did not perceive the scenarios as challenging and could not learn from them. As for the design of the intervention, participants underlined the experience as realistic and immersive.

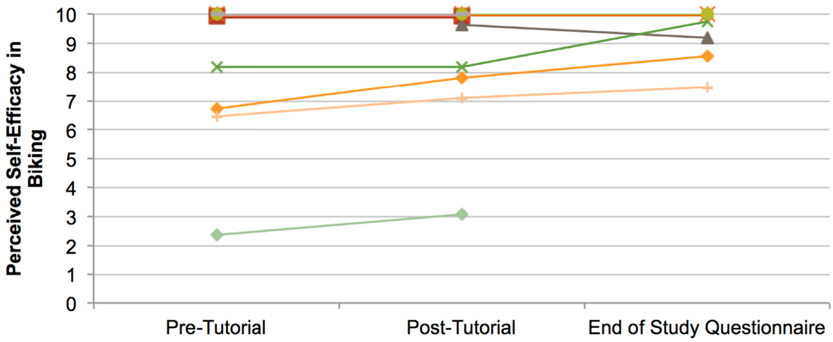


Fig. 4. Perceived self-efficacy in biking. Each line represents one participant.

4.4 Strategy III: Bike Buddy Program

Due to difficulties with scheduling, only six participants did the bike buddy rides in week 4 or after. Because the rides took place so late, no post-intervention trip data is available. The participants reported a positive experience with their "bike buddies" and perceived this strategy to be valuable for new bikers. In addition, they provided several suggestions on how this strategy could be improved. However, no clear rise in intention to bike more in the future emerged. The bike buddies were overall satisfied with helping less experienced bikers to overcome their fears and barriers. This indicates that voluntary work can be utilized in this matter.

5 Discussion

Strategy I (Frequent Biking Challenge) resulted in a significant increase in bike use above control group levels. To improve this strategy, the individual effect of the included principles should be studied. The gathered qualitative data at least suggests that competition and the collective goal elements should be designed in a way to allow social comparison also with familiar besides unknown participants. Notably, this strategy was the only one out of three that lasted for a long period of time (20 days), conveyed several main principles of socially influencing systems [20], and it used messages dependent on the actual behavior of participants, as recommended by Gatersleben & Appleton [13].

Although the design of strategy II (Virtual Bike Tutorial) was described as an immersive experience, it should be examined how to achieve a higher rise in perceived self-efficacy. Furthermore, it remains unclear if this change will actually lead to an increase in bike use. Due to scheduling issues and subsequent low actual participation the effect of strategy III (Bike Buddy Program) could not be examined by the quantitative trip data. The conducted interviews with the participants suggest that this one time biking experience did not change their intention to bike. Therefore it must be assumed, that this intervention design does not lead to a sufficient behavior change.

No change in perceived risk and perceived safety in biking could be identified. This indicates, that as most participants were already used to biking, they already had an estimate on the related safety and risk aspects. A change in actual bike use did not lead to a subsequent change of the individual evaluation of risk and safety associated with biking, at least not in the short term of this study. Our future research will therefore focus on more and other aspects that influence biking (e.g. bike availability, experience of biking or general attitudes). Furthermore, it will emphasize the use of qualitative methods to better assess why interventions show certain outcomes.

6 Conclusions and Future Research

This paper provides several contributions. Three persuasive strategies were designed for persuading people to bike as a low-energy mode of transportation and an evaluation of these were presented. The Frequent Biking Challenge showed an increase in bike use. Future research can focus on the individual principles applied as well as the analysis, for whom these are effective under which circumstances. The Virtual Bike Tutorial and Bike Buddy Program got promising feedbacks, but no clear conclusions about their outcomes can be drawn yet. Further evaluation of these strategies is needed and future research should focus more on novice bikers and evaluate the potential of these strategies to encourage them to bike. More elaborated technologies (e.g. immersive virtual environments) to simulate biking could improve the persuasive power of this design. This may be combined with a virtualized bike buddy experience, providing guidance to a user.

Overall, the presented study explored a set of strategies and features that shall act as a valuable base for future research on how to design and implement persuasive technologies [8] and socially influencing systems [20] in the field of mobility.

Acknowledgments. The authors gratefully acknowledge Kent Larson and Geraldine Fitzpatrick for their advice and support within this research project. We thank Bernhard Dorfmann for providing his expertise in tutoring novice bikers. Our special acknowledgement is due to Sandra Richter for her contributions to this research project and her help in performing the presented experiments.

References

1. Ajzen, I.: The Theory of Planned Behavior. Organizational Behavior and Human Decision Processes 50(2), 179–211 (1991)
2. Bandura, A.: Guide For Constructing Self-Efficacy Scales. In: Pajares, F., Urdan, T.C. (eds.) Self-Efficacy Beliefs of Adolescents, pp. 307–337. Information Age Publishing, Greenwich (2006)
3. Bandura, A.: Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice-Hall, Englewood Cliffs (1986)

4. Baumeister, R.F.: The Self. In: Gilbert, D.T., Fiske, S.T., Lindzey, G. (eds.) The Handbook of Social Psychology, pp. 680–740. McGraw-Hill, New York (1998)
5. Chittaro, L.: Changing User's Safety Locus of Control through Persuasive Play: An Application to Aviation Safety. In: Spagnolli, A., Chittaro, L., Gamberini, L. (eds.) PERSUASIVE 2014. LNCS, vol. 8462, pp. 31–42. Springer, Heidelberg (2014)
6. Cialdini, R.B.: Influence: The Psychology of Persuasion. HarperCollins ebooks (2007)
7. Flüchter, K., Wortmann, F., Fleisch, E.: Digital Commuting: The Effect of Social Normative Feedback on E-Bike Commuting – Evidence From A Field Study. In: Proceedings of the European Conference on Information Systems ECIS, pp. 1–14. Tel Aviv (2014)
8. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann, San Francisco (2003)
9. Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J.A.: UbiGreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1043–1052. ACM, New York (2009)
10. Gabrielli, S., Maimone, R.: Are Change Strategies Affecting Users' Transportation Choices? In: Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI, pp. 1–4. ACM, New York (2013)
11. Gärling, T., Fujii, S.: Travel behavior modification: Theories, Methods, and Programs. In: The Expanding Sphere of Travel Behaviour Research, pp. 97–128. Emerald Group (2009)
12. Gass, J.S., Seiter, R.H.: Persuasion, Social Influence, and Compliance Gaining. Allyn & Bacon, Boston (2010)
13. Gatersleben, B., Appleton, K.M.: Contemplating Cycling to Work: Attitudes and Perceptions in Different Stages of Change. Transportation Research Part A: Policy and Practice 41(4), 302–312 (2007)
14. Heinen, E., van Wee, B., Maat, K.: Commuting by Bicycle: An Overview of the Literature. Transport Reviews 30(1), 59–96 (2010)
15. Malone, T.W., Lepper, M.: Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning. In: Snow, R.E., Farr, M.J. (eds.) Aptitude, Learning and Instruction: III. Conative and Affective Process Analyses, pp. 223–253. Erlbaum, Hillsdale (1987)
16. Mayring, P.: Qualitative Content Analysis. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research 1(2), Art. 20 (2000), `http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204`
17. NOAA (2014), `http://www1.ncdc.noaa.gov/pub/orders/cdo/435898.csv`
18. Prochaska, J.O., DiClemente, C.C.: The Transtheoretical Approach: Crossing Traditional Boundaries of Change. Dow Jones/Irwin, Homewood IL (1984)
19. Stibe, A.: Socially Influencing Systems: Persuading People to Engage with Publicly Displayed Twitter-based Systems. Acta Universitatis Ouluensis (2014)
20. Stibe, A.: Towards a Framework for Socially Influencing Systems: Meta-Analysis of Four PLS-SEM Based Studies. In: MacTavish, T., Basapur, S. (eds.) Persuasive Technology. LNCS, vol. 9072, pp. 171–182. Springer, Heidelberg (2015)
21. United Nations, Department of Economic and Social Affairs, Population Division. World Urbanization Prospects: The 2014 Revision Highlights (2014)
22. Wills, T.A.: Downward Comparison Principles in Social Psychology. Psychological Bulletin 90(2), 245–271 (1981)
23. Wilson, S.R., Benner, L.A.: The Effects of Self-Esteem and Situation upon Comparison Choices During Ability Evaluation. Sociometry 34(2), 381–397 (1971)

The Art of Online Persuasion through Design: The Role of Issue Involvement as it Influences Users based on Prior Knowledge

Dianne Cyr
Simon Fraser University
cyr@sfu.ca

Milena Head
McMaster University
headm@mcmaster.ca

Eric Lim
University of New South Wales
e.t.lim@unsw.edu.au

Agnis Stibe
MIT Media Lab
agnis@mit.edu

ABSTRACT

With a goal to investigate the dynamics of online persuasion, this research extends the Elaboration Likelihood Model to determine the relative effects of argument quality as a central route to influence attitude change versus design elements (specifically image appeal, navigation design, and connectedness) as peripheral route cues to attitude change. Results emanating from this research are based on a broad sample of 390 participants who viewed a website about the merits of the Keystone XL oil pipeline. The findings indicate that in addition to argument quality, the design of the website can influence attitude change. Further, there are differences in how those with high or low prior knowledge of a persuasion topic are influenced. Of interest, change in issue involvement is less important for the high knowledge group, but for the low prior knowledge group it mediates the peripheral route impacts on attitude change.

Keywords

website design; Elaboration Likelihood Model; online persuasion, issue involvement

INTRODUCTION

Persuasion refers to human communication designed to *influence* the autonomous judgments and actions of another (Simons et al., 2001), who must make optimal compromises among conflicting forces (McGuire, 1973). In the information technology (IT) realm, persuasive systems may be defined as computerized software or information systems designed to reinforce, change or shape attitudes or behaviors, or both, *without* using coercion or deception (Fogg, 2003; Oinas-Kukkonen and Harjumaa, 2009). Unlike earlier information systems research when the IT artifact was assumed to be "neutral" and designed to serve the needs of the user, a persuasive IT has potential to change the user's attitudes. Historically, there has been considerable investigation of persuasion in sociology, psychology, and marketing. However, there are few instances when persuasion has been specifically examined in the context of information technology, and more particularly as an IT artifact. However, there is a compelling case for this work when, in an increasingly online environment, the aim is to convince individuals or groups to take a course of action not normally pursued. For example, in e-business this might include creating a web environment when users are prompted to purchase sustainable products. In an e-health setting, there are important social implications if users are persuaded to immunize their children against infectious diseases.

As a theoretical underpinning to this research, we use the Elaboration Likelihood model, or ELM (Petty and Cacioppo, 1984). The basic premise of the ELM is that persuasion may occur through a central route based on strength of arguments presented, or a peripheral route based on cues such as those provided in this case by the design of the website. The extent to which an individual chooses to scrutinize the data provided about the persuasion object depends on the "elaboration likelihood". In high elaboration likelihood states people are more likely to engage in careful scrutinization or thoughtful processing of an *informational message* and therefore tend to be more persuaded by argument quality than by peripheral cues. In contrast, those in the low elaboration likelihood states (i.e. lacking the motivation or ability to deliberate thoughtfully) also tend to be motivated by peripheral cues. In this research such cues are the design elements that appear on a website such as image appeal, navigation design, or connectedness with others on the site. In sum, peripheral cues will have a stronger influence on attitude change or persuasion when someone is in a low elaboration likelihood state compared to those in high elaboration likelihood states.

An interesting caveat to the ELM, and important to this research, is that the level of personal relevance with a topic influences the likelihood of elaboration, and the content of arguments is important for persuasion (Petty and Cacioppo, 1984). For those with low involvement, the quality of arguments is less important (Ibid.) and the source attractiveness (or in this case website design) has particular impact. Involvement moderates the main effects of argument quality and source credibility on perceived usefulness (Bhattacherjee and Sanford, 2006). Further, when product involvement is low, consumers are less likely to believe content claims, and instead focus on graphic motivators (Vaughn, 1986). Therefore, individuals in the low elaboration state are expected to be most influenced by peripheral cues - with changes in issue involvement as a precursor to attitude change. Following on these findings, in the current research we are interested to further explore conditions of differing levels of prior knowledge. Further, we are interested to explore if change of issue involvement by a user will result in attitude change, and how this may differ to those with low versus high knowledge of the topic of persuasion. This will serve to extend the ELM in the particular context of a design environment.

More specifically, this research is focused on the following issues:

1. Argument quality is an important mechanism in activating the central route to influence change in attitude depending on an individual's level of prior knowledge of the target issue. For individuals with low prior knowledge of the target issue, we postulate that argument quality might be less salient in its impact on attitude change, as compared to those of peripheral cues from specific design features such as Image Appeal, Navigation Design, or Connectedness. The reverse is postulated for individuals with high prior knowledge of the target issue.
2. Change in issue involvement moderates the effects of argument quality and characteristics of the source on perceived usefulness as outlined above. In this study, we are interested to examine if change in issue involvement potentially mediates the relationships of argument quality and design elements to attitude change.
3. To what extent do the three design elements chosen for this study impact attitude change? If so, what is the relative impact of design compared to argument quality? Do differences exist based on prior knowledge of a topic?

THE RESEARCH MODEL

Based on the preceding, a theoretical model is developed in Figure 1 to test the relative impact of Argument Quality and Design Elements (Image Appeal, Navigation Design, Connectedness) to Attitude Change and Change in Issue Involvement. We are likewise interested to ascertain whether or not Change in Issue Involvement will mediate Argument Quality or Design Elements to Attitude Change. Refer to Figure 1.

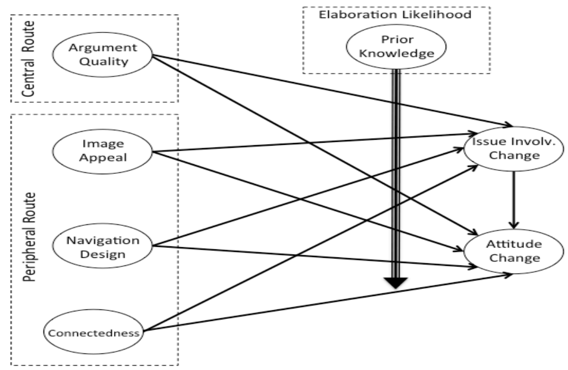


Figure 1. Research Model

HYPOTHESIS DEVELOPMENT

Argument Quality, Attitude Change, Issue Involvement

Bhattacherjee and Sanford (2006) used the ELM to investigate the relative importance of the central information processing route operationalized as argument quality and the peripheral route operationalized as source credibility resulting in attitude change. They found argument quality has a positive effect on a user's perceived usefulness ultimately leading to IT usage intention. Angst and Agarwal (2009) used the ELM to examine individual persuasion in the adoption of electronic health records. These researchers were particularly interested to determine how argument framing and issue involvement interact to influence individual persuasion. Based on the established role of Argument Quality as it relates to Attitude Change, we also explore the role of Change in Issue Involvement on Attitude Change. Further, we are interested to determine if there are any differences in these relationships depending on the role of user expertise (as considered by Bhattacherjee and Sanford, 2006) operationalized in our study as Low Prior Knowledge or High Prior Knowledge users. This leads to our first hypotheses:

Hypothesis 1: Argument Quality positively influences Attitude Change.
Hypothesis 2: Argument Quality positively influences Change in Issue Involvement.
Hypothesis 3: Change in Issue Involvement positively influences Attitude Change.

Image Appeal, Navigation Design, Connectedness

Visual Design of a website is important because it improves website aesthetics and emotional appeal (Djamasbi et al., 2010), which may in turn lead to more positive user attitudes (DeWulf et al., 2006). This encompasses Web elements such as use of images, photographs, colors, shapes, or font type (Garrett, 2003). Based on the advertisement involvement model by Vaughn (1986) and tested in numerous studies, (p. 58). "source attractiveness/likeability primarily serves as a peripheral cue, having a greater effect on persuasion when elaboration likelihood is low rather than high" (Hanafizadeh and Behboudi, 2012, p. 56). We would therefore expect that Prior Knowledge would influence the relationship of Image Appeal to Change in Issue Involvement and Attitude Change, and that the moderation effect would be stronger when prior knowledge (elaboration likelihood) is low. In this case, image appeal will serve as peripheral cue that serves to change an individual's perception and involvement with a topic especially when knowledge of a topic is limited, which in turn results in attitudinal change. This expectation is represented in the hypotheses below:

Hypothesis 4: Image Appeal positively influences Change in Attitude.
Hypothesis 5: Image Appeal positively influences Change in Issue Involvement.
Hypothesis 6a: Prior Knowledge will attenuate the relationship of Image Appeal with Change in Issue Involvement.
Hypothesis 6b: Prior Knowledge will attenuate the relationship of Image Appeal with Change in Attitude.

Navigation Design refers to the navigational scheme used to help or hinder users as they access different sections of a website (Garrett, 2003). For example, this could include whether text is horizontal or vertical, and the number of drop down menus or submenus. Similar to the rationale for Image Appeal (above), we expect that Navigation Design is a design element which contributes to Attitudinal Change and Change in Issue Involvement through the ELM peripheral route. Also similar to Image Appeal, we expect the moderation effect for those in the low Prior Knowledge group will be higher than for those with existing and substantial knowledge of the topic of persuasion. This leads to the following hypotheses

Hypothesis 7: Navigation Design will positively influence Change in Attitude.
Hypothesis 8: Navigation Design will positively influence Change in Issue Involvement.
Hypothesis 9a: Prior Knowledge will attenuate relationship of Navigation Design with Change in Issue Involvement.
Hypothesis 9b: Prior Knowledge will attenuate the relationship of Navigation Design with Change in Attitude.

Online users increasingly expect to engage and connect with others on a website or other medium. In a consumer driven context, some researchers have suggested online interactivity including **Connectedness** helps vendors to build good customer relations (Ghose and Dou, 1998). In the current investigation, Connectedness refers to the extent visitors to the persuasive website share views, feel they benefit from the community of visitors to the website, and share a common bond with others on the website. Following the same logic as for Image Appeal, and Navigation Design (above), we propose the following hypotheses:

Hypothesis 10: Connectedness will positively influence Change in Attitude.
Hypothesis 11: Connectedness will positively influence Change in Issue Involvement.
Hypothesis 12a: Prior Knowledge will attenuate the relationship of Connectedness with Change in Issue Involvement.
Hypothesis 12b: Prior Knowledge will attenuate the relationship of Connectedness with Change in Attitude.

RESEARCH METHODOLOGY

Participants

The sample was recruited via a market research firm with access to a broad pool of participants. This was a balanced stratified sample by the North American age demographic (over the age of 18) consisting of 52.3% males/47.7% females and 77.7% having a college education or higher. The total sample size is 390.

Experimental Task and Design

The Keystone XL oil pipeline (http://www.keystone-xl.com) website was chosen due to its aesthetic and human focused design. Participants were asked about their knowledge, attitudes and involvement with the Keystone XL oil pipeline issue. Once participants completed their browsing of the Keystone XL website, they were asked three content-specific questions as a manipulation check to ensure attention was paid to the website. Only those participants who correctly answered all three questions were allowed to proceed to the survey questions. The survey asked participants to respond to items for Argument Quality, Image Appeal, Navigation Design and Connectedness. Participants were again asked about their attitudes and involvement with the Keystone XL pipeline issue in order to determine their potential change in these constructs from viewing the website. The average completion time for the experiment (website browsing and survey questions) was approximately 29 minutes.

ANALYSIS

In this research, survey items were adapted from previously validated work, Therefore, content validity for these constructs was established through past research. A PLS approach to confirmatory factor analysis (CFA) was used to assess the psychometric properties of the multi-item constructs and determined our instrument demonstrated both discriminant and construct validity. Structural Equation Modeling (SEM) was conducted using PLS, which is a component-based approach.

Although results for the overall model were determined (n=390) due to space constraints we will focus on a comparison of the low versus high prior knowledge groups. The ELM posits that individuals in low elaboration likelihood states (i.e. low Prior Knowledge) tend to be motivated by peripheral cues (design elements) whereas those in high elaboration likelihood states (i.e. high Prior Knowledge) are more persuaded by Argument Quality. Prior Knowledge was a three-item construct measured on a 7-point Likert scale, and the sample was divided into Low Prior Knowledge (0 and 4.2) and High Prior Knowledge (4.3 and 7) based on average participant scores. Figure 2a and 2b show the results of our PLS analysis for the Low Prior Knowledge group (n=216) and High Prior Knowledge group (n=125), respectively. Only significant moderation paths are shown on these figures.

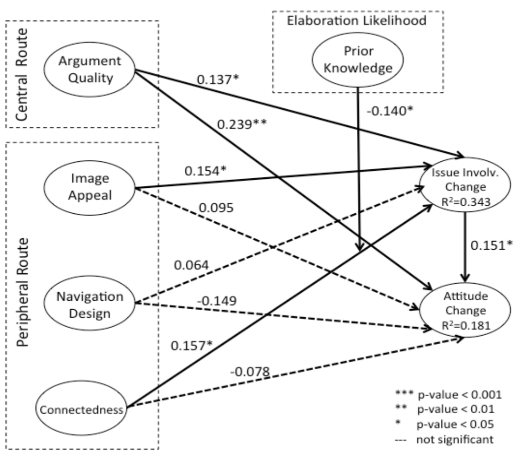


Figure 2a: Low Prior Knowledge Groups (n=216)

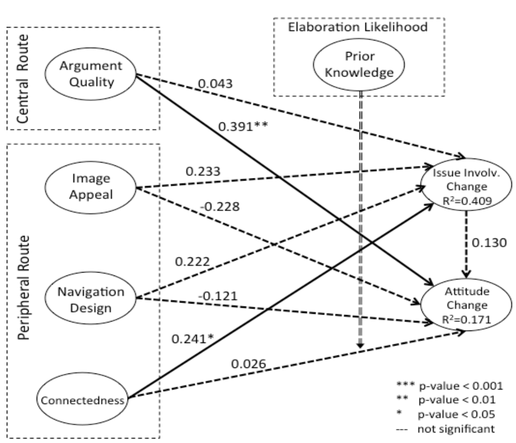


Figure 2b: High Prior Knowledge Groups (n=125)

In sum, for the Low Prior Knowledge Group significant paths indicate: Argument Quality results Issue Involvement Change and Attitude Change; Image Appeal results in Issue Involvement Change; Connectedness results in Issue Involvement Change; and Issue Involvement Change results in Attitude Change. For the High Prior Knowledge Group significant paths indicate: Argument Quality results in Attitude Change; and Connectedness results in Change Issue Involvement Change. The only significant moderating effect of Prior Knowledge for either group was between Connectedness and Change in Issue Involvement for the Low Prior Knowledge Group.

DISCUSSION AND CONTRIBUTIONS

Based on the preceding framework and results, there are several theoretical contributions that emanate from this investigation. First, as already confirmed, argument quality is important as a central route to influencing change in attitude. However, new to this research is that for those with low prior knowledge, argument quality directly impacts attitude change but it is also mediated through change in issue involvement. For those with high prior knowledge, argument quality directly impacts attitude change but does not impact change in issue involvement. Therefore, the role of issue involvement plays a role in affecting attitudinal change. One explanation is that for those with the least prior knowledge, they also have the least commitment to the topic; then as they become more involved, their attitudes are most likely to be mediated by change in issue involvement leading to attitude change.

Second, it appears that change in issue involvement is of less importance to the high knowledge group (it does not have a significant effect on change in attitude). But change in issue involvement is important for the low prior knowledge group and mediates the peripheral route design impacts on attitude change. While involvement is known to influence commitment (Vaughn, 1986), we believe this is the first study in which issue involvement is studied in tandem with both argument quality and design features.

Third, this is one of the first investigations to study the relative impact of argument quality versus design elements on attitude change. While not all design elements had a direct impact on attitudinal change (e.g. navigation design for either the low or high knowledge groups), it should be noted that the peripheral (design) route is more important for the low prior knowledge group. Further, connectedness and image appeal have a direct significant effect on change in issue involvement (which in turn has a direct significant effect on change in attitude). The feeling of connectedness is important to influence issue involvement, regardless of prior knowledge level. However for those with less prior knowledge, connectedness plays an even stronger influential role on changing issue involvement. Not only does connectedness have a significant direct effect on change in issue involvement, but it also moderates this relationship such that those with the lowest prior knowledge would see the connectedness of the site as having an even stronger influence on change in issue involvement. In sum, this signals the important role that design elements may have in the attitudinal change process.

From a practical standpoint, this research is relevant in a variety of online settings (e.g., e-government, e-health, e-commerce), when the goal is to change user behavior or attitudes. It provides insights for designers to design elements that can potentially persuade users. While three design elements were investigated in this study, there is much scope for designers to determine which design elements are most appropriate for the goal of persuasion. Based on the results of this study, increasing user knowledge through design will in turn increase issue involvement leading to potential attitudinal change.

A limitation to this research is that a single website was used. It may be that users would respond differently depending on the site, and that a website on the topic of an oil pipeline may not be of interest to many of the users. So while this particular website was carefully selected, other websites could possibly yield differing results depending on the level of user knowledge and involvement with the topic. Likewise, there are numerous website features such as interactivity, perceived social presence, or "Inspiration" (Fogg, 2003) that could have been tested. While a sample of design elements was tested in this investigation, this also becomes an avenue for future research. These design elements should also be investigated across cultures. Although there is previous research into cross-cultural elements of website design (e.g. Cyr et al., 2009) and how users in different cultures have different preferences for design features, to our knowledge there is little if any research that investigates specific persuasive design features in different cultural contexts.

In sum, and as the world increasingly moves to online media as a form of communication, then how to appeal to the sensibilities of users - including how to persuade - is a topic of immense interest. A focus on how attitude change through both words and through design is expected to be a topic of continuing investigation.

ACKNOWLEDGMENT

This research was funded by the Social Sciences and Humanities Research Council of Canada.

REFERENCES

1. Angst, C. and Agarwal, R. (2009). Adoption of Electronic Health Records in the Presence of Privacy Concerns: The Elaboration Likelihood Model and Individual Persuasion, *MIS Quarterly*, 36(2), 339-373.
2. Bhattacherjee, A. and Sanford, C. (2006). Influence Processes for Information Technology Acceptance: An Elaboration Likelihood Model. *MIS Quarterly*, 30(4), 805-825.
3. Cyr, D., Head, M., Larios, H. and Pan, B. (2009a). Exploring human images in website design: A multi-method approach. *MIS Quarterly*, 33(3), 539-566.
4. DeWulf, K., Schillewaert, N., Muylle, S. & Rangarajan, D. (2006). The Role of Pleasure in Web Site Success. *Information and Management,* 43, 434-446.
5. Djamasbi, S., Siegel, M., Tullis, T. and Dai, R. (2010). Efficiency, Trust, and Visual Appeal: Usability Testing through Eye Tracking, *Proceedings of the 43rd Hawaii International Conference on System Sciences.* pp. 1-10.
6. Fogg, B.J. (2003). Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann Publishers.
7. Garrett, J. (2003). The Elements of User Experience: User-Centered Design for the Web. Thousand Oaks, CA: New Riders.
8. Ghose, S. and Dou, W. (1998). Interactive functions and their impacts on the appeal of Internet presence sites, *Journal of Advertising Research,* 38(2), 29-43.
9. Hanafizadeh, P. and Behboudi, M. (2012). *Online Advertising and Promotion.* IGI Global.
10. McGuire, W.J. (1973). "Persuasion" in G.A. Miller (ed.) *Communication, Language, and Meaning Psychological Perspectives.* New York: Basic Books, 242-255.
11. Oinas-Kukkonen, H. and Harjumaa, M. (2009). Persuasive Systems Design: Key Issues, Process Model, And System Features. *Communications of the Association for Information Systems*, 24(28), 485-500.
12. Petty, R.E. and Cacioppo, J.T. (1984). Source Factors and the Elaboration Likelihood Model of Persuasion. *Advances in Consumer Research*, 11, 668-672.
13. Simons, H.W., Morreale, J. and Bronbeck, B. (2001). *Persuasion in Society*. Thousand Oaks: Sage Publications, Inc.
14. Vaughn, R. (1986). Cited in Hanafizadeh, P. and Behboudi, M. (2012). *Online Advertising and Promotion.* IGI Global.

CHALLENGED TO BIKE: ASSESSING THE POTENTIAL IMPACT OF GAMIFIED CYCLING INITIATIVES

Matthias Wunsch*
AIT Austrian Institute of Technology, Dynamic Transportation Systems
Giefinggasse 2, 1210, Vienna, Austria
Vienna University of Technology, Human Computer Interaction
Argentinierstrasse 8, 1040, Vienna, Austria
Tel: +43 50550-6614 Fax: +43 50550-6439 Email: matthias.wunsch.fl@ait.ac.at

Alexandra Millonig
AIT Austrian Institute of Technology, Dynamic Transportation Systems
Giefinggasse 2, 1210, Vienna, Austria
Tel: +43 50550-6672 Fax: +43 50550-6439 Email: alexandra.millonig@ait.ac.at

Stefan Seer
AIT Austrian Institute of Technology, Dynamic Transportation Systems
Giefinggasse 2, 1210, Vienna, Austria
Tel: +43 50550-6478 Fax: +43 50550-6439 Email: stefan.seer@ait.ac.at

Katja Schechtner
MIT Massachusetts Institute of Technology, Media Lab
20 Ames Street, E15-368D, Cambridge, MA, 02139, USA
Tel: + 1 617-253-9880 Fax: +1 866-740-3158 Email: katjas@mit.edu

Agnis Stibe
MIT Massachusetts Institute of Technology, Media Lab
20 Ames Street, E15-368, Cambridge, MA, 02139, USA
Tel: + 1 617-253-1000 Email: agnis@mit.edu

Ryan C.C. Chin
MIT Massachusetts Institute of Technology, Media Lab
20 Ames Street, E15-368D, Cambridge, MA, 02139, USA
Tel: + 1 617-253-6828 Fax: +1 617-258-6264 Email: rchin@mit.edu

*) corresponding author

Word count: 6,245 words text + 4 tables/figures x 250 words (each) = 7,245 words

Submission date: November 15, 2015

ABSTRACT

Cycling is promoted by a variety of initiatives and events all around the world. Given the popularity of such campaigns, there is surprisingly limited literature assessing their effectiveness and investigating the involved social processes. This paper presents findings from two case studies of gamified biking initiatives which have been evaluated using both quantitative-statistical and qualitative-interpretative methods: (1) a small-scale research experiment ("Frequent Biking Challenge") and (2) an annual national biking campaign ("Bike to Work").

The two case studies provide insights into their general effects, on changes for different groups of bikers, the effectiveness of different elements of the interventions, and the methodological requirements for evaluating short and long-term impacts of campaigns aiming at increasing biking in cities. Both initiatives successfully encouraged people to bike more in the short run, 25% in the experimental group and 36% participating in the campaign increased their amount of biking during the intervention period. However, effects vary for different groups of bikers and only the "Bike to Work" campaign shows a lasting effect with 26% of those participating in previous years continuing to bike more. The results show that in order to induce long-term behavior changes, gamified biking initiatives have to be embedded into everyday life, enable social interactions and provide mutual encouragement.

Keywords: Mobility behavior change, cycling, gamification, soft policy measures, bike to work, evaluation

INTRODUCTION

Growing cities as well as increasing greenhouse gas emissions are driving factors for a global rethinking of transportation systems. Biking can be seen as a central element to that, as it is a low-emission, low-cost, healthy and space efficient mode of transportation. To increase bike usage, cities are investing in cycling infrastructure and are implementing bike share programs. Additionally, numerous initiatives, publicity campaigns and events are being supported to promote biking. The aims of such "soft" policy measures typically include motivating car users to voluntarily switch to more sustainable modes of transport, e.g. cycling (1). Evidence on the effectiveness of such interventions is, however, either inconsistent, can hardly be generalized, or is non-existent at all (2).

With growing popularity of gamification approaches for triggering behavior changes in different contexts, the use of game elements like incentives or rewards in products, services and campaigns aiming at modal shift is also on the rise (3). Some examples are bike programs, such as "bike-to-work" events, which can be found globally (4). Such promotional programs often use elements like competition, lotteries, team experience or awards, adding an emotional quality to the more objective arguments for biking, such as health benefits, time saving or climate change mitigation.

Yet, just as for other soft policy measures, studies on the actual effectiveness of gamified biking campaigns are still scarce: Rose & Marfurt (5) evaluated the impacts on travel behavior change of a major one-day bike-to-work and found that 27% of those riding to work for the first time due to the event continued to do so. For a similar one-day initiative to promote active transport to school no clear effect could be shown (6). Piatkowski et. al. (7) identified different groups participating in one-day bike-to-work events and identified barriers to increased commuter cycling. Despite the small number of existing literature on the general effects of gamification in the mobility context, the findings indicate its usefulness to provoke behavior and attitude change. Playful elements enrich the user's motivation and engagement in specific activities as well as in exploring new possibilities and options (8–10). Concurrently, the success of interventions strongly depends on the nature of the gamified system, the applied game mechanics and the types of players (8). Due to the contextual complexity of daily life and daily mobility routines, also other behavior than originally intended may be triggered by gamified interventions (9, 11). Hence, a better understanding of potential reactions to soft policy initiatives is key to designing successful interventions.

The contribution of this paper is to examine the strategies employed in gamified biking campaigns with the aim of changing accustomed mobility behavior patterns. Results from from two case studies focusing on promoting commuting by bike through the use of game mechanics are presented. The first study describes an experimental intervention called "Frequent Biking Challenge" which we carried out in a small-scale randomized controlled trial over a period of four weeks during fall 2014 in the Cambridge/Boston area in Massachusetts, USA (12). It was complemented by a long-term qualitative evaluation nine months after the initial trial. The second case study presents an evaluation study accompanying the 2015 Austrian national cycling initiative called "Bike to Work", a one-month campaign that promotes bike commuting which has been conducted annually since 2011. Both examples are comparable in terms of objectives, intervention period and gamification elements, although the nature of the two examples (small-scale experiment vs. established national event) causes significant differences in the sample sizes and hence explanatory power of the results. Still, each intervention and the comparison of them provide valuable insights into drivers of the success of biking campaigns.

Based on the hypotheses that gamified biking initiatives can trigger mobility behavior changes for specific target groups, the case studies are discussed with respect to the following questions:

- How and to what extent are gamified biking initiatives affecting behavior change?
- Which aspects of gamified biking initiatives are most effective for inducing behavior change?
- Which groups can be identified within participants of by gamified biking initiatives to increase biking on their work trips?
- Which methods can be used for assessing the potential short- and long-term impact on mobility behavior?

The following two sections provide comprehensive descriptions of the case studies, each focusing on the intervention design, the methods used for evaluating the effects, and the results of the analysis. In the subsequent section, the main research questions are discussed by comparing the dominant characteristics and outcomes of the case studies. The last part of the paper includes conclusions and provides an outlook on further research, which will build on the findings presented in this paper.

CASE STUDY 1: FREQUENT BIKING CHALLENGE

Intervention Design

As part of a larger study to promote biking with different persuasive strategies we carried out small-scale experiments in fall 2014 *(12)*. The "Frequent Biking Challenge" was one intervention within these experiments designed for encouraging participants to collect points for their bike rides in order to achieve different experience levels. We developed a web-based application (see figure 1) to visualize level names and descriptions, badges and a competition element in form of a leaderboard. At the beginning of the study all participants were invited to jointly reach a total number of points. After this collective goal was fulfilled, a social comparison feature was installed, showing the number of individual bike rides in comparison to the average rides of all participants and to the current front-runner's number of rides.

Information was communicated to the participants via email. During the first week, participants received email updates on a daily basis; after that the frequency was reduced to three to four times per week. The emails were sent in the evenings in order to influence the planning of the next day by various messages. Figure 1 illustrates the visualization of levels and an example of an information update.

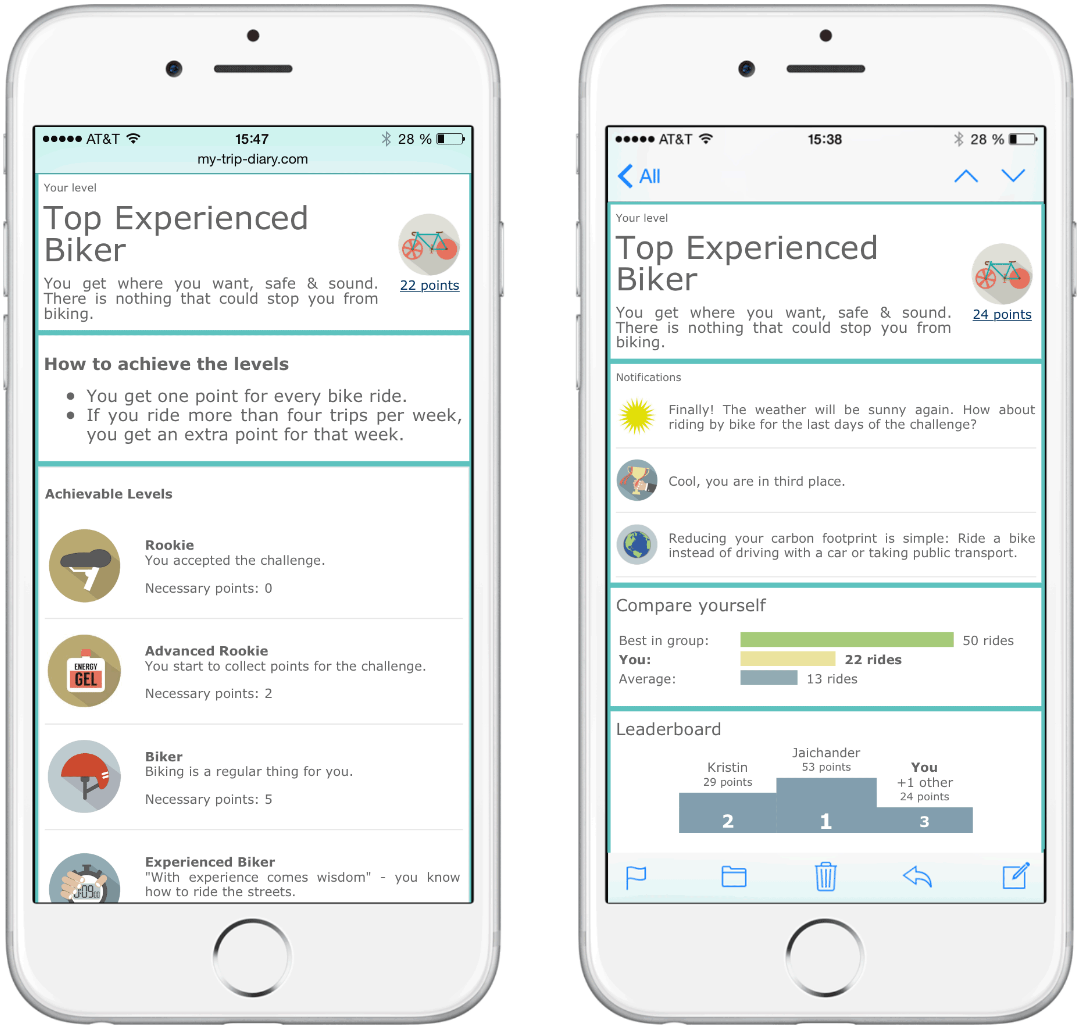


FIGURE 1 Screenshots showing the levels (left) and information updates (right) of the "Frequent Biking Challenge"

In particular, this approach focused on the following persuasive elements and hypotheses:

- Triggering: Mobility mode choices are altered by suggestions to ride a bike.
- Competition: Behavior change is facilitated through a leaderboard that allows individuals to see where they stand in an overall ranking *(13)*.
- Social comparison: Enabling participants to compare their performance with the best and the average performance in a group motivates individuals to change their behaviour *(13)*.
- Raised awareness: By regularly reporting mobility choices participants become aware of otherwise intuitive behavior and start to question existing habits.

Methods

The experiment was conducted over a period of four weeks (29 days) in October 2014 in Greater Boston (MA). The first 12 days were used to gather baseline data (pre-intervention phase). After that, the experimental group participated in the "Frequent Biking Challenge" for 17 days (during-intervention phase). The control group did not receive any intervention. This small-scale experiment was intended to provide first insights into potential reactions and requirements for subsequent large-scale interventions.

Study Sample

Participants were recruited through mailing lists at the Massachusetts Institute of Technology (MIT) and subsequently consisted mainly of students and to a lesser extend staff members. Based on a questionnaire including current mobility patterns, only people biking less than three times a week were selected as participants for the study. The resulting 29 participants were randomly assigned to the experimental group (n=13) or the control group (n=16).

For the analysis, only data from participants who had reported at least 80% of their daily trips was used in order to achieve a minimum level of data quality (experimental group: n=12; control group: n=10), consisting of 11 women and 11 men. To ensure access to bikes, all participants were provided with a one-month membership for the local bike-sharing scheme (Hubway) and a helmet if required.

Collection and Analysis of Data

All participants had to report their daily trips using a web-based application. For each trip, participants were instructed to select a trip purpose and the transport modes they used (single or multi-modal): "walking" (minimum 5 minutes walks), "biking", "public transport", "car" and "other". Based on the collected data the share of biking among all modes y_d was computed per person per day. In order to outweigh other effects than the introduced intervention influencing mode choice (e.g. weather), a controlled share of bike trips per day $x_{g,d}$ between the daily bike share of each participant of the experimental group $y_{g,d}$ and the mean of daily bike share within the control group $\bar{y}_{c,d}$ was computed for each day by

$$x_{g,d} = y_{g,d} - \bar{y}_{c,d} \,. \tag{1}$$

To assess the effect of the intervention, we determine the average controlled share of biking for the time before the intervention $z_{g,pre}$ and the average controlled share of biking during the intervention $z_{g,during}$ given by

$$z_{g,pre} = \frac{1}{N_{pre}} \sum_{d=1}^{N_{pre}} (x_{g,d}),\; z_{g,during} = \frac{1}{N_{during}} \sum_{d=1}^{N_{during}} (x_{g,d}) \,, \tag{2}$$

with N_{pre} denoting the number of days before and N_{during} the number of days during the the "Frequent Biking Challenge". The pre- and during-intervention values of the average controlled share of biking were then tested for normality with a Shapiro-Wilk test *(14)* and compared with a one-sided paired-sample t-test. Per participant changes of bike share were tested with a two-sample t-test comparing daily bike shares of the pre-intervention and during-intervention days.

Additionally, qualitative data was obtained by a post-intervention online survey, through interviews with four participants (two from each the control and experimental group) and finally an open question email exchange with the study participants nine months after the initial study.

Results

There was a statistically significant ($p<0.05$) increase of the share in biking in the individuals' modal split for three out of the 12 participants (25%) in the experimental group after starting the intervention. The comparison of the controlled share of biking for the pre-intervention phase $z_{g,pre}$ [M=-7.9%, SD=11.8%] and the controlled share of biking during the intervention $z_{g,during}$ [M=5.5%, SD=30.8%] showed an increase of 13.5 percentage points [$t(11)=2.079$, $p = 0.031$]. This underlines that the "Frequent Biking Challenge" lead to an increase in bike use above control group levels during the intervention.

The qualitative data collected directly after the intervention supported this result as participants reported, for instance, an increased awareness of biking as an alternative mode of transportation and to be slightly more engaged in doing so due to participating in the "Frequent Biking Challenge". However, due to the overall small sample size these results must be taken as first insights that require further validation.

The long term evaluation, which was conducted nine months after the initial study, showed that two out of the three participants with increased bike usage during the intervention (participants #61, #62, and #66) returned to their initial mobility habits afterwards. Quitting the newly achieved habit had different reasons: one participant mentioned seasonal influences as the reason why she got back to her initial mobility routines: *"I got out of the habit of daily biking over the winter, and now commute primarily on foot. I bike maybe once a week to run an errand or attend church, more or less what I was doing before the survey started" (#61).* For another participant access to bikes was the main issue, as participants have been provided with access to the local bike sharing scheme, and keeping their increased levels of biking has also been a question of renewing their subscription or buying a bike. *"I have since transitioned back to mostly walking. I feel that bike-share programs are not worth the cost, and unfortunately also lack a good place to store a personal bike" (#66)*. The third participant with an increased level of biking has been able to sustain this habit over time: *"I renewed the Hubway membership subscription you gave me and have been biking about the same ever since" (#62)*. However, although there was this one participant with a lasting behavior change in the experimental group, it is questionable if the intervention itself caused the transition or the mere provision of access to bikes was the decisive factor for the participant's new biking habit. Therfore no clear conclussions can be drawn for the question if the overall intervention design of the "Frequent Biking Challenge" did lead to a long lasting behavior change.

CASE STUDY 2: BIKE TO WORK

Intervention Design

The "Bike to Work" initiative (in German: "Radelt zur Arbeit", www.radeltzurarbeit.at) is designed to encourage people to cycle on their commute. It is conducted by an Austrian biking advocacy group with financial support from the Austrian government. "Bike to Work" has been held for the fifth time in 2015 and is taking place each year during May. In 2015 over 14,800 participants joined the campaign, making it the largest of this kind in Austria.

Participants are required to sign up on the campaign's website and to form teams of two to four. They can also volunteer as so-called company coordinators which help their colleagues with registration and the creation of teams. Furthermore, they typically promote the campaign within the company. The initiative provides incentives in form of different prizes (e.g. vouchers for leisure activities or bike gear). These are distributed by (1) having a lottery for all teams achieving a minimum of 50% bike commutes for each member and (2) giving prizes to participants who are randomly called, if they biked on the day they receive the call.

Participants can either track their daily commutes with a smartphone app, or log into a website to insert the distances covered by bike. The team performance and the individual mileages are shown within the app and on the website. Figure 2 shows the tracking tab and the statistics tab of the "Bike to Work" smartphone app.

FIGURE 2 Screenshots of the "Bike to Work" (German: "Radelt zur Arbeit") app showing the tracking tab (left) and the statistics tab (right).

For this case study, the following persuasive elements and hypotheses are regarded:

- Commitment: Participants agree to bike commute within the month of the initiative and strive to remain consistent with this commitment, which is also likely to be reinforced by other team-members.
- Competition: Behavior change is motivated by comparing individual performance to other team members and by comparing team performance to other teams.
- Tangible incentives: Behavior change is motivated through the chance to qualify for a lottery by reaching a concrete goal (minimum of 50% bike commutes per member in each team during the intervention period).
- Raised awareness: By regularly reporting mobility choices participants reflect on otherwise intuitive behavior and start to question existing habits.

Methods

Study Sample

Out of the 14,809 subscribers of this year's initiative 498 individuals participated in the study by answering a standardized online questionnaire. Out of all survey respondents, 157 provided additional qualitative data by responding to open questions in the questionnaire. A question regarding mode shifts was added to the questionnaire at a later time during running the survey resulting in n=92 for this item. Within the sample, nine participants have been additionally enlisted for semi-structured qualitative in-depth telephone interviews.

Study participants were recruited with a link on the campaign's website and on different social media channels and they were eligible to participate in a lottery offering the winners bike racks for their work place. This opportunity may have caused a selection bias, as regular bikers might be more interested in having sufficient biking infrastructure and might be generally more interested in the initiative.

Collection and Analysis of Data

The standardized online questionnaire included questions about the respondents' overall perception of the "Bike to Work" campaign, biking behavior before and during the initiative, potential long-term effects, and the influence of different motivational factors encouraging biking. In addition, participants could provide supplementary comments. The content of the in-depth telephone interviews was audio-recorded and partly transcribed for the analysis process. The qualitative data was analyzed using thematic analysis according to Braun & Clarke *(14)*.

In order to compare the distribution of motivational factors between groups of participants Pearson's chi-squared tests with Yates' continuity correction were applied. Estimates of proportions within all participants are based on sample means. 95% binomial confidence intervals are based on Clopper-Pearson *(16)* and fulfill the requirements according to Brown et al. *(17)*. The statistical analyses were conducted using R *(18)*.

Results

General Satisfaction with the Initiative and Mode Shifts

Participants of "Bike to Work" have been very positive about the initiative. The question: "How do you like 'Bike to Work 2015'?" with a scale ranging from 1 "Not at all" to 7 "Very good" the initiative achieved a mean rating of 6.33 [SD=1.00]. This result has been confirmed by the in-depth interviews were all respondents expressed their satisfaction with "Bike to Work". Regarding mode shifts due to participation a reduction of car use was reported by 38.5% [95% CI 28.4%, 49.2%] of the respondents and a reduction in public transportation use was reported by 34.0% [95% CI 24.5%, 44.7%].

Changes in Bike Commuting Frequency

"Bike to Work" brings together people with very different rates of bike commuting. As can be seen in table 1, the general biking behavior is mainly structured as follows: 54.6% report that they usually commute by bike almost on a daily basis ("daily bikers"), 23.9% bike several times a week ("regular bikers"), and the remaining 21.5% ("occasional bikers") bike less often than that. The teams described in the qualitative data reflected this mix of participants, resulting in a group dynamic of having regular bikers motivate occasional bikers.

Table 1 furthermore shows the change in bike commuting frequency during the campaign. Within the full sample 36.1% increased their frequency of bike commuting during „Bike to Work“.

TABLE 1 Overview of group sizes for regularity of biking and change in frequency of bike commuting during the campaign

	Occasional bikers	**Regular bikers**	**Daily bikers**	**Full Sample**
Usual frequency of bike commuting	never to (almost) weekly	several times a week	(almost) daily	
Number of respondents*	107	117	272	496
Share among all respondents	21.4%	23.9%	54.6%	100.0%
Change during Bike to Work				
Biked more often	77.6%	52.1%	12.9%	36.1%
Biked the same	21.5%	47.9%	87.1%	63.7%
Biked less often	0.9%	0.0%	0.0%	0.2%
Total	100.0%	100.0%	100.0%	100.0%

* Two respondents skipped this question.

Among the occasional bikers, more than three out of four (77.6%) biked more often than usual during the month of the initiative. This group, which can be regarded as the ones drawn into more regular biking by the campaign, account for 16.7% of all "Bike to Work" participants [95% CI 13.6%, 20.3%]. One reason for this change was that the campaign acted as a trigger to try bike commuting. Interviewee #99 who moved to a new town six weeks before joining "Bike to Work" used to go to work by car despite the short distance of 1.7 km (1.1 mi). Getting an information about the campaign from her employer, knowing about other teams that participated in the previous year and having a colleague who already biked daily to team up with made her start to ride the bike for her commute. Participant #351 rediscovered biking due to the campaign: *"I used to live in a small town and biked a lot back then, but as I moved to a bigger city and had small kids I almost never used the bike. Now I have rediscovered biking".* Another reason for the high share of increased bike use during the initiative within the occasional bikers can be found in the eligibility criteria for the lottery (minimum of 50% bike commutes per team member), which required a significant change of the participants' mobility behavior. This led to a temporary change in order to achieve the goal: *"It is okay for one month. But for me it is also very cumbersome" (#367).*

Although the general frequency of bike commuting in the group of regular bikers is already comparatively high, still 52.1% increase their bike use during the initiative. The analysis suggests that this stems mainly from changes in situations where participants usually would not have chosen to ride the bike to work, mostly regarding days with bad weather. Typical statements for this group include: *"[Bike to Work] is indeed motivating us all to ride the bike, even when the weather is bad" (#106), "I ride now even when it is raining whereas previously I would have taken public transport" (#246), "Before [Bike to work] I took the car when there was rainy weather. Now I am riding my bike and even if there are some scattered showers I will be fine" (#224).*

Among the group of daily bikers a comparatively low number of 12.9% increased their bike use, which can be expected given the fact that this group is already mainly using a bike to get to work. The increase can be mainly attributed to the same weather related effects.

Long-Term Effects

Participants of our survey who also participated in the "Bike to Work" initiative in previous years (n=381) were asked for the resulting long-term effect. Based on the responses it is estimated that 26.4% [95% CI 23.8%, 29.0%] of "Bike to Work" participants increase their level of bike commuting in the long run after they experienced the benefits of bike commuting due to the initiative, which is much in line with findings of Rose and Marfurt *(5)*. *"I will extend 'Bike to Work' and will continue to bike to my workplace. I realized through this initiative how great it is to bike to my work. I am very enthusiastic and enjoy it every day" (#099)* In retrospective respondent #229 notably stated: *"Bike to Work 2014 has been the trigger to switch completely to the bike within the city and closer distances. Meanwhile, I am using the car just for trips over 100 km [62.1mi] or for hauling. I sold my own car and I almost never need public transport".* Although such drastic changes might be the exception rather than the rule, the data indicates a long lasting effect of the initiative. Additionally, 4.7% of the respondents stated that they had increased their bike commuting for a short period of time after "Bike to Work" but subsequently returned back to their usual mobility patterns.

Motivational Aspects

The results of the analysis of questionnaire data and qualitative data provide valuable insights into the motivational drivers of behavior changes in the course of the "Bike to Work" campaign and show that some of them differ depending on the usual frequency of bike commuting. Table 2 shows the compliance to different motivational factors according to the responses in the questionnaire survey and highlights the factors which are dependent on the usual rate of biking.

TABLE 2 Relevance of motivational factors during participation in "Bike to Work"

Group	***Occasional bikers***	***Regular bikers***	***Daily bikers***	***Full Sample***	χ^2	***p-value***
Usual frequency of bike commuting	never to (almost) weekly (n=103)	several times a week (n=116)	(almost) daily (n=266)			
Motivational factor						
Environmental protection	57.3%	54.3%	53.8%	54.6%	0.378	0.828
Health benefits	**69.9%**	**59.5%**	**45.9%**	**54.2%**	18.982	**<0.001**
Team spirit	**58.3%**	**45.7%**	**41.0%**	**45.8%**	8.927	**0.011**
Motivate others to bike	**14.6%**	**42.2%**	**44.0%**	**37.3%**	29.056	**<0.001**
Prizes	31.1%	33.6%	38.7%	35.9%	2.228	0.328
Biked distance	31.1%	39.7%	35.3%	35.5%	1.762	0.414
Individual statistics	38.8%	32.8%	30.8%	33.0%	2.158	0.340
Biking enthusiasm of colleagues	16.5%	20.7%	15.4%	16.9%	1.616	0.446

p-values are the results of Chi-Squared test for independency between all three groups. Bold numbers indicate significant dependences of the factors on the groups (p<0.05).
Survey question: "What motivated you while you were participating in Bike to Work?" Participants could select the items that they agreed with (dichotomus scale).

The motivator with the highest compliance rate among all factors was "environmental protection". At the same time, this motivator does not stand out for its persuasive effect on non-regular bikers as it is equally important for all three groups.

Health benefits achieved the second highest compliance rate showing that bike commuting is seen as a way to live a healthier life and to include physical activity into everyday routines. This has also been confirmed by statements within the interviews: *"Otherwise I would often be too lazy to do some sports in the evening after I got home from work. But by commuting by bike one has to use the bike to get home as well and one has to pedal" (#224).* Here bike commuting and the decision to bike is even a way of committing oneself to do a physical activity. Commuting by bike may take more time than available alternatives, but it acts as a physical workout at same time. By that the choice for bike commuting can be a reasonable one as the loss in time is compensated by a gain in health and fitness. If doing sports is substituted with bike commuting on long distances altogether this can even be time efficient: *"Although I need over an hour one way, overall I am saving time" (#082)*. In that light "Bike to Work" provides a framework to start or increase bike commuting as a way of 'doing something' about one's health. As health benefits were mentioned significantly more often by occasional bikers (69.9%) and regular bikers (59.5%), this aspect is a more prominent motivator during the Bike to Work participating for them than for daily bikers at 45.7% [χ^2=16.2, n=369, p<0.001 and χ^2=5.46, n=382, p=0.019].

Team spirit (i.e. doing something together as a team) was at least related to three social processes influencing participants: First, commitment, as participation required each team member to obey the 50% bike trips rule in order to qualify for the lottery. Second, increased visibility of mode choices, as daily transportation choices become subject to judgements by others. *"On days I commute by car I'll get some 'friendly' remarks from colleagues" (#224)* Third, competition, as for some of the participants being in a team also meant competing with others. *"We have two teams and this year there was quite some competition going on" (#310). "Teams within our company that biked about the same as we did sent an email stating: 'Look, we have overtaken you in the ranking" (#224).* Being motivated by team spirit was mentioned significantly more often by occasional bikers at 58.3% compared to regular bikers (45.7%) or daily bikers (41.0%) [χ^2=2.96, n=219, p=0.08 and χ^2=8.24, n=369, p=0.004], showing the importance of participating together with colleagues and friends and that this is more relevant and more motivating for occasional bikers.

Motivating others to bike was a driving factor primarily for regular and daily bikers to join "Bike to Work" (p<0.001). *"I daily bike to work anyway, but with this initiative more people get motivated and some may stick with it" (#089).* Several respondents reported on the effectiveness in this respect, e.g. *"Because of Bike to Work I could motivate two colleagues to bike commute" (#210). "Since we have started two years ago I could persuade my brother in law as well as colleagues of mine to bike commute, especially as I am riding even during winter" (#106)*. This result also demonstrates the interrelation of motivational factors for frequent bikers (daily and regular bikers) and occasional bikers, as "Bike to Work" provides them with an opportunity to promote biking and get non-bikers and occasional bikers excited about the idea of bike commuting. This effect can be seen in Table 2 in the higher compliance for "motivate others to bike" for the regular and daily bikers and the higher proportion of "team spirit" for the occasional bikers. This result has been present in most of the qualitative data as well: *"This initiative is ideal to raise my colleagues' awareness for biking" (#437)* The formation of teams consisting of occasional and regular bikers (*"Some in my team are already biking a lot; others were not so much before the initiative." #310*) may also set norm for the regularity of bike commuting. This role of already frequently biking people is comparable to previous findings pointing at their role for awareness raising and motivating others *(7)*.

Although prizes and the possibility to win them in the lottery are a main feature of this initiative, they rank only fifth within the list of motivators and are relevant for 35.9% of participants. This leads to the conclusion that the social interactions and mutual engagement introduced by the campaign are the true top features. However, prizes could be an important trigger to motivate people to join the campaign in first place. For them as for the remaining motivational factors (personal statistics, biked distance, enthusiasm of colleagues) no significant differences between the three groups of bikers emerged.

Another relevant aspect it the potential role of companies: They can support behavior changes of their employees as they may seek to encourage them to engage in healthy activities or contribute to increase ecological sustainability. In these situations, “Bike to Work” provided a useful framework. One company had created a corporate social responsibility project based and inspired on “Bike to Work” (#271). Another company was actively encouraging employees to join “Bike to Work” by providing information material, helping with the set-up process of the teams and even organizing a kick-off event (#411). Besides that, many private and public organizations help at least to distribute information about “Bike to Work” to their employees.

DISCUSSION

The two case studies presented in this paper share several characteristics: both aim at increasing the share of biking in the daily mode choices of participants (though one is only focusing on commuter trips), in both cases the intervention took place over a limited period of time (four weeks), and both used persuasive strategies and game mechanics for achieving behavior change. At the same time, the two examples also differ in specific aspects, particularly due the types of intervention: the “Frequent Biking Challenge” was a nonrecurring, small-scale scientific experiment with limited sample size, and “Bike to Work” is an annual national campaign attracting thousands of participants each year. The subsequent discussion follows the initial questions of this study.

How and to what extent are gamified biking initiatives affecting behavior change? Which aspects of gamified biking initiatives are most effective for inducing behavior change?

As gamification elements like competition, points and rewards are effective measures for encouraging people to get involved in an intervention, participants of the “Frequent Biking Challenge” on one side did show higher levels of biking during the intervention period. However, these elements by themselves might not be sufficient for inducing enduring behavior change. On the other side, biking campaigns such as “Bike to Work”, unite long-standing frequent bikers, occasional and even non-bikers. They embed the idea of bike commuting into the everyday social context, such as collaborating with colleagues and put gamification elements on top of that. Furthermore, the reqirement for participants to form small teams of two to four seems to even enhance all of these effects. Combining social interactions, mutual encouragement and gamification elements, caused many participants to increase the amount of biking trips in their daily routine and it can be assumed that a considerable share of 26% of participants will keep their increased levels of bike commuting in the long run.

Which groups can be identified within participants of by gamified biking initiatives to increase biking on their work trips?

It turned out that the effect of raised awareness of individual mobility habits is of importance, particularly for participants who have rarely biked before enrolling in an initiative. In both case studies participants were required to observe their own behavior by collecting data on their daily mode choices, which eventually disclosed previously unconscious behavior patterns. This enabled participants to deliberately reflect on their mobility habits and discover unexpected qualities of alternative modes of transport. This novel awareness paved the way for perceiving the value of specific benefits connected to biking. In the case of the "Bike to Work" initiative the two main drivers the participants discovered were health gains and social benefits (team spirit).

Which methods can be used for assessing the potential short- and long-term impact on mobility behavior?

Regarding methodological aspects for assessing the impact of biking initiatives, the two case studies show that both quantitative and qualitative methods need to be combined for describing and understanding behavior changes. The "Frequent Biking Challenge" provides an approach for controlled assessment of biking share changes based on trip data delivering valid interpretations of actual behavior change. Applying its methods to a large-scale randomized controlled trial would allow to draw conclusions on the causal effect of an intervention on actual bike usage. The combination of quantitative and qualitative methods for evaluating the "Bike to Work" campaign provides a deeper understanding on the social processes that are introduced by the campaign and to have a better interpretation of quantitative survey data. The experiences presented in this paper should be taken up in future research for evaluating comparable initiatives in order to contribute to a better understanding of the potential effects of soft policy measures, in particular regarding approachable target groups (which social groups can be motivated in which way) and achievable modal shifts (which modes will be replaced by biking).

CONCLUSIONS AND OUTLOOK

In summary, this work showed that gamified biking campaigns can be very effective for raising awareness for biking and encouraging participants increase their level of cycling. Furthermore, the importance of social elements that foster mutual encouragement for adopting and maintaining new behavior was examined.

Future work must explore the quantifiable effects of gamified interventions on mode shifts in more depth in order to answer a variety of crucial questions, e.g. how to robustly estimate the reduction of distances driven by cars or use of public transport. Furthermore, the data from "Bike to Work" shows that people commute by bike even though they usually use alternatives that are taking substantially less time. One explanation can be found in an offsetting effect of other motivational elements, such as health benefits. However, as this should be taken into consideration when transportation mode choices are discussed future research should look into the relativity of travel times as well.

The experiences drawn from the presented studies will be collated in order to prepare a large scale experiment that will employ gamified elements in a "Biking Tourney" within the the greater Boston area. The campaign will be scientifically evaluated using a combination of quantitative and qualitative methods. The results are expected to provide deeper insight into the social dynamics of creating/stimulating behavior change in the absence of any tangible incentives.

ACKNOWLEDGEMENTS

We gratefully acknowledge Sandra Richter and Chengzen Dai (all MIT) for their contribution in the "Persuasive Urban Mobility" project including the "Frequent Biking Challenge", and Geraldine Fitzpatrick (TU Vienna) and Kent Larson (MIT) for their advice and support throughout this research. Furthermore, we thank Francesco Pilla (Trinity College Dublin) and Norbert Brändle

(AIT) for their feedback on this paper. We thank Alexander Roth (AIT) for his support in the desk research regarding bike-promoting campaigns. Special thanks go to the Austrian biking advocacy group Radlobby, which is organizing the "Bike to Work" campaign for supporting the presented evaluation study.

REFERENCES

1. Richter, J., M. Friman, and T. Gärling. Soft Transport Policy Measures: Gaps in Knowledge. *International Journal of Sustainable Transportation*, Vol. 5, No. 4, 2011, pp. 199–215.

2. Ogilvie, D. Promoting walking and cycling as an alternative to using cars: systematic review. *BMJ*, Vol. 329, No. 7469, 2004, pp. 763–0.

3. Millonig, A., and K. Mitgutsch. Playful Mobility Choices: Motivating informed mobility decision making by applying game mechanics. *ICST Transactions on Ambient Systems*, Vol. 1, No. 4, 2014, p. e3.

4. Pucher, J., R. Buehler, and M. Seinen. Bicycling renaissance in North America? An update and re-appraisal of cycling trends and policies. *Transportation research part A: policy and practice*, Vol. 45, No. 6, 2011, pp. 451–475.

5. Rose, G., and H. Marfurt. Travel behaviour change impacts of a major ride to work day event. *Transportation Research Part A: Policy and Practice*, Vol. 41, No. 4, 2007, pp. 351–364.

6. Crawford, S. *Promoting active transport to school : evaluation of the Ride2School program*. Deakin University, 2011.

7. Piatkowski, D., R. Bronson, W. Marshall, and K. J. Krizek. Measuring the Impacts of Bike-to-Work Day Events and Identifying Barriers to Increased Commuter Cycling. *Journal of Urban Planning and Development*, 2014.

8. Hamari, J., J. Koivisto, and H. Sarsa. Does gamification work?–a literature review of empirical studies on gamification. 2014.

9. McCall, R., V. Koenig, and M. Kracheel. Using Gamification and Metaphor to Design a Mobility Platform for Commuters: *International Journal of Mobile Human Computer Interaction*, Vol. 5, No. 1, 2013, pp. 1–15.

10. Jylhä, A., P. Nurmi, M. Sirén, S. Hemminki, and G. Jacucci. Matkahupi: a persuasive mobile application for sustainable mobility. 2013.

11. Ecker, R., P. Holzer, V. Broy, and A. Butz. EcoChallenge: a race for efficiency. 2011.

12. Wunsch, M., A. Stibe, A. Millonig, S. Seer, C. Dai, K. Schechtner, and R. C. C. Chin. What Makes You Bike? Exploring Persuasive Strategies to Encourage Low-Energy Mobility. In

Persuasive Technology (T. MacTavish and S. Basapur, eds.), Springer International Publishing, 2015, pp. 53–64.

13. Stibe, A. Towards a Framework for Socially Influencing Systems: Meta-Analysis of Four PLS-SEM Based Studies. In *Persuasive Technology* (T. MacTavish and S. Basapur, eds.), Springer International Publishing, 2015, pp. 172–183.

14. Shapiro, S. S., and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 1965, pp. 591–611.

15. Braun, V., and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, Vol. 3, No. 2, 2006, pp. 77–101.

16. Clopper, C. J., and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 1934, pp. 404–413.

17. Brown, L. D., T. T. Cai, and A. DasGupta. Interval Estimation for a Binomial Proportion. *Statistical Science*, Vol. 16, No. 2, 2001, pp. 101–133.

18. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.

Persuasive Backfiring: When Behavior Change Interventions Trigger Unintended Negative Outcomes

Agnis Stibe[1(✉)] and Brian Cugelman[2,3]

[1] MIT Media Lab, Cambridge, MA, USA
agnis@mit.edu
[2] Statistical Cybermetrics Research Group,
University of Wolverhampton, Wolverhampton, UK
brian@alterspark.com
[3] AlterSpark, Toronto, ON, Canada

Abstract. Numerous scholars study how to design evidence-based interventions that can improve the lives of individuals, in a way that also brings social benefits. However, within the behavioral sciences in general, and the persuasive technology field specifically, scholars rarely focus-on, or report the negative outcomes of behavior change interventions, and possibly fewer report a special type of negative outcome, a backfire. This paper has been authored to start a wider discussion within the scientific community on intervention backfiring. Within this paper, we provide tools to aid academics in the study of persuasive backfiring, present a taxonomy of backfiring causes, and provide an analytical framework containing the intention-outcome and likelihood-severity matrices. To increase knowledge on how to mitigate the negative impact of intervention backfiring, we discuss research and practitioner implications.

Keywords: Backfire · Taxonomy · Behavior change · Intention-outcome matrix · Likelihood-severity matrix · Persuasive technology · Intervention design

1 Introduction

Scholars have focused on the ways in which technology can produce positive outcomes, such as increasing users' physical activity [29], reducing binge drinking [10], quitting smoking [26], or managing mood and anxiety disorders [14]. There is considerable research on this topic, with several systematic reviews and meta-analyses that focus on a wide variety of positive outcomes [9, 38].

However, few papers report negative outcomes, and possibly fewer report a special class of negative outcomes, called a *backfire*, which we define as an intervention that triggers audiences to adopt the opposite target behavior, rendering the intervention partially responsible for causing the behavior it was designed to reduce or eliminate.

Examples of backfiring interventions include drug use reduction programs that trigger drug use by accidently creating a social norm that triggers some youth to feel like everyone else is trying drugs except for them; traffic safety campaigns that use

A. Meschtscherjakov et al. (Eds.): PERSUASIVE 2016, LNCS 9638, pp. 65–77, 2016.
DOI: 10.1007/978-3-319-31510-2_6

shame which accidently triggers denial and resentment which leads to increased dangerous driving; binge drinking screeners that trigger some youth who drink less than average to feel pressured to catch up to their peers; or a tobacco industry sponsored anti-smoking campaign that encouraged parents to lecture their children on not smoking which triggered more youth to smoke [20].

In this paper, we aim to start a wider scientific discussion on intervention backfiring, to provide analytical frameworks to help structure this discussion, to support academic research on the topic, and to raising practitioners' awareness of the potential risks of intervention backfires so they can better manage the potential risks.

2 Background

Numerous fields provide recommendations on how to design behavior change interventions, including social marketing [1, 25], evidence-based behavioral medicine [12, 13], several health behavior change approaches [2, 7, 33], socially influencing systems [36, 37], persuasive technology [15], and classic persuasion literature [32].

However, the application of evidence-based intervention design frameworks does not guarantee that intervention designers will achieve positive outcomes. In practice, scientific models are more likely to inspire interventions, rather than to dictate exactly how they are implemented [31].

Quite often, interventions that start-out with a solid theoretical underpinning, lose their theoretical roots after adapting to real-world necessities, implementation complexities, budget limits, stakeholder feedback, market testing, political tampering, etc.… Then after moving through the process of translating abstract behavioral science principles into concrete intervention materials, such as translating the health belief model into a health app, it can be difficult to make clear links between the foundational theory and applied interventions.

No matter how promising an intervention may have appeared on paper, in practice, they always have the potential to exert unforeseen, and possibly negative outcomes. This is one of the reasons why new behavior change interventions typically undergo stringent monitoring and evaluation, before they are widely disseminated.

Scholars and practitioners, who report that they have disseminated a backfiring technology, can easily feel embarrassed, or worse, find themselves not just stigmatized, but potentially unfunded. Without doubt, there are many practical incentives and disincentives that may motivate people to not publish any information that details how their digital interventions backfired, causing even a small degree of unintended harm, even in marginal population.

We believe that this stigma has created a climate where the existing body of scientific literature may possess a significant degree of publication bias, resulting from too many published studies that only report positive outcomes, and too few studies reporting negative outcomes. This stigma has the potential to create a climate where scientists and practitioners are at greater risk of disseminating harmful interventions.

3 Framework

To aid discussions on intervention backfiring, this section presents two matrices that define categories of backfires while clarifying their potential to undermine the efficacy of behavior change interventions. Figure 1 presents the intention-outcome matrix, which describes different types of outcomes.

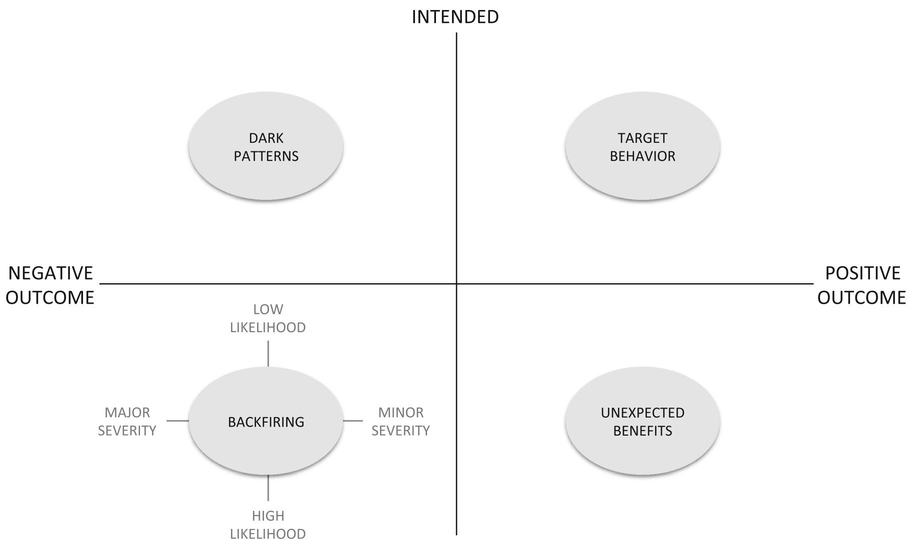


Fig. 1. Intention-Outcome matrix

3.1 Axes of the Intention-Outcome Matrix

The intention-outcome matrix has two axes with four quadrants, which are described below, and illustrated through a fictional binge-drinking reduction intervention.

- **Intended Outcome.** An outcome that was intended by the intervention designer. For example, getting university students to reduce their binge-drinking.
- **Unintended Outcome.** Any outcome that the intervention designer did not intend, whether it is a positive or negative outcome. For example, when a binge-drinking screener accidently motivates some students to adopt other health outcomes, or triggers some students to drink more alcohol after they use it in a competition to see who can achieve the highest binge-drinking score.
- **Positive Outcome.** An outcome that serves the interest of both the intervention designer and the target audience. This is a win-win situation, where both parties benefit. For example, a binge-drinking screener where the target audience achieves reduced alcohol consumption.

- **Negative Outcome.** An outcome that does not serve the interest of the target audience. For example, a binge-drinking screener that causes small segments to drink more alcohol, or exposes them to greater risk.

3.2 Quadrants of the Intention-Outcome Matrix

The intention-outcome matrix has four quadrants, that describes the four types of outcomes, as described below.

- **Target Behavior.** The primary intended positive behavioral outcome being sought, and typically reported.
- **Unexpected Benefits.** A positive behavioral outcome that was not intentionally sought, but which was positive nonetheless, and may be reported as a complimentary benefit of the intervention.
- **Backfiring.** This category includes a number of negative outcomes, when an intervention causes the opposite of the outcome (e.g. more binge drinking instead of less). It also includes “side effects”, when the primary behavior is achieved, but it also triggers unintended negative outcomes (e.g. using peer pressure to influence behavior which lowers audiences’ self-esteem). This quadrant is further subdivided into a risk management matrix that contrasts the likelihood of backfiring (low to high) with the potential severity (minor to major).
- **Dark Patterns.** When an intervention is used for the benefit of the developer, at the expense of the target audience. This is in the realm of unethical applications, such as coercion, deception, and fraud. For instance, scholars draw attention to dark game design patterns that developers use to provide negative experiences to players, not in their best interests, and without their consent [24, 27].

4 Method

The taxonomy of persuasive backfires presented in this paper was derived through a grounded theory methodology [6, 17], based on a corpus of academic, applied, and personal experiences with backfiring behavior change interventions [3].

We began the process by defining and limiting our selection criteria to interventions that backfire, that cause the opposite behavior, or unanticipated negative consequences that were contrary to the intentions of the program. We excluded flawed interventions that did not contribute towards any outcomes, positive or negative, as these programs constituted poor implementations, not backfires. The first type of flawed intervention includes programs that audiences did not find motivating, as these programs lacked the capacity to achieve significant outcomes. Similarly, we excluded interventions that faced implementation barriers, or created barriers among its target audiences, as it was not possible to assess the outcomes of these interventions.

To gather qualifying sources, we ran a call for references and examples across several academic, professional, and personal networks. The types of references we collected included journal papers, articles, program evaluations, and personal experiences. In total, we collected 47 responses.

We systematically reviewed all sources, and only included submissions that qualified as having demonstrated an unintended negative outcome, or which were presented as having contributed an unintended negative outcome. We also included behavior change interventions that were not implemented within technology per se, but were reasonably implemented in a context that could be applied to online behavior change campaigns or digital products. We also received submissions of backfiring legislation, which is often used in conjunction with communication campaigns, to elicit social change. In total 30 responses were included in our qualitative analysis.

We carried out a qualitative assessment of the corpus, with a view to developing taxonomy of triggers for backfiring interventions.

5 Findings

The findings of our study are presented in Fig. 2, a risk management framework called the likelihood-severity matrix. The four quadrants of backfires include (A) low likelihood and minor severity, (B) high likelihood and minor severity, (C) low likelihood and major severity, and (D) high likelihood and major severity. Through our qualitative assessment of backfires and their potential risks, we have synthesized twelve types of backfires, within six categories: inexperience, fineprint fallacy, personality responses, credibility damage, poor judgment, and social psychology.

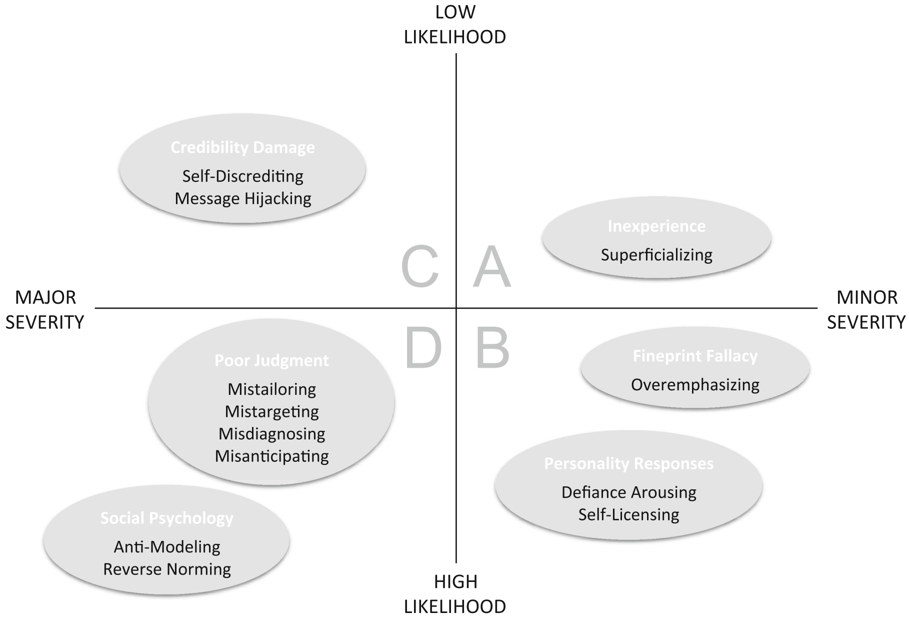


Fig. 2. Likelihood-Severity matrix

The taxonomy of persuasive backfires, presented in Table 1, provides twelve types of backfires within six broad categories.

Table 1. Taxonomy of persuasive backfires

Description	Examples
A: LOW likelihood and MINOR severity	
Inexperience	
Superficializing	
The superficial application of theory, such as copying surface tactics without understanding the underlying strategies, principles, or reasoning.	Research on social media, particularly on Foursquare, looks at driving user motivation through gamification mechanics that compete with intended target behaviors of information sharing [23]. More generally, research that focuses on how gamification elements can drive extrinsic motivation, may cease to be persuasive over the long term, and may potentially deplete intrinsic motivation.
B: HIGH likelihood and MINOR severity	
Fineprint Fallacy	
Overemphasizing	
Motivate people to take action for one strongly emphasized benefit, while omitting (or hiding) harmful factors that are in the fine print.	Stressing "low fat", while still including several unhealthy factors, such as high-sugar, or trans fats. Asking people to take one pill a day to lower their risk of contracting HIV, which causes some groups to engage in riskier sex (and stop using condoms), contributing to an increase in other sexually transmitted diseases [22].
Personality Responses	
Defiance Arousing	
Resistance to messages that are incompatible with a person's self-identity, that can induce unpleasant cognitive dissonance, leading the audience to reject the message, or oppose it.	Guilt and shame messages in anti-drinking ads for drunk drivers are ignored, and in some cases, backfire, when the message is incompatible with how individuals view themselves. Persuasive messaging involving an authority might provoke opposition from groups and individuals who are sensitive or resistant to authorities. One anti-smoking campaign failed due to advocating the message that "teens shouldn't smoke… because they're teens." Youths in the 10th-12th-grade were 12 percent more likely to smoke for each parent-targeted ad they had seen in the last

(*Continued*)

Table 1. (*Continued*)

Description	Examples
	30 days. According to developmental psychologists, teens 15 to 17 years old tend to reject authoritative messages because they believe they are independent, which renders Philip Morris' ad campaign largely useless [16]. The presence of peer information decreases the savings of nonparticipants who are ineligible for automatic enrollment in a saving plan, while higher observed peer savings rates also decreased savings. Discouragement from upward social comparison seems to drive this reaction [4].
Self-Licensing	
When someone does something good in one area, they sometimes feel like they have a license to misbehave in other.	After donating to charity, people may feel licensed to behave less morally in subsequent decisions [30]. For instance, donating to charity may have a dark side to it, as it negatively affects subsequent, seemingly unrelated moral behavior, such as the intention to be environmentally friendly.
C: LOW likelihood and MAJOR severity	
Credibility Damage	
Self-Discrediting	
When the source disseminates discrediting information, causing a misalignment of source and message credibility.	People report less favorable thoughts and attitudes towards a source, after reading weak arguments presented by a high vs. low expertise source. Too much fear mongering may discredit a campaign to the point of disbelief or humor [21].
Message Hijacking	
Third party actors re-contextualize the source's message, bringing a new meaning, which in many cases undermines the intervention by turning it into a public joke.	Creative works designed to cause fear, become a trendy meme with a different meaning, such as humorous cigarette ads of smoking children, reefer madness, fashionable heroin chic [19], or the ad campaign for TV PSA "This is your brain and this is our brain on drugs" which triggered numerous parodies [7]. England's Beat Bullying Campaign triggered bullying and violence. The campaign was so popular at its launch that supplies of the "Beat Bullying" wristbands quickly sold

(*Continued*)

Table 1. (*Continued*)

Description	Examples
	out. Because of the scarcity of the bracelets, and theme of the campaign, some kids were bullied for wearing the bracelet [11].
D: HIGH likelihood and MAJOR severity	
Poor Judgment	
Mistailoring	
When a tailored messaging system provides information that produces negative outcomes in some users, that could have been avoided with an appropriate message.	A drinking screener showed both low and high drinking students how much they consume in comparison to an average consumption. Those that were above the norm felt encouraged to drink less, while those below received an implied message to drink more. And a boomerang effect [35], when a descriptive social norm was not accompanied by an injunctive social norm in a similar way as described above.
Mistargeting	
When a message that was intended for one audience segment is misinterpreted by another group of people.	A one-size-approach to persuasive messaging can deter healthy eating behavior change, leading to a negative change in attitudes towards healthy eating over time. The Playpump program was meant for children to pump water while playing, but resulted in adults using the playground pump, leading to back injuries and other health problems [5].
Misdiagnosing	
When a behavior change intervention does not properly diagnose user behavior or psychological processes.	Gaze tracking software, designed to provide proactive help as patients read medical documents, used fixation time as a cue to identify when users were struggling with the material, and might need help. During the trial, the participants with low health literacy had a slower reading rate, causing the system to inappropriately offer help continually, which just annoyed the users, leading to lower comprehension compared to the control condition.
Misanticipating	
Changes in policies or directives that lead to unanticipated shifts in beliefs, attitudes or behaviors.	In the Netherlands, the drinking age was changed from 16 to 18, which news stations linked to an increase in drug use among citizens in this age group.

(*Continued*)

Table 1. (*Continued*)

Description	Examples
	Nebraska's "Give Us Your Troubled Child" law backfired, as the Nebraska Safe Haven law enabled parents to drop off kids of any age and the state for the state to take in, resulting in some parents who dropped off grownups [Gra2008].
Social Psychology	
Anti-Modeling	
Demonstrating negative behavior, which exposes people to memory triggers of bad behaviors or temptations, potentially in moments of the greatest susceptibility.	Being exposed to others experiencing the stress of quitting smoking, for example, triggers people to want to smoke. Anti-bad-behavior interventions can remind people of the bad behavior, thus potentially spark their motivation. A cookie company that introduced a 100-calorie packs of snacks, triggered people to eat far more of the small snack packs [28].
Reverse Norming	
Interventions that use examples of popular bad behaviors can establish the bad behavior as a social norm.	An anti-littering program used campaign posters to stress how widespread littering was, accidently contributed to a social norm for littering, by demonstrating how many people were littering. Two separate studies indicated that D.A.R.E. program was ineffective and in some cases, pushed kids toward drug use and lowered self-esteem. Researchers suspected that the intervention's message made some kids want to try drugs as a way of fitting in, stating and the program's message could be misinterpreted by youth as conveying that "peer pressure is around every corner, because everyone is doing drugs but you!" [34].

6 Scientific Considerations

We have authored this paper to initiate a scientific discussion on intervention backfiring, and to encourage scholars to examine this phenomenon in greater depth. We hope this will contribute to more transparency within the academic community, leading to more well-rounded research on persuasive design and its application.

The scientific contribution of this paper includes the persuasive backfiring framework, its two matrices (intention-outcome and likelihood-severity), and taxonomy. These frameworks can be used to define, discuss, and further research behavior change interventions that trigger unintended negative outcomes. Although we have not discussed ethics in this paper, it provides a system to further define ethical and unethical uses of persuasive technology.

Perhaps the most important contribution of this paper, is raising awareness of the elephant in the room, the well-known but rarely discussed fact that applied interventions targeting good outcomes, occasionally produce negative outcomes. It also raises questions on the social stigma attached to the people and organizations that deploy backfiring interventions, and how this may contribute to under reporting negative outcomes, or omitting them altogether.

More concerning, this investigation has identified a potentially large source of publication bias, as we believe that there are many incentives and disincentives, including stigma and embarrassment, that motivate researchers to avoid publishing research on negative outcomes.

We believe that scholars will need to develop innovative new research methods to overcome the research barriers that surround this subject, due to the stigma associated with reporting on interventions that backfire, and the limited ability of scientific studies to identify backfires that are more likely to become apparent in evaluations of real-world interventions, and often informally known by intervention staff but not publicly reported.

For instance, given the ability of technology to employ tailoring techniques, where content can be personalized, persuasive technology scholars are better equipped to undertake research on backfiring psychology, and use this knowledge to advise intervention designers when they need to omit influence principles that may be counterproductive to particular segments.

7 Practitioner Considerations

To assist practitioners, the taxonomy and tools presented in this paper provides a list of risks that can be addressed, if identified before they occur. Moreover, these risks can be used to build more effective interventions, by teaching practitioners how to avoid, mitigate, or manage backfires. The examples in this study suggest that backfires may originate from political tampering during the intervention design phase, evaluations that do not distinguish between groups, overusing a principle to the point of triggering mistrust in a message ("your brain on drugs"), misdiagnosis that leads to subscribing the wrong intervention leading to garbage-in-garbage-out interventions.

However, one ethically questionable practice we discovered was the potential intentional use of backfiring as a dark pattern, designed to deploy an intervention that superficially appears to promote a healthy behavior, but which actually promote unhealthy behavior. When corporations are obliged, or volunteer to carry out public health interventions to warn the public against their product, these corporations can easily benefit from the intentional use of backfiring interventions.

For instance, the "Talk: They'll Listen" campaign is frequently cited as an example of a cleaver antismoking ad campaign that on the outside appeared to be a legitimate antismoking campaign, but which in practice caused an increase in youth smoking [19]. Consequently, policy makers and regulators who empower tobacco, alcohol, and pharmaceutical companies to run their own interventions, need to be fully informed about the potential intentional use, and abuse of backfire-based campaigns, that superficially look effective, but at a deeper level, have been engineered to encourage the opposite effect.

Finally, we believe that backfires are normally present to some degree in all interventions, and that they can never be fully eliminated. It may be more practical to focus on strategies to reduce the riskiest backfires, to manage those that cannot be fully eliminated, and to continually monitor and improve interventions over time.

8 Conclusions

The stigma associated with reporting behavior change interventions that trigger negative outcomes, has relegated the topic of intervention backfiring to an informal observation that is widely known, but rarely discussed or reported. This has created a climate where scholars routinely overemphasize positive outcomes, while failing to report the fact that the same principle, can also lead to unforeseen negative outcomes.

In this paper, we discussed multiple ways how behavior change interventions can backfire. We provided a framework to help facilitate the discussion of this topic, presented tools to aid academics in the study of this realm, and offered advice to practitioners about potential risks. We encourage researchers to build on this work, and take a more systematic look on approaches involving the design of behavior change interventions.

In the future, researchers will need to innovate new ways to study this subject, and extend our scientific and practical knowledge of what pitfalls need to be avoided when designing technology-supported behavior change interventions. We advocate that researchers and practitioners adopt an honest and open attitude towards identifying and removing backfires as soon as possible, and to disseminating strategies to reduce their occurrence, before they cause more harm than good.

References

1. Andreasen, A.: Social marketing in the 21st century. Sage Publications Inc, Thousand Oaks (2006)
2. Bartholomew, L.K., Parcel, G.S., Kok, G., Gottlieb, N.H.: Planning health promotion programs: an intervention mapping approach. Wiley, San Francisco (2011)
3. Beer, M., Eisenstat, R.A., Spector, B.: Why change programs don't produce change (1990)
4. Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C., Milkman, K.L.: The effect of providing peer information on retirement savings decisions. J. Finan. **70**(3), 1161–1201 (2015)
5. Chambers, A.: Africa's not-so-magic roundabout (2009). http://www.theguardian.com/commentisfree/2009/nov/24/africa-charity-water-pumps-roundabouts
6. Charmaz, K.: Constructing grounded theory. Sage, Thousand Oaks (2014)
7. Chatterjee, S., Price, A.: Healthy living with persuasive technologies: Framework, is- sues, and challenges. J. Am. Med. Inf. Assoc. **16**(2), 171–178 (2009)

8. Crano, W.D., Burgoon, M., Oskamp, S. (eds.): Mass media and drug prevention: Classic and contemporary theories and research. Psychology Press, Mahwah (2001)
9. Cugelman, B., Thelwall, M., Dawes, P.: Online interventions for social marketing health behavior change campaigns: A meta-analysis of psychological architectures and adherence factors. J. Med. Internet Res. **13**(1), e17 (2011). http://doi.org/10.2196/jmir.1367
10. Cunningham, J.A., Wild, T.C., Cordingley, J., Van Mierlo, T., Humphreys, K.: A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction **104**(12), 2023–2032 (2009)
11. Curtis, P.: Anti-bullying wristband scheme backfires (2004). http://www.theguardian.com/education/2004/dec/08/schools.uk2
12. Davidson, K., Goldstein, M., Kaplan, R., Kaufmann, P., Knatterud, G., Orleans, C., Whitlock, E.: Evidence-based behavioral medicine: What is it and how do we achieve it? Ann. Behav. Med. **26**(3), 161–171 (2003)
13. Embry, D., Biglan, A.: Evidence-based kernels: Fundamental units of behavioral influence. Clin. Child Family Psychol. Rev. **11**(3), 75 (2008)
14. Farvolden, P., Denisoff, E., Selby, P., Bagby, R.M., Rudy, L.: Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. J. Med. Int. Res. **7**(1), e7 (2005). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1550639/
15. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann, San Francisco (2003)
16. Fraser, J.: Anti-smoking ads cleverly boost smoking among teens (2006). http://www.naturalnews.com/020996.html
17. Glaser, B.G., Strauss, A.L.: The discovery of grounded theory: Strategies for qualitative research. Transaction Publishers, New Jersey (2009)
18. Graham, J.: Father leaves nine children at Nebraska hospital (2008). http://newsblogs.chicagotribune.com/triage/2008/09/father-leaves-n.html
19. Hastings, G.: Social marketing: why should the devil have all the best tunes?. Butterworth-Heinemann, Oxford (2007)
20. Healey, B., Zimmerman, R.S.: The new world of health promotion: New program development, implementation, and evaluation. Jones & Bartlett Learning, Sudbury (2009)
21. Hinkley, K.: 5 Ridiculous Anti-Drugs Posters (2014). http://www.talkingdrugs.org/5-anti-drugs-campaigns
22. Holpuch, A.: Truvada has been called the ‘miracle’ HIV pill–so why is uptake so slow (2014). http://www.theguardian.com/world/2014/sep/18/truvada-mircle-pill-prevent-hiv-controversy
23. Kietzmann, J.H., Hermkens, K., McCarthy, I.P., Silvestre, B.S.: Social media? Get serious! Understanding the functional building blocks of social media. Bus. Horiz. **54**(3), 241–251 (2011)
24. Kirkland, A., Metzl, J.M. (eds.): Against health: How health became the new morality. NYU Press, New York (2010)
25. Kotler, P., Roberto, N., Lee, N.: Social marketing: improving the quality of life, 2nd edn. Sage Publications Inc, California (2002)
26. Lenert, L., Munoz, R., Perez, J., Bansod, A.: Automated e-mail messaging as a tool for improving quit rates in an internet smoking cessation intervention. J. Am. Med. Inf. Assoc. **11**(4), 235–240 (2004)
27. Linehan, C., Harrer, S., Kirman, B., Lawson, S., Carter, M.: Games against health: a player-centered design philosophy. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 589–600. ACM (2015, April)

28. London, L.: Guiltless Girls: Unpacking 100-calorie Snacks (2009)
29. Marshall, A., Leslie, E., Bauman, A., Marcus, B., Owen, N.: Print versus website physical activity programs A randomized trial. Am. J. Prev. Med. **25**(2), 88–94 (2003)
30. Meijers, M.H., Verlegh, P.W., Noordewier, M.K., Smit, E.G.: The dark side of donating: how donating may license environmentally unfriendly behavior. Social Influence, 1–14 (2015). doi:10.1080/15534510.2015.1092468
31. Michie, S., Abraham, C.: Interventions to change health behaviours: evidence-based or evidence-inspired? Psychol. Health **19**(1), 29–49 (2004)
32. O'Keefe, D.: Persuasion: Theory and Research. Sage Publications Inc, London (2002)
33. Prochaska, J.O., Velicer, W.F.: The transtheoretical model of health behavior change. Am. J. Health Promot. **12**(1), 38–48 (1997)
34. Reaves, J.: Just Say No to DARE (2001). http://content.time.com/time/education/article/0,8599,99564,00.html
35. Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., Griskevicius, V.: The constructive, destructive, and reconstructive power of social norms. Psychol. Sci. **18**(5), 429–434 (2007)
36. Stibe, A.: Towards a framework for socially influencing systems: Meta-analysis of four PLS-SEM based studies. In: MacTavish, T., Basapur, S. (eds.) PERSUASIVE 2015. LNCS, vol. 9072, pp. 172–183. Springer, Heidelberg (2015)
37. Stibe, A.: Advancing typology of computer-supported influence: moderation effects in socially influencing systems. In: MacTavish, T., Basapur, S. (eds.) PERSUASIVE 2015. LNCS, vol. 9072, pp. 253–264. Springer, Heidelberg (2015)
38. Webb, T., Joseph, J., Yardley, L., Michie, S.: Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J. Med. Internet Res. **12**(1), e4 (2010). http://doi.org/10.2196/jmir.1376

Persuasive Cities: Health Behavior Change at Scale

Agnis Stibe

MIT Media Lab
agnis@mit.edu

Abstract. Can you imagine a city that feels, understands, and cares about your wellbeing? Future cities will reshape human behavior in countless ways. New strategies and models of urban spaces are required for creating future cities to properly respond to human activity, environmental conditions, and market dynamics. Persuasive urban systems will play an important role in making cities more livable and resource-efficient by addressing current environmental problems and enabling healthier routines. Drawing on socio-psychological theories and integrating them with new concepts for urban design, the persuasive cities research focuses on improving wellbeing across societies. This research presents an ecosystem of future cities, describes three generic groups of people depending on their susceptibility to persuasive technology, explains the process of defining behavior change, and provides tools for social engineering of persuasive cities. Further research should continue exploring how urban design in combination with socially influencing systems could encourage healthy and sustainable behaviors at scale.

Keywords: persuasive cities, social engineering, socially influencing systems, behavior change, wellbeing, health, persuasive technology

1 Perspective

As population in cities continue grow exponentially the architecture and design of future urban places will become more dominant in impacting human behavior. According to social cognitive theory [1], any well-designed environment can become a strong influencer of what people think and do. There is an endlessly dynamic interaction between a person, a particular behavior, and an environment in which that behavior is performed. The persuasive cities research leverages this knowledge to engineer persuasive environments for altering human behavior on societal levels.

The proposed research reflects on novel ways of how persuasive technology [2] and socially influencing systems [3-4] enable mechanisms to perpetually support motivation of individuals comparing to conventional methods, such as those that are based on carrots and sticks. Instead, persuasive urban systems harness social influence from crowd behavior to craft influential messaging aimed at shifting behavior and attitude of an individual, who naturally is an integral part of the same crowd. Such continuous interplay can ultimately result in an ongoing process that reshapes communities and societies without any other incentives.

2 Emergence of Persuasive Cities

Ongoing research streams focus on *sensible cities* (researching sensing technologies to read human behavior in urban spaces) and *smart cities* (analyzing big data to classify groups of people based on their distinct behavioral patterns), however there is a lack of knowledge about perspective ways to achieve persistent behavioral changes at scale. Therefore, the proposed research extends an ecosystem of future cities (Table 1) by introducing the notion of *persuasive cities* that aims to advance and refine influential strategies designed for intentionally reshaping how people think and act in urban environments.

Table 1. Ecosystem of future cities

Role	Character	Technology
	PERSUASIVE	
Change	Care	Socially Influencing Systems
	SMART	
Classify	Understand	Big Data Analytics
	SENSIBLE	
Read	Feel	Sensor Networks

Each *layer* of future cities has its *role*, *character*, and supportive *technology*. Sensible cities employ sensor networks to read crowd behaviors. In other words, these cities feel human movements. These crowd behaviors further serve as an input for big data analytics that smart cities apply to classify groups of people according to similar behavioral patterns (profiles). When that is accomplished, the groups having better routines can be exemplified to other underperforming groups through intentionally designed socially influencing systems, which are at the core of persuasive cities.

3 Susceptibility to Persuasive Technology

People generally can fall into one of the three generic categories depending on their susceptibility to persuasive technology (Fig. 1). *Self-contained* people (the red circle) most likely are not open for changing anything in them. They are fully satisfied with who they are and what they do on daily basis, thus many behavioral interventions might fail in attempts to influence this group of individuals. *Self-driven* people (the green circle) typically have comparatively high levels of motivation and can achieve everything that they have envisioned. Thus, these people most likely are not looking for additional sources of encouragement, and therefore persuasive technologies might become unnecessary for this group.

However, there is another group of people that oftentimes would like to change their routines, but rarely succeed in doing so. That reminds of New Year's resolutions that in many cases end around February. Therefore, this group is entitled as January 1st (the yellow circle) and seem to be the most welcoming towards technology supported behavioral interventions designed to help achieving target behaviors. Although, Fig. 1 presents all three groups as equal circles, in reality the size of each group might significantly vary depending on the context and particular behavior.

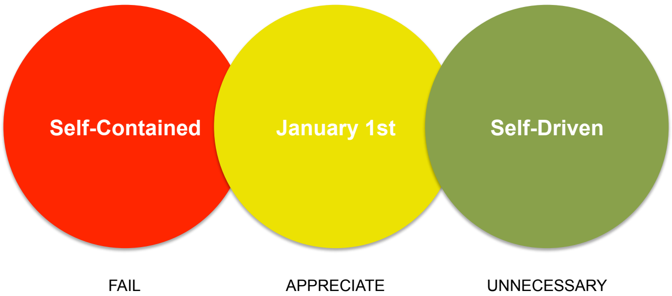


Fig. 1. Susceptibility to persuasive technology

4 Defining Behavior Change

To achieve an envisioned target behavior, the process and components of behavior change have to be well understood and clearly defined. In the process of defining behavior change, there are three main components, namely the target group, its present behavior, and its envisioned future behavior (Table 2).

Table 2. The three main components for defining behavior change

Target Group	Current Behavior	Future Behavior
	Description	
A group of people currently having an unsatisfactory behavior. It is important to narrow down the target group as precise as possible.	A certain behavior of the target group that currently is not in line with an envisioned future behavior in a given context.	An ultimate future behavior of the target group that is envisioned to be more beneficial for everyone.
	Example	
There are MIT faculty members.	Who currently commute alone in their private cars.	They could commute by bicycles instead whenever possible.

5 Tools for Social Engineering

Earlier research on persuasive technology [2] describes several ways how social dynamics can influence human behavior, which have been further refined and structured as a framework for Socially Influencing Systems (SIS) [3], depicted in Fig. 2. The SIS framework is a useful tool for scholars and practitioners aiming at improving future cities by introducing persuasive urban interventions targeted to support wellbeing.

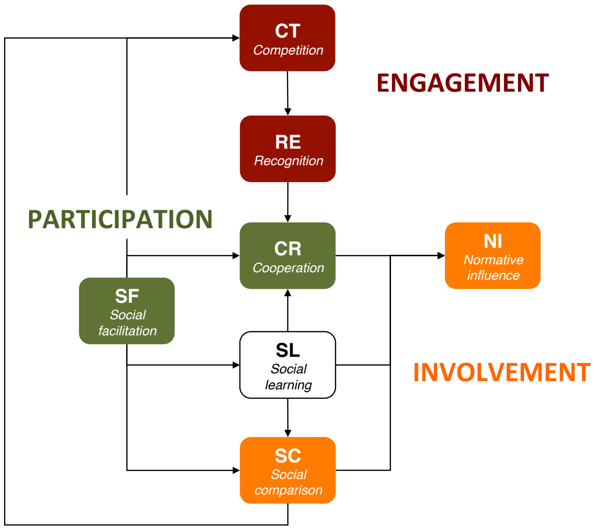


Fig. 2. Socially Influencing Systems (SIS) framework

The framework describes seven socially influencing principles that can support persuasive urban interventions. The principles are interlinked and have potential to exert stronger effects depending on the context of a particular behavioral challenge. Normative influence and social comparison seem to be more effective to achieve involvement of the target group as the two principles focus on attitudinal changes. Cooperation and social facilitation seem to be more effective to make individuals participate and do the envisioned future behavior even without a formed attitude towards it. Competition and recognition seem to be more effective in engaging the target group to do the future behavior as the principles focus on both attitude and behavior simultaneously. For example, the effects several socially influencing principles have already been studied in the context of urban mobility, e.g. bicycling [5].

6 References

1. Bandura, A.: Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice Hall, Englewood Cliffs (1986)
2. Fogg, B. J.: Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann (2003)
3. Stibe, A.: Towards a Framework for Socially Influencing Systems: Meta-Analysis of Four PLS-SEM Based Studies. In: MacTavish, T., Basapur, S. (eds.) Persuasive Technology. LNCS, vol. 9072, pp. 171–182. Springer, Heidelberg (2015)
4. Stibe, A.: Advancing Typology of Computer-Supported Influence: Moderation Effects in Socially Influencing Systems. In: MacTavish, T., Basapur, S. (eds.) Persuasive Technology. LNCS, vol. 9072, pp. 251–262. Springer, Heidelberg (2015)
5. Wunsch, M., Stibe, A., Millonig, A., Seer, S., Dai, C., Schechtner, K, and Chin, R.C.C. What Makes You Bike? Exploring Persuasive Strategies to Encourage Low-Energy Mobility. Lecture Notes in Computer Science, 9072, Persuasive Technology, pp. 53-64. (2015)

Empowering Cities for Sustainable Wellbeing

Agnis Stibe[1], Samir Chatterjee[2], Katja Schechtner[1,4], Matthias Wunsch[3,5], Alexandra Millonig[3], Stefan Seer[3], Ryan C.C. Chin[1], Kent Larson[1]

[1] MIT Media Lab, Cambridge, MA, USA
{agnis,katjas,rchin,kll}@mit.edu
[2] Claremont Graduate University, Claremont, CA, USA
samir.chatterjee@cgu.edu
[3] Austrian Institute of Technology, Vienna, Austria
{Matthias.Wunsch.fl,Alexandra.Millonig,Stefan.Seer}@ait.ac.at
[4] University of Applied Arts, Vienna, Austria
[5] Vienna University of Technology, Human Computer Interaction, Vienna, Austria

Abstract. Quality of life in cities can be improved through reshaping and advancing urban spaces with seamless persuasive and socially influencing strategies for empowering people to succeed in achieving better lifestyles. This vision aims at helping people to acquire healthy and resource-efficient everyday routines, thus leading to prosperous and sustainable societies. Urban engineers and technology developers are oftentimes unaware of how diversely their innovations will affect lives of many people. Therefore, this research is focused on investigating and designing ways how urban environments can be retrofitted and complemented with persuasive technology and socially influencing systems to facilitate societal changes at scale. Outcomes of this research are instrumental for various contexts, including health, mobility, education, energy and water conservation, safety, emergency management, ecology, and economy. Ultimately, more refined scientific knowledge on how to design empowering cities has to be generated and translated into applicable guidelines for practice to foster their emergence.

Keywords: persuasion, empowerment, socially influencing systems, wellbeing, mobility, behavior change, sustainability, gamification, persuasive technology

1 Motivation

Future cities will reshape human behavior in countless ways. Persuasive urban systems will play an important role in making cities more livable and resource-efficient by addressing current environmental problems and enabling healthier routines. Therefore, future research should be directed towards exploring how urban design might be combined with persuasive technology [3] and socially influencing systems [4-5] to encourage healthy behaviors at scale.

More effort has to be put into studying how quality of life and the health of the individual and communities might be improved through the creation of empowering cities, streets, buildings, homes, and vehicles. Information technology and computer systems are increasingly designed to support everyday routines and advance user experience in multiple ways [2]. Novel computer systems can be also intentionally designed to influence how users think and behave. Theories of persuasion and social influence provide various strategies for the developers of such systems to facilitate desired effects on users.

Research on empowering cities seeks to advance urban spaces to facilitate societal changes. This research is primarily focused on socially engaging environments for supporting entrepreneurship and innovation, reshaping routines and behavioral patterns in dense urban districts, intelligent outdoor sensing for shifting mobility modes, enhancing environmentally friendly behaviors through social norms, interactive public feedback channels for affecting attitudes, engaging residents through socially influencing systems [4-5], exploring methods for designing persuasive neighborhoods, testing agent-based models and simulations of persuasive interventions, and fostering adoption of novel urban systems.

2 Importance

The proposed research direction is highly important, as it will directly influence everyone living in future cities. Environmental, personal, and behavioral factors are locked into triadic reciprocal determinism [1], meaning that all three are strongly interconnected and continuously reshaping each other. Thus, environmental design is strong influencer on human behavior and attitude. In other words, quite often it is merely sufficient to improve urban spaces to help people become healthier and to create sustainable communities. This is very powerful vision as it encompasses transformation of human behavior and urban environments at scale. Moreover, since persuasion sometimes has a negative connotation, we will explore new ways to empower communities and cities.

The proposed research will reflect on novel ways of how socially influencing systems [4-5] enable perpetual mechanisms to foster user motivation as compared to conventional methods, such as those that are based on the principle of carrots and sticks. Earlier research on motivation discusses methods that have substantial limitations. For example, monetary incentives are mostly effective only as long they are provided, so people tend returning to their earlier behavior after the motivators are taken away. Instead, persuasive urban systems harness social influence from crowd behavior to craft influential messaging aimed at shifting behavior and attitude of an individual, who naturally is an integral part of the same crowd. Such continuous interplay can ultimately result in an ongoing process that reshapes communities and societies without any other incentives.

3 Agenda

The empowering cities research agenda is focused on reshaping and redesigning three main urban areas: outdoor environments, indoor environments, and mobility in cities.

Public spaces can be advanced in many ways, e.g. supermarkets can project a portion of how many healthy products have been purchased that day, week, or month. Responsive environments can use ambient lights to provide feedback about behavioral patterns of crowds. For example, streetlights can change color depending of how many joggers have been on that street on that morning. The window frames of residential buildings can be illuminated for those apartments, which have changed regular light bulbs to energy-efficient ones.

Computer-supported strategies [4] can be implemented indoors to motivate using stairs instead of an elevator. For example, a situated display that represents various comparisons of what can happens when stairs or an elevator is chosen. Strategies can be introduced to increase water intake in offices. For example, a situated display can present an increase of water consumption, which can be used to compete with other offices. New ways can be designed for office workers to increase socializing among individuals from various groups and departments. For example, specific game-like activities can be set up for employees to promote socializing.

Mobility within dense urban districts can be reshaped in multiple ways, for example, by introducing influential strategies to facilitate bicycle commuting. Street signage can be used to display how many bicyclists have ridden over a bridge today, for instance. Mobile apps can be developed to engage bicycle riders in reporting experiences with bike lanes and their quality in a selected urban area. Electric bicycles can be complemented with influential strategies to attract more riders and persuade them to pedal.

4 Application

There are various ways how the concepts of empowering cities can be designed and applied in urban contexts to support wellbeing. For example, it is necessary to design interventions for promoting bicycling and walking, because these mobility modes are carbon neutral, provide major health and financial benefits, and require less space for parking. Besides investing in infrastructure, cities can work on shifting mobility patterns towards bicycling and walking through publicly engaging urban interventions, especially designed leveraging persuasive technology [3] and socially influencing systems [4-5].

Earlier research traditionally reports how various interventions to promote modal shift can be effective. However, most of them rather follow one of the traditional approaches like publicity campaigns, engineering measures, or financial incentives. Therefore, more research should be done with regards to how behavioral and attitudinal changes can be achieved through persuasion and social influence [6]. Empowering cities have potential to significantly contribute to this effort, for example, through publicly displayed street signage with interactive computer-supported [4] social comparisons of cars versus bicycles.

To promote walking, city planners can potentially make many modifications to the urban environment. Besides meeting the requirements of safe walking pathways, there are other proven methods of fostering behavioral and attitudinal changes through social influence [5]. Persuasive urban systems can leverage the principles of normative influence, social learning, and social facilitation to affect the way people think about walking, which might presumably lead to increased physical activity. The principles of recognition, competition, and cooperation can be incorporated to build on the initial levels of walking and promote sustainable adoption.

Empowering cities can make walking experience more engaging, for example, by combining a mobile phone app that interacts with retrofitted traffic light junctions. So, when waiting at traffic light junctions, people would be invited to do certain meaningful activities. To represent the level of activity at each junction, each traffic light can be overlaid with an interactive color strip, which would display rankings on a particular day or week.

5 Future

Fundamentally new strategies must be found for creating the places where people live and work, and the mobility systems that connect these places, in order to meet the profound challenges of the future. Novel models for urban architecture and personal vehicles should be more responsive to the unique needs and values of individuals though the application of disentangled systems and smart customization technology. Technology has to be designed to understand and respond to human activity, environmental conditions, and market dynamics. The design of future cities requires optimal combinations of automated systems, just-in-time information for personal control, and interfaces to persuade people to adopt sustainable behaviors.

Drawing on socio-psychological theories [1] and integrating them with new concepts for urban design and technology [3], the proposed empowering cities research aims at advancing livability in future cities.

6 Literature

1. Bandura, A.: Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice Hall, Englewood Cliffs (1986)
2. Chatterjee, S. and Price, A.: Healthy Living with Persuasive Technologies: Framework, Issues, and Challenges. Journal of the American Medical Informatics Association (JAMIA) 16, 171–178 (2009)
3. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann (2003)
4. Stibe, A.: Advancing Typology of Computer-Supported Influence: Moderation Effects in Socially Influencing Systems. In: MacTavish, T., Basapur, S. (eds.) Persuasive Technology. LNCS, vol. 9072, pp. 251–262. Springer, Heidelberg (2015)
5. Stibe, A.: Towards a Framework for Socially Influencing Systems: Meta-Analysis of Four PLS-SEM Based Studies. In: MacTavish, T., Basapur, S. (eds.) Persuasive Technology. LNCS, vol. 9072, pp. 171–182. Springer, Heidelberg (2015)
6. Wunsch, M., Stibe, A., Millonig, A., Seer, S., Dai, C., Schechtner, K, and Chin, R.C.C. What Makes You Bike? Exploring Persuasive Strategies to Encourage Low-Energy Mobility. Lecture Notes in Computer Science, 9072, Persuasive Technology, pp. 53-64. (2015)

The Bologna Ringway Dataset: Improving Road Network Conversion in SUMO and Validating Urban Mobility via Navigation Services

Luca Bedogni, Marco Gramaglia, Andrea Vesco, *Member, IEEE*, Marco Fiore, *Member, IEEE*, Jérôme Härri, and Francesco Ferrero

***Abstract*—The current lack of reference datasets of road traffic mobility for network simulation jeopardizes the reliability and reproducibility of vehicular networking research. We contribute to the ongoing effort to develop dependable and publicly available mobility traces through the following: 1) implementing an original version of the Simulation of Urban Mobility (SUMO) road network conversion tool that allows importing OpenStreetMap (OSM) data in a neat automated fashion; 2) generating an original dataset of road traffic in Bologna, Italy; and 3) providing a novel validation methodology that builds on open data provided by navigation service, which we leverage to assess the quality of the proposed Bologna dataset. These three contributions are expected to benefit the whole research community, since they not only provide a new ready-to-use realistic dataset that can be input to network simulators but ease the generation and validation of further vehicular mobility traces for networking research as well.**

***Index Terms*—Data set, mobility traces, open source, road traffic, simulation, simulation of urban mobility (SUMO), validation methodology.**

I. Introduction

CONNECTED vehicles are at the center of a telecommunication revolution. On one hand, long-foreseen vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications are expected to hit the market soon: dedicated frequency bands have been allocated in Europe, the U.S., and Japan [1]; standards have been finalized [2]–[5]; and political actors and regulators are already crafting proposals enforcing all new vehicles to embed V2V radio interfaces [6]. On the other hand, the connected vehicle vision is a reality: cellular connectivity is an increasingly popular feature in cars, even budget models [7].

Many networking solutions and mobile services building on both V2V/V2I communication and connected vehicles are expected to operate at very large scales. This is the case of, e.g., opportunistic offloading [8], floating car data management [9], delay-tolerant message dissemination [10], or multihop routing [11].

Despite the deployment of a few first large-scale field tests [12], [13], the cost and the complexity of experimental studies force the performance evaluation of the aforementioned solutions to rely on simulation studies. Within this context, a reliable modeling of road traffic has been repeatedly proven to be paramount to the dependability of the results [14]–[16]. However, the availability of vehicular mobility datasets that are designed for networking purposes and that are publicly available is, today, limited to a few scenarios [17]–[19]. In turn, this inconvenient situation risks severely limiting the reliability and reproducibility of networking results [20].

The main reasons for the lack of reference mobility datasets are that 1) the tools to generate realistic road traffic are complex to configure and operate and that 2) real-world input data to be fed to such tools are hard to retrieve. Concerning the first point, specialized software, such as the widely adopted Simulation of Urban Mobility (SUMO) [22], typically has a steep learning curve, particularly when it comes to large-scale scenarios that involve high levels of detail and realism. This induces vehicular networking researchers to favor simple mobility scenarios in their simulations, at the cost of reliability. As for the second aforementioned item, a dependable road traffic simulation requires precise information on a number of aspects, including, e.g., the road network (down to per-lane settings or traffic light parametrization) and the trajectories of vehicles over it. These data are not easy to collect and correctly implement in the mobility simulation.

In this paper, we propose a threefold contribution to ease the development of a reference set of realistic ready-to-use datasets of vehicular mobility for networking studies, as follows.

- We improve the usability of software for the generation of road traffic datasets. Specifically, our focus is on a

Manuscript received April 15, 2015; revised July 15, 2015; accepted August 19, 2015. Date of publication September 1, 2015; date of current version December 14, 2015. The research leading to these results was supported by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n.630211 ReFleX. This work was also supported by the COLOMBO Project, which is funded, in part, by the European Commission under the Seventh Framework Program, FP7-ICT-2011-8, Grant 318622, as well as the support of its industrial members, namely, BMW Group Research and Technology, IABG, Monaco Telecom, Orange, SAP, SFR, ST Microelectronics, and Symantec. The review of this paper was coordinated by the Guest Editors.

L. Bedogni is with the Department of Computer Science and Engineering, University of Bologna, 40126 Bologna, Italy (e-mail: luca.bedogni4@unibo.it).

M. Gramaglia is with the Institute of Electronics, Computer, and Telecommunications Engineering (IEIIT), National Research Council (CNR), 10129 Torino, Italy (e-mail: marco.gramaglia@ieiit.cnr.it).

A. Vesco and F. Ferrero are with the Smart City Strategic Program, Istituto Superiore Mario Boella, Torino, Italy (e-mail: vesco@ismb.it; fferrero@ismb.it).

M. Fiore is with the Institute of Electronics, Computer, and Telecommunications Engineering (IEIIT), National Research Council (CNR), 10129 Torino, Italy, with Inria, 78153 Le Chesnay, France, and also with INSA Lyon and INRIA, 69621, Lyon, France (e-mail: marco.fiore@ieiit.cnr.it).

J. Härri is with EURECOM, 06904 Sophia Antipolis, France (e-mail: jerome.haerri@eurecom.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2015.2475608

key phase of the generation process, i.e., the conversion of road networks from map databases to mobility simulators. We consider state-of-the-art open-source tools, i.e., SUMO [22] and OpenStreetMap (OSM) [21], and propose some modifications to the NETCONVERT tool, which allows linking the two.

- We present an original methodology to validate synthetic road traffic datasets using easily accessible real-world information provided by navigation services. Our approach provides a rapid way to verify if fundamental features of the synthetic mobility, i.e., routing paths and travel times, are realistic enough. We demonstrate a specific implementation of the methodology, which is based on Google Maps Application Program Interfaces (APIs) [27].
- We apply our original road network converter and validation methodology to a use case scenario, i.e., a vast portion of the city of Bologna, Italy. The generation process employs realistic information on the macroscopic traffic flows in the region, which is collected by the iTetris European collaborative project [26]. The resulting dataset represents the mobility of more than 22 000 vehicles in a 25-km area during one traffic peak hour in a typical morning, and its features are shown to match well those inferred from navigation services. Coherently with the objective of building a reference set of mobility scenarios for vehicular networking purposes, we make this novel dataset open to the research community.[1]

These contributions foster the development and verification of new vehicular mobility datasets for networking research, by providing novel tools, guidelines, and methodologies that can be largely reused. We provide a proof of concept of their practical usage, by generating and analyzing an original dataset, which is made available to the research community.

This paper is organized as follows. After an introductory discussion of the generation process of synthetic road traffic, in Section II, we present our novel road network converter and compare it with legacy solutions in Section III. We apply our converter to a specific use case, i.e., the Bologna Ringway scenario, in Section IV. Using a realistic traffic-peak travel demand allows generating an original dataset. We validate such a scenario by leveraging a reusable methodology based on navigation service data, in Section V. Finally, Section VI presents related work, and Section VII concludes this paper.

II. Synthetic Vehicular Mobility: A Primer

The generation of dependable datasets of road traffic for networking studies requires bringing together several different tools and data sources.

Road Network: The first critical component is a comprehensive representation of the road network, which is not limited to the street layout, but includes detailed information on the nature of road segments (e.g., number of lanes per direction or speed limits) and road junctions (e.g., presence and synchronization of traffic lights or priorities among incoming roads).

Microscopic Mobility Models: The second element is represented by validated models of the driver's behavior, which describe, at a microscopic level, his decisions in terms of acceleration, deceleration, lane changing, and, generally, his reactions to the surrounding environment.

Macroscopic Traffic Flows: The third major ingredient is a faithful model of the road traffic flows in the target region, so that the resulting mobility is realistic not only at a microscopic level but from a macroscopic viewpoint as well. That implies gathering information on the traffic demand [i.e., the start time, the origin, and the destination of each vehicle's trip, stored in an origin–destination (O–D) matrix], as well as running appropriate traffic assignment techniques that determine the routes followed by each driver to reach his destination.

Validation: It is important that the synthetic mobility is validated by comparison against real-world data, including, e.g., traffic counts collected by induction loops, traffic volumes computed through cameras, or traffic flows from surveys or floating car data samples.

All such components need to be integrated via a federated software, which generates a trace of the movement of each vehicle in the considered scenario with a high spatial (order of meters) and temporal (order of hundreds of milliseconds) accuracy. We refer the interested reader to the detailed descriptions in [17] and [18] for a thorough presentation of the generation process of road traffic datasets for networking research.

In this paper, we focus on the current state-of-the-art federating tool, i.e., SUMO [22]. The specific problems we tackle are the first component mentioned earlier, i.e., the representation of the road network, and the last, i.e., the validation of the simulated road traffic.

III. Improved Road Network Conversion in Simulation of Urban Mobility

SUMO can import road network information from a number of sources, the most popular being, by far, OSM [21]. The OSM project provides crowdsourced maps of cities worldwide contributed by a vast user community, which are commonly regarded as the highest quality road data publicly available to date.

However, OSM is not designed for vehicular mobility simulation, and the conversion process is not trivial, as discussed in Section III-A. Since an incorrect representation of the road layout and features risks to impair the whole generation process, *a posteriori* manual corrections of conversion errors are often required. The latter are, however, extremely time consuming, and as a consequence, they are simply overlooked most of the time. To address this issue, we propose modifications to NETCONVERT, which make OSM-SUMO linkage significantly easier, in Section III-B, and we provide guidelines on common errors within OSM data that can impair the conversion process in Section III-C.

A. Limitations in the Legacy Conversion

As many other microscopic traffic simulators, SUMO models the road network topology as a graph composed of sets of edges (i.e., the road segments) and nodes (i.e., the intersections). Both

[1] Available at http://www.cs.unibo.it/projects/bolognaringway/

edges and nodes feature complex internal structures: Edges are structured into lanes, whereas nodes define all possible connections among the lanes they interconnect, as well as the parametrization of traffic lights that possibly regulate them. Manual configuration of such composite representations is unfeasible at scales larger than a few intersections, which calls for tools that can perform an automated configuration of the road network. In the case of SUMO, the reference tool is NETCONVERT, which builds on data from several cartographic sources, among which OSM.

Specifically, OSM data contain information that can be used by NETCONVERT to determine the internal structure of edges, i.e., the number, direction, and speed limit of lanes. However, data from OSM—or any other common map database—do not include the information required to build the internal structure of nodes, i.e., intersections, in the SUMO representation. There, NETCONVERT implements a dedicated algorithm to figure out the correct connections among the lanes entering and exiting a road intersection.

The algorithm works on a per-intersection basis; its baseline operation is described in [23], whereas refinements are presented in [24]. In brief, it first orders all incoming and outgoing edges according to their priority, inferred using some of the OSM lane attributes, such as the road type or the speed limit. Then, it determines edge-to-edge, lane-to-edge, and lane-to-lane associations, in this order, by serving the highest priority edges first. Finally, it rechecks the internal structure for flaws such as disconnected lanes. The same ordered list of edges is used also to assign cycles to traffic lights governing the intersection. We refer the interested reader to [23] and [24] for a comprehensive description of the algorithm.

The solution previously sketched aims at accommodating the conversion of many different types of intersection using a same heuristic, and, conceptually speaking, does so in a sensible way. However, after some hands-on experiments on different OSM datasets, we found that the NETCONVERT output is affected by inaccuracies that can drastically reduce the capacity of the road network. That is, we spotted the following common conversion errors in the scenarios we considered.

- The legacy NETCONVERT algorithm tends to disable many lane-to-lane connections in the presence of multilane edges. For example, in a four-way intersection joining roads that feature three lanes each, it may happen that only the middle lane of each road is allowed to go straight, whereas the outermost ones are used for left and right turns only. An example can be observed in Fig. 3(a), where only one of the two lanes in the road heading northwest is allowed to go straight at the top part of the intersection. While this can be acceptable in some cases, it is a configuration that generally reduces the straight traffic throughput at the intersection.
- Turning is only allowed on a lane-to-lane basis, i.e., incoming traffic from one left-turn lane is not allowed to flow into multiple lanes of the outgoing road on the left. Again, an example is shown in Fig. 3(a), where the rightmost lane on the east-heading road is connected to just one lane of the two-lane road heading northwest. Enabling lane-to-multilane turning would not generate conflicts in traffic flows and would significantly increase the turning traffic throughput at intersections.
- Left turns are allowed almost everywhere, even in multiple-lane boulevards. When not supported by a coherent traffic light synchronization, this tends to reduce the intersection outflow, up to the point where deadlocks (i.e., situations where vehicles are stuck inside the intersection) appear. In fact, in many real-world situations, left turns are forbidden on main arteries, where drivers are forced to use service roads or perform 270° turns to go left. Left-turn rules are, in general, difficult to figure out and depend on the underneath traffic demand, as recommended, e.g., by the U.S. Highway Capacity Manual [25].
- Traffic light synchronization is generally more complex than needed. The dedicated NETCONVERT algorithm builds on the composite connection structure aforementioned, which leads, at times, to long traffic light cycles that create significant queuing at intersections. An example is shown in Fig. 3(b), where the green traffic light phase allows concurrent crossing by conflicting flows.

The overlapping of all these problems makes the SUMO road network prone to congestion. In some cases, it induces severe deadlocks that the simulator tries to solve by teleporting stuck vehicles, often when the traffic flow stability is already compromised. We remark that similar problems were encountered under different configurations of the current release of NETCONVERT, within SUMO version 0.22, as also proven by the results of our comparative analysis in Section IV-B.

B. NETCONVERT Modifications

Our work focuses on improving the quality of the conversion of intersections and roundabouts, as well as redefines the traffic light synchronization logic, as detailed next.

1) Connections: Determining the rules according to which vehicles can traverse road junctions, such as intersections and roundabouts, plays a critical role in a smooth microscopic-level simulation of road traffic.

Fig. 1(a) outlines the notation that is at the base of the algorithm used by our original NETCONVERT, henceforth named *trigonometric*, to perform intersection conversion. For each incoming edge,[2] we calculate the angle α between the edge itself and each of the outgoing edges. Angle α is used to infer if an outgoing edge is going straight $(-45^\circ \leq \alpha \leq 45^\circ)$ or turning right $(-135^\circ < \alpha \leq -45^\circ)$ or left $(45^\circ \leq \alpha < 135^\circ)$ with respect to the selected incoming edge. We then apply the following rules to create the internal connections.

- The rightmost lane of an incoming edge is connected to all lanes of an outgoing edge going right.
- Each lane i of an incoming edge is connected to lane i of an outgoing straight edge. Let n_{in} and n_{out} be the number of lanes of the incoming and outgoing edges, respectively.

[2] An edge maps to an unidirectional road segment. A single bidirectional road segment is thus mapped into two distinct unidirectional edges.

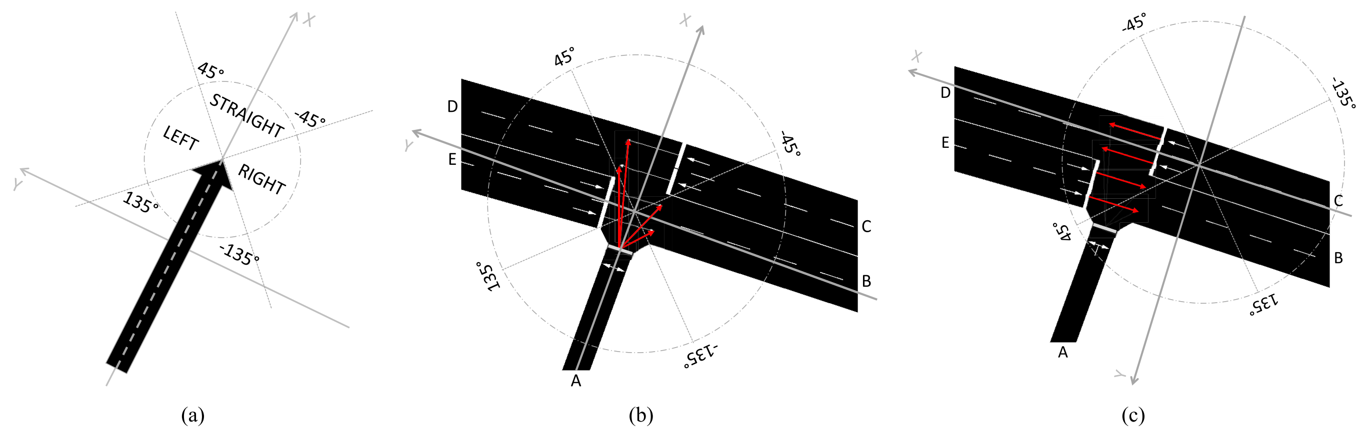


Fig. 1. Trigonometric NETCONVERT. (a) Straight, left, and right directions with respect to an incoming road. (b) Example of conversion for left and right connections at an intersection. (c) Example of conversion for straight connections at an intersection.

If $n_{\text{in}} > n_{\text{out}}$, then the exceeding lanes of the incoming edge are not connected to any lane of the outgoing edge. This rule forces lane changing before entering the intersection and avoids collisions during the simulation: the microscopic driver models in SUMO cannot handle situations where two vehicles enter together the same lane. Instead, when $n_{\text{in}} \leq n_{\text{out}}$, all lanes of the incoming edge are connected to those of the outgoing edge.

- The leftmost lane of an incoming edge is connected to all lanes of an outgoing edge going left.

Fig. 1(b) shows an example of left- and right-turn connections at an intersection with five edges. The bright red arrows between the lanes of the edges represent the enabled connections, i.e., the paths a vehicle can take to cross the intersection. When considering the incoming edge A, edge B has a relative angle $(-135^\circ < \alpha_{\text{AB}} \leq -45^\circ)$ with respect to A, and edge D has a relative angle $(45^\circ \leq \alpha_{\text{AD}} < 135^\circ)$ with respect to A. Consequently, the right turn is enabled from A's right lane to B and the left turn from A's left lane to D. Fig. 1(c) shows the configuration of connections between incoming and outgoing edges along the same direction. When considering the incoming edge C, edge D has a relative angle $(-45^\circ \leq \alpha_{\text{CD}} \leq 45^\circ)$; thus, the two edges are interconnected lane by lane, as previously discussed.

The trigonometric NETCONVERT also modifies the conversion logic in the presence of roundabouts. Specifically, our proposed algorithm leverages the OSM data to flag all edges identified as roundabout. The flag simplifies the rules when creating connections between adjacent roundabout edges, as well as between roundabout and outgoing edges. The following rules apply.

- Each lane i of an incoming edge, i.e., a nonroundabout edge, is connected to the lane i of the roundabout edge.
- Each lane i of a roundabout edge is connected to the lane i of the next roundabout edge.
- The n_{out} rightmost lanes of the roundabout edge are connected to the rightmost lanes of the outgoing edge, where n_{out} is the number of lanes of the outgoing ones.

The highest priority, as defined by SUMO, is assigned to all roundabout edges. This assignment is important to preserve the left precedence rule governing ingress in the roundabout. In our experience, when the legacy NETCONVERT is used, the precedence rule is not always satisfied at simulation run time.

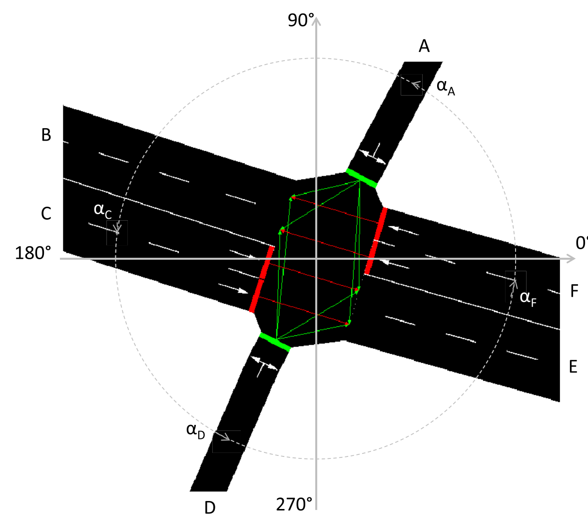


Fig. 2. Trigonometric NETCONVERT. Example of traffic light.

2) Traffic Lights: Automated traffic light configuration is a much desirable feature in the road conversion process. A reliable algorithm can dramatically reduce the manual configuration time of the simulation and, at the same time, ensure that the in- and outflow capacity of intersections is fully exploited.

The traffic light synchronization logic implemented in the trigonometric NETCONVERT builds on the identification of all incoming edges at an intersection, including their angle with respect to a common reference system, as depicted in Fig. 2. The rationale is that edges with an angle between 0° and 90° and between 180° and 270° should share the same traffic light phases. The edges with an angle between 90° and 180° and between 270° and 360° also share the same phases among them, but in opposition to the previous ones. In other words, this procedure aggregates the incoming edges into two clusters and configures coherent but opposite traffic light phases between clusters. With reference to the example in Fig. 2, edges A and D belong to a first cluster, whereas C and F belong to another.

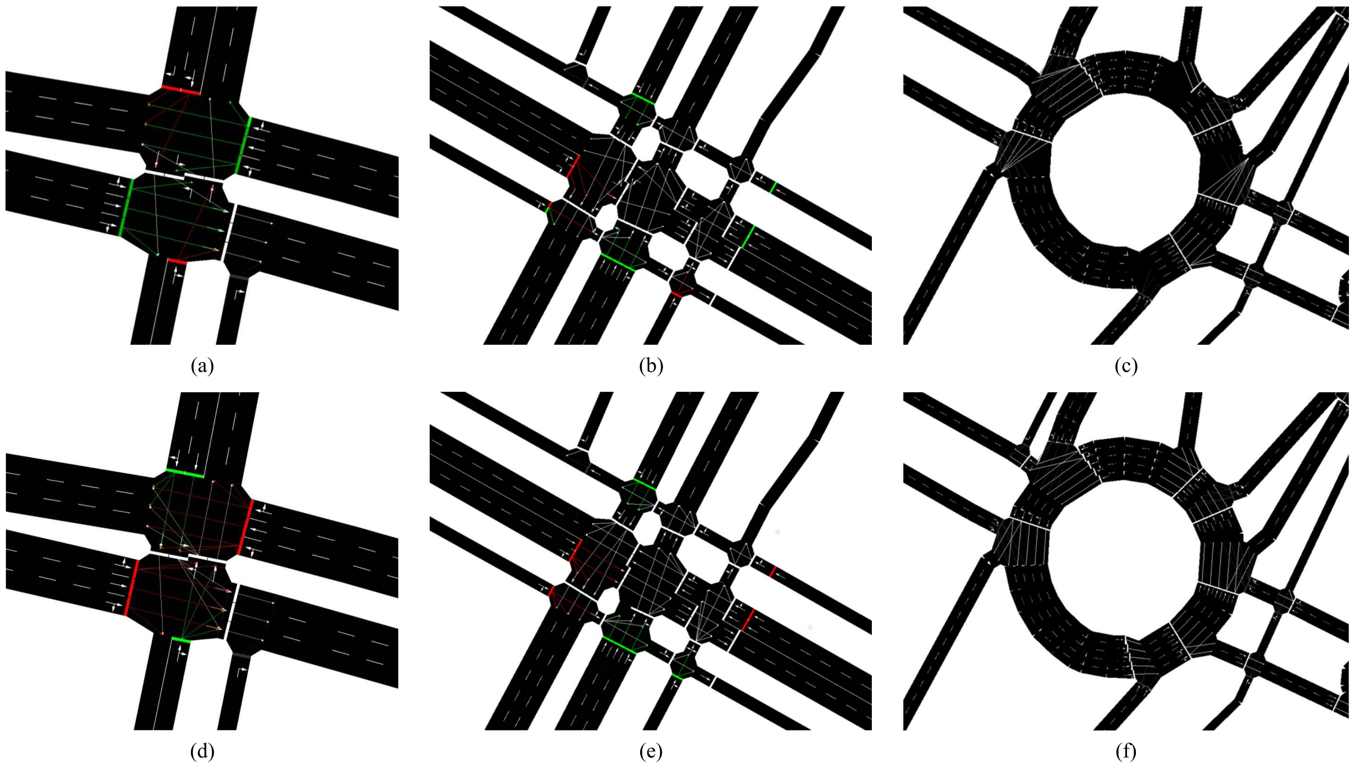


Fig. 3. Examples of road conversion from OSM to SUMO using the (top) legacy and (bottom) trigonometric NETCONVERT. (a) and (d) Dual-carriageway intersection. (b) and (e) Multiple-parallel-carriageway intersection. (c) and (f) Five-entry–five-exit roundabout.

We remark that some borderline scenarios can lead to synchronization errors. For example, an edge with an angle equal to 91° and another one with an angle equal to 179° are aggregated into the same cluster, even if they are almost perpendicular. These cases could be solved by increasing the number of clusters at the cost of adding traffic light phases. However, in our experience, two clusters are enough to handle the vast majority of cases, without adding traffic light phases that would reduce the fraction of time dedicated to each phase and, thus, slow down the vehicle flow.

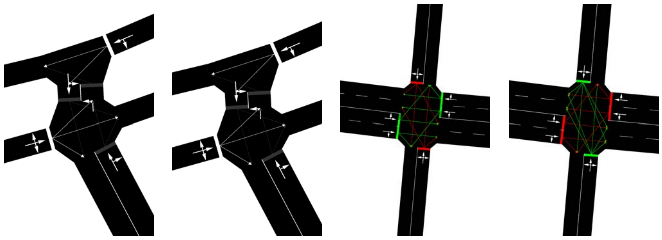
3) Practical Examples: Several representative examples of OSM-to-SUMO road network conversions are presented in Fig. 3. We compare the results obtained with the legacy NETCONVERT (top) and the trigonometric version (bottom).

Fig. 3(a) shows an example of an intersection merging a dual-carriageway road with a single-carriageway road, as converted by the legacy NETCONVERT. The conversion is flawed because vehicles incoming from the north direction cannot turn left on the avenue going southeast. This type of conversion error leads to increased route lengths, as vehicles must undergo long detours to enter the avenue. Fig. 3(d) shows the conversion of the same intersection by the trigonometric NETCONVERT. It is worth noting that all turns are correctly represented in this case. Moreover, vehicles can now employ both lanes of the outgoing edge toward north, whereas one lane was enabled by the legacy NETCONVERT. These improvements lead to significantly increased in- and outflow throughputs.

Fig. 3(b) shows an example of a more complex intersection converted by the legacy NETCONVERT. In this case, one dual-carriageway road merges with a multiple-parallel-carriageway road. Multiple evident errors affect the converted road network in Fig. 3(b). The road along the northwest-to-southeast heading has two lanes in each direction; however, in both directions, vehicles are not allowed to move straight through the junction; instead, they have to undergo multiple lane changes, and in the end, the capacity of the intersection to manage straight traffic is halved. Moreover, the legacy NETCONVERT leaves the possibility for drivers to turn on their left once they are in the middle of the intersection; at simulation run time, this leads to deadlocks among vehicles occupying the center of the intersection. Finally, the traffic light synchronization determined by the legacy NETCONVERT allows vehicles traveling along perpendicular incoming roads to cross the junction at the same time (note the green lights in the picture), which results in congestion and accidents. Fig. 3(e) depicts the result of the conversion of the same intersection, which was performed with the trigonometric NETCONVERT. Here, all lanes are coherently connected, which enables the utilization of the full capacity of the intersection and makes the traffic simulation more fluid and realistic. In addition, left turns from the center of the intersection are now prohibited,[3] which avoids deadlocks. Last but not the least, Fig. 3(e) shows that the traffic lights are now properly synchronized in all directions.

Fig. 3(c) shows an example of a roundabout with five inbound and five outbound roads, which was converted from OSM to SUMO with the legacy NETCONVERT. A major error affects, in this case, the conversion: There are no connections

[3]The trigonometric NETCONVERT allows the user to enable or disable left turns on the avenues through a specific command-line option.



(a) (b) (c) (d)

Fig. 4. Examples of test suite road conversions using the (a) and (c) legacy and (b) and (d) trigonometric NETCONVERT.

between consecutive edges along the roundabout, which basically means that no route exists from any entry point to any exit points of the roundabout. In other words, the roundabout is excluded from all routes computed by SUMO at run time, which may induce huge detours in vehicle trips and a critical loss of capacity in the road network. Fig. 3(f) shows the conversion results with the trigonometric NETCONVERT. In this case, connections within the roundabout are correctly modeled, and the roundabout becomes part of the simulated road network.

As a final remark, we mention that we ran the trigonometric NETCONVERT on the SUMO scenario test suite, and in most cases, we obtained identical results to those produced by the legacy conversion algorithm, as shown in Fig. 4(a) and (b). The only differences emerged in a few cases where the trigonometric NETCONVERT enabled lane-to-multilane turns, as shown in Fig. 4(c) and (d). As discussed in Section III-A, the additional connections created by our modified tool do not create conflicts but do increase the intersection capacity.

C. OpenStreetMap Modifications

Despite the improvements it brings over the legacy version, our proposed NETCONVERT cannot make up for issues that pertain to the original OSM data. The latter need to be fixed before the OSM-to-SUMO road network conversion.

To correctly model large-scale urban scenarios, the input OSM data should be as close as possible to the actual road infrastructure to let the simulated road traffic flow smoothly. This is particularly important when the macroscopic travel demand mimics that observed in the real world, and inaccuracies in the representation of the road network make it impossible to accommodate a realistic demand.

Among all possible inconsistencies between the modeled road infrastructure and the real one, there are two main aspects that need to be taken into account, based on our experience, as follows.

- *Correct number of lanes*. Several roads do not report, or wrongly report, the number of lanes of which the corresponding edges are composed. The OSM format assigns a `highway` tag to all roads, along with a value indicating the road category, e.g., motorway, trunk, and primary. Depending on the category to which it belongs, each road is assigned a default number of lanes. However, the default configuration is, most of the time, different from the real-world one; this leads to an unrealistic, and typically underestimated, road capacity that is not compatible with the actual travel demand. The `lanes` tag in OSM data overrides the default category-based number of lanes, and its correct setting is thus paramount to obtain a converted road layout in SUMO that is coherent with the real one.
- *One-way-only roads*. Roads in OSM data are assigned, at times, a wrong value of the `oneway` tag, which determines whether the road is one way or not. Such a misconfiguration prevents real-world routes from being used in the simulation, forcing vehicles to travel much longer paths to reach their destination. Depending on the `highway` value used for the road, the `oneway` tag might be not necessary. However, it has to be used every time the road presents differences with the standard behavior.

We thus strongly suggest that OSM data be carefully polished before the conversion is performed. Particular attention should be paid to fixing all instances of the aforementioned two issues.

IV. Use Case: The Bologna Ringway Dataset

We assess the quality of the novel OSM-to-SUMO road network conversion, by targeting one specific use case scenario, i.e., the Bologna Ringway dataset. Bologna is a middle-sized city of around 380 000 inhabitants, which is located in central-northern Italy. As shown in Fig. 5(a), the urban area is composed of a limited-access downtown, which is surrounded by a ringway that features three lanes per direction; the peripheral regions of the city lay outside the ringway. The typical morning traffic consists in commuting drives from residential areas in the city outskirts, toward downtown, where offices and commercial activities are located. Commuters tend to drive around the city using the fast-transit ringway and only enter the inner part of the city once close to their destination. Thus, most of the vehicular traffic flows along the ringway.

A. Simulation Configuration

Our use case scenario encompasses the downtown region, the ringway around it, and part of the periphery; and it is aptly named Bologna Ringway. The layout is shown in Fig. 5(a).

1) Road Network: The information on the road infrastructure is extracted from up-to-date OSM data, downloaded in February 2015. We had to perform a number of fixes in the OSM data, based on the recommendations in Section III-C. In particular, we corrected critical errors on the number of lanes and bidirectionality of road segments along the main ringway, which completely impaired the simulation of vehicular mobility. Overall, we fixed 163 roads that did not report the correct number of lanes, seven roads that were defined two times in the same area, and six roads that did not exist in the real network.

We employ the trigonometric NETCONVERT presented in Section III to convert OSM data in the Bologna Ringway scenario into a SUMO road network. For the sake of completeness, we test other conversion approaches as well, which allows appreciating the impact of the different conversions on

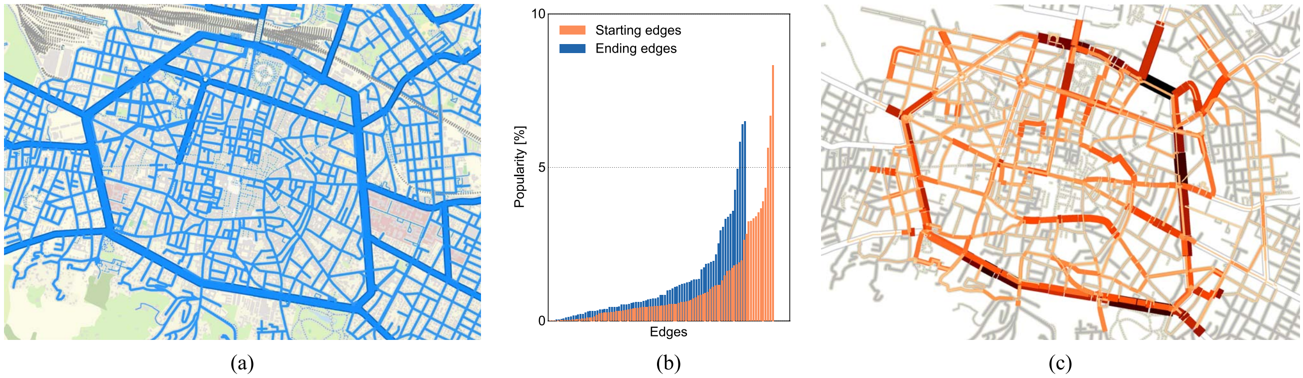


Fig. 5. Bologna Ringway. (a) Road network map. (b) Popularity of start and end edges. (c) Route distribution over the network.

the microscopic mobility simulation. We consider the following four configurations.

- *iTetris*. The iTetris dataset is the original road traffic trace developed in the collaborative research project of the same name [26]. The dataset is obtained by feeding the iTetris travel demand presented earlier, as well as a polished road network map derived from OSM data, to SUMO version 0.12, i.e., the current version during the project execution, in 2010. Interestingly, the dataset also features optimized hand-configured traffic light synchronization that mimics that in the real world.
- *Legacy*. The Legacy dataset is obtained by combining the iTetris travel demand and polished OSM data with the latest release of SUMO, i.e., version 0.22. Therefore, it employs the legacy `NETCONVERT` and the automated traffic light synchronization algorithm present in that version. We consider two flavors: **A** uses basic conversion options, whereas **B** adds the recommended `NETCONVERT` options `–roundabouts.guess` and `–junctions.join`.
- *Trigonometric*. The Trigonometric dataset is the same as the Legacy one, but for the trigonometric `NETCONVERT`. It thus uses SUMO version 0.22, but it features the enhanced road network conversion and traffic light parametrization presented in Section III.

2) Microscopic Mobility Models: The microscopic-level models adopted are the default ones implemented by SUMO, i.e., Krauss' car-following model [28] and Krajzewicz's lane-changing model [29]. These control drivers' acceleration and overtaking decisions, respectively, by taking into account a number of factors (e.g., the distance to the leading vehicle, the traveling speed, and the acceleration and deceleration profiles).

3) Macroscopic Traffic Flows: The travel demand we consider describes 1 h of traffic in the Bologna conurbation during the morning rush hour. The dataset was provided by the iTetris project [26] and represents the most challenging conditions for our performance evaluation: indeed, if the road network can support the 8 A.M. traffic peak, it will also take in the travel demand observed during any other time of the day. The travel demand is structured as an O–D matrix, which is converted into individual trips of vehicles: it thus describes the departure and arrival locations of each vehicle in the region during the simulated hour. A total of 22 213 individual trips is present in the dataset, starting at 93 different edges and ending at 81 edges around the city.

Fig. 5(b) shows the distribution of trips over the starting and ending edges: we observe a limited number of edges that gather most of the vehicle departures and arrivals, with roughly the same four starting edges and five ending edges present in more than 20% of the total trips. The vast majority of edges are featured in a small percentage of trips: 2% or less each. Fig. 5(c) presents instead the distribution of traveled routes over the street layout, as recoded in the synthetic trace. Thicker darker road segments are part of the routes of a larger number of vehicles. We observe the importance of the ringway for traffic flows in the area, as the relative segments support most of the travel demand. The main entry points to downtown can also be spotted, as they feature significant traffic.

To perform a more general analysis, we generate mobility datasets under four different road traffic loads[4]: 1) the complete travel demand, i.e., 22 213 trips, which maps to the traffic peak hour between 8 A.M. and 9 A.M.; 2) 90% of the travel demand, i.e., 19 940 trips, which can be mapped to traffic between 7 A.M. and 8 A.M.; 3) 80% of the travel demand, i.e., 17 806 trips, which can be mapped to traffic between 9 A.M. and 10 A.M.; and 4) 70% of the travel demand, i.e., 15 483 trips, which can be mapped to the hourly traffic from 10 A.M. through 12 A.M. Considering different traffic loads allows simulating conditions other than the complete saturation observed during the morning rush hour.

In all settings, the final mobility is obtained by iterating Gawron's traffic assignment algorithm [30], until a user equilibrium is reached. Gawron's algorithm determines the least expensive (i.e., fastest) route for each vehicle and then computes a cost for each road segment based on its occupancy level. At each iteration, it then moves part of the traffic to alternate less congested paths and recomputes road segment costs, until new iterations bring no advantage, i.e., the road capacity is exploited to a maximum.

B. Dataset Analysis

A first representative metric of the synthetic mobility is the time series of the number of vehicles circulating in the

[4]The mapping between the load percentages and the daytime hours is inferred from nonpublic statistics provided by the transportation agency in Bologna.

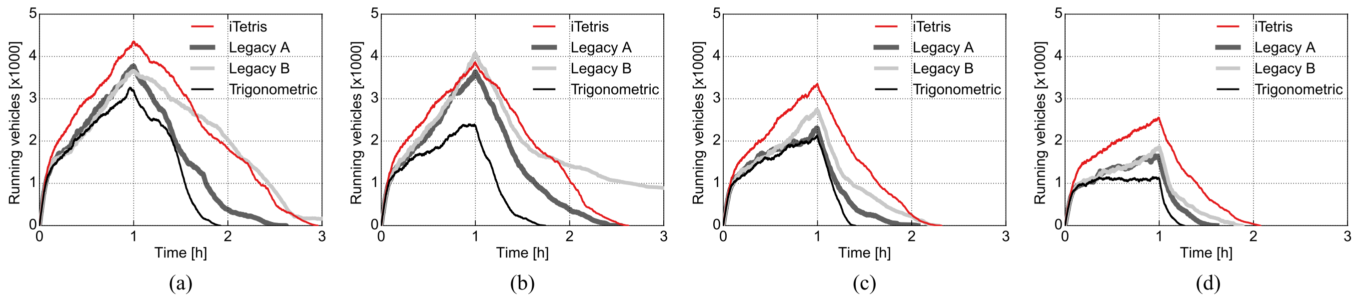


Fig. 6. Bologna Ringway. Concurrently traveling vehicles under different road traffic loads. (a) 100% travel demand. (b) 90% travel demand. (c) 80% travel demand. (d) 70% travel demand.

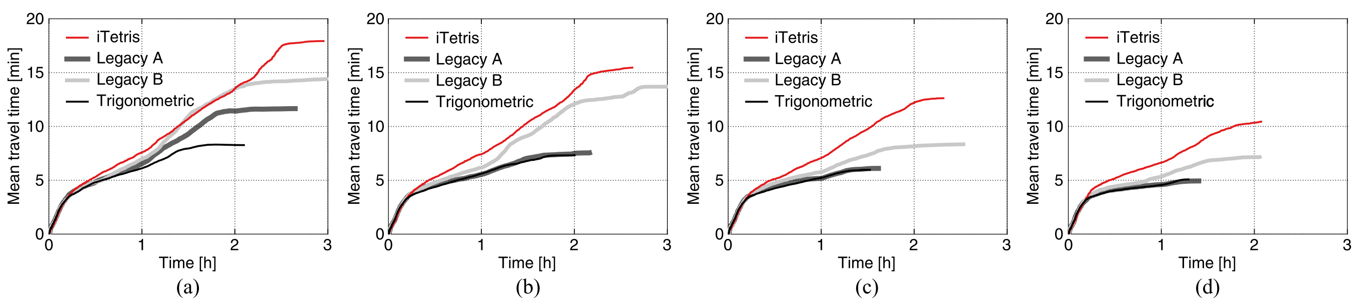


Fig. 7. Bologna Ringway. Mean travel time under different road traffic loads. (a) 100% travel demand. (b) 90% travel demand. (c) 80% travel demand. (d) 70% travel demand.

simulated region, as shown in Fig. 6. Each plot refers to a different road traffic load, which is expressed as a percentage of the iTetris travel demand. The curves in each plot refer instead to the simulation tool configurations, i.e., iTetris, Legacy A and B, and Trigonometric.

The number of vehicles always starts at zero, since the road network is empty at simulation startup. The simulation reaches a steady state, where vehicles are spread over the whole road network, after 15 min from the start of the first trip. In addition, a peak is reached after 1 h; then, vehicles stop being injected into the simulation, which runs until all cars already present in the road network reach their destination.

By comparing plots in Fig. 6(a)–(d), we note that the number of running vehicles decreases as the travel demand is reduced: heavier traffic loads induce more severe congestion, which, in turn, increases travel times and forces a larger number of vehicles to coexist in the road network.

When comparing instead the results obtained with different tool configurations, we observe that, in all travel demand settings, the original iTetris configuration yields a larger number of running vehicles than the Legacy and Trigonometric configurations. The enhancements introduced by SUMO version 0.22 over the version 0.12 originally used by iTetris allow better accommodating the road traffic, reducing congestion, and thus the number of concurrently traveling vehicles. A direct comparison between the Legacy and Trigonometric configurations highlights how the latter further reduces the number of vehicles present in the road network at the same time. The improvement is typically the most evident at the traffic peak occurring after around 1 h of elapsed time, and it ranges between 10% and 40%, depending on the travel demand settings. It is also interesting that, under the Trigonometric configuration, simulations tend to complete much faster, again with gains ranging between 15% and 40%. The very long tails in the Legacy B case are due to a few intersections, unable to manage the traffic load and generating important queues that take a long time to be served.

In terms of road network conversion, these results demonstrate that the trigonometric `NETCONVERT` can bring significant advantages over the legacy approaches in the current SUMO release. Indeed, the SUMO representation of the road infrastructure provided by the trigonometric `NETCONVERT` makes traffic significantly smoother, which reduces congestion and allows discharging the road network much more fluidly once no new vehicles are injected.

Confirmations come from other road traffic performance metrics. Fig. 7 details the mean travel time of vehicles in the Bologna Ringway scenario, as recorded over simulation time. We observe that, in all cases, travel times tend to grow over time: this is coherent with the results in Fig. 6, showing that the road network becomes increasingly congested as simulation time elapses, and thus, vehicles travel at lower speed and take more time to reach their destination. This effect becomes less pronounced as the travel demand is reduced, since traffic then becomes more fluid. The very low travel time recorded when close to the x-axis origin is a simulation artifact, due to the fact that the road network is initially empty, and thus, only vehicles traveling over short distances are included in the computation. Again, we can remark that the simulation reaches a steady state after around 15 min, which corresponds to a change in the slope of all curves.

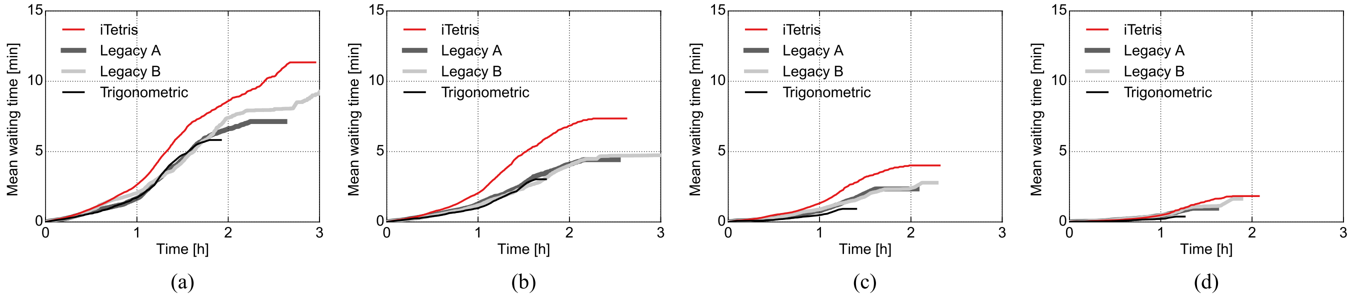


Fig. 8. Bologna Ringway. Mean waiting time under different road traffic loads. (a) 100% travel demand. (b) 90% travel demand. (c) 80% travel demand. (d) 70% travel demand.

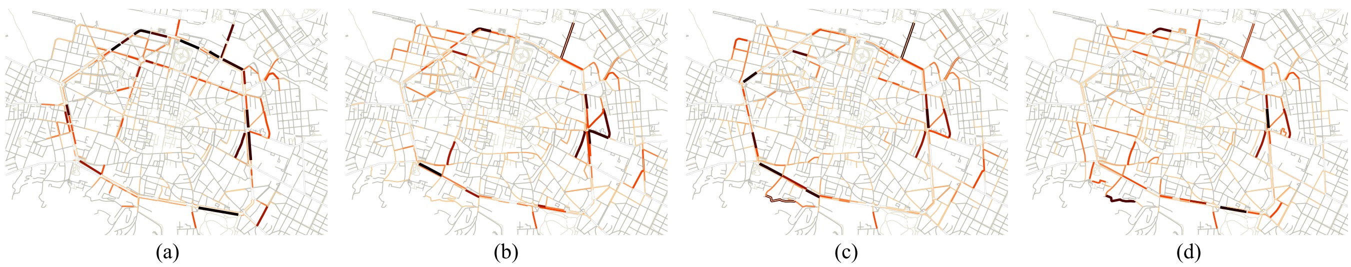


Fig. 9. Bologna Ringway. Road traffic density over the network at the simulation peak traffic time. (a) iTetris. (b) Legacy A. (c) Legacy B. (d) Trigonometric.

The comparison between the iTetris, Legacy, and Trigonometric configurations underscores the advantage yielded by the latter, with average travel times that are reduced by up to 2 min. The gain is significant, considering the size of the scenario, as well as the fact that trips in the area typically last between 5 and 10 min.

Fig. 8 presents the mean time spent by vehicles in standstill situations, which may be generated by queuing at intersections and ramps or by generic congestion within the urban area. Different plots refer to varying road traffic loads. Clearly, it is desirable that such a waiting time is the lowest possible, since this implies a smoother road traffic; at the same time, the waiting time cannot be zero in a realistic scenario, since some queuing is unavoidable due to, e.g., traffic lights.

A cross comparison of the four plots indicates that the waiting time is reduced along with the travel demand, which is intuitive, since we already observed that a lower demand implies less congestion in the road network. In addition, both the Legacy and Trigonometric configurations outperform the original iTetris settings, attaining a significantly lower mean waiting time. Under heavy load, the iTetris configuration forces vehicles to spend, on average, 40% more time in standstill situations than what happens with the Legacy and Trigonometric configurations. When comparing the Legacy and Trigonometric cases, performances are identical in the presence of heavy traffic conditions, as shown in Fig. 8(a). However, as the traffic load is reduced, some diversity emerges, with the Trigonometric setting accommodating the demand more smoothly than the Legacy configuration. The gain can reach peaks of 60%, in terms of waiting time reduction, when 70% of the original travel demand is considered.

Overall, we observe that, throughout all the results presented here, the trigonometric `NETCONVERT` presented in Section III produces a SUMO road network that is more efficient than that generated by the legacy `NETCONVERT`. The more consistent road infrastructure representation has an important impact on the simulated microscopic-level mobility, resulting in more fluid traffic and reduced queuing and congestion.

In particular, we remark that the Trigonometric configuration is the only configuration capable to attain stable traffic conditions when considering 70% of the iTetris travel demand, as shown in Figs. 6(d)–8(d): there, the number of running vehicles is steady throughout the whole simulation, with nearly constant mean travel time and very low (10–20 s) average waiting times. This result is particularly meaningful, since, as discussed in Section IV-A3, a 70% traffic load corresponds to the moderate traffic observed during the morning hours in the Bologna area we considered. Unlike what happens during the rush hours represented by the 80%–100% demands, where some congestion is unavoidable, the road network shall be able to accommodate the ordinary load observed during most of the day. In fact, the Trigonometric configuration is the only configuration capable of reproducing that realistic behavior.

Finally, Fig. 9 presents the road traffic density recorded in the target area, 1 h after the simulation start. Each plot refers to a different tool configuration, which allows appreciating how differences are not only related to the aggregate traffic statistics in the previous figures but to the geographical distribution of congestion as well. For instance, the iTetris configuration leads to a large traffic jam along the whole northeastern section of the ringway around downtown Bologna. The Legacy A and B configurations generate instead congestion in the eastern and southwestern regions of the city, respectively. Coherently with the previous results, the Trigonometric case is that yielding the mildest problems in terms of road traffic.

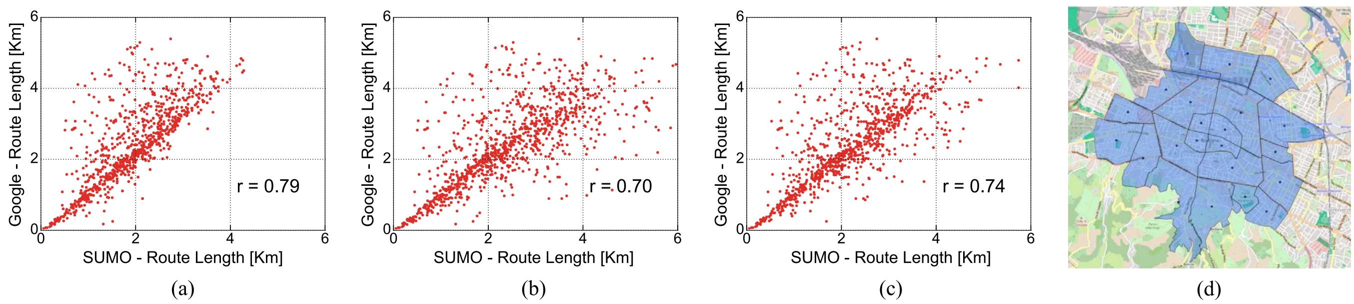


Fig. 10. Navigation-service-based validation. (a)–(c) Routing path lengths provided by Google Maps versus those recorded in simulation. (a) One simulated probe vehicle, (b) complete iTetris travel demand, corresponding to the morning road traffic peak, and (c) 70% of the iTetris travel demand, corresponding to typical moderate morning traffic. (d) Twenty-two statistical areas of Bologna as an OSM overlay; dots map to area centroids.

V. Dataset Validation via Navigation Services

A major issue with synthetic mobility traces, such as the Bologna Ringway dataset presented earlier, is that they are the result of a simulation and thus need to undergo some validation, as discussed in Section II. This aspect is often overlooked, and only a few previous studies proposed validation techniques that are, however, based on visual inspection [18] or city-specific hard-to-retrieve data [31].

A. *Navigation-Service-Based Validation*

We propose an original approach to the validation of synthetic road traffic traces. Our methodology builds on publicly accessible data that are available for the vast majority of urban areas worldwide and can thus be replicated for the validation of any mobility dataset.

Specifically, we leverage navigation services that also offer routing and travel time estimation functions, such as Google Maps. The results returned by such services are based on a number of real-world sources, including statistics produced by transportation authorities, floating car data generated by in-vehicle navigation systems, and crowdsourced data collected by GPS-enabled smartphones that run dedicated applications. The resulting information is extremely accurate and can be retrieved via dedicated APIs. In the following, we will consider Google Maps as our reference navigation service, due to its popularity.

Routing and travel time information from navigation services can be used as a ground truth for the equivalent data extracted from a synthetic mobility dataset. More precisely, navigation service data can be leveraged for the validation of two key performance metrics of road traffic, as follows.

- *Routing paths*. The first step of our proposed navigation-service-based validation focuses on the analysis of the routes traveled by vehicles in simulation. By using dedicated Google Maps APIs, one can input the origin and destination pair of each trip in the synthetic dataset to the Google Maps routing function and retrieve the corresponding path indicated by the navigation service. The latter can be compared with that followed by vehicles in simulation, to verify if traveled routes are realistic.
- *Travel times*. The second phase of the validation process concerns travel times. The procedure is similar to that adopted for the routing paths but involves querying Google Maps APIs for trip duration estimates. This information can be then used as a benchmark for the equivalent data recorded in the synthetic dataset.

B. *Validation of the Bologna Ringway Dataset*

The aforementioned two-phase validation process was run on the Bologna Ringway dataset, under the Trigonometric configuration. Once more, we remark that the methodology is based on open-access data, and it is thus not specific to this scenario, but it is replicable in most other urban areas.

The results of the routing path validation are presented in Fig. 10(a). In all scatterplots, each point represents one trip: its x-axis value is the corresponding route length measured in simulation, whereas its y-axis value is the same route length retrieved from the Google Maps service. In the ideal case, the routes assigned in simulation are the same ones suggested by the navigation service, and all points shall lie on the bisector line. This would map to a Pearson product–moment correlation coefficient r equal to 1.

Fig. 10(a) refers to a simplified scenario, where only one vehicle is traveling at a time, between a selected pair of origin and destination locations. Thus, each point in the plot corresponds to one independent simulation. The rationale for this setting is that having a lone vehicle travel in the road network eliminates all bias due to queuing and traffic jams; as a matter of fact, the latter phenomena can induce vehicles to modify their routes, according to Gawron's algorithm presented in Section IV-A3, to avoid congested areas—an aspect that Google Maps APIs do not take into account.[5] We observe a good match between navigation service data and simulation results: The vast majority of points lie along the bisector, and the correlation is rather strong at $r = 0.79$. Minor discrepancies are mainly related to the fact that, in some cases, Google Maps suggests routes longer than those computed in simulation. Such an effect is induced by a slightly stronger tendency of Google Maps routing to prefer paths over high-speed roads, when compared with SUMO routing. This comes at some cost in terms of traveled distance but only occurs on a small set of cases, which explains the limited dispersion in the scatterplot.

Fig. 10(b) and (c) assesses the reliability of the simulated routing paths once they are determined by Gawron's traffic

[5]Routing and travel time estimation in presence of traffic information are only available to business customers, through Google Maps APIs.

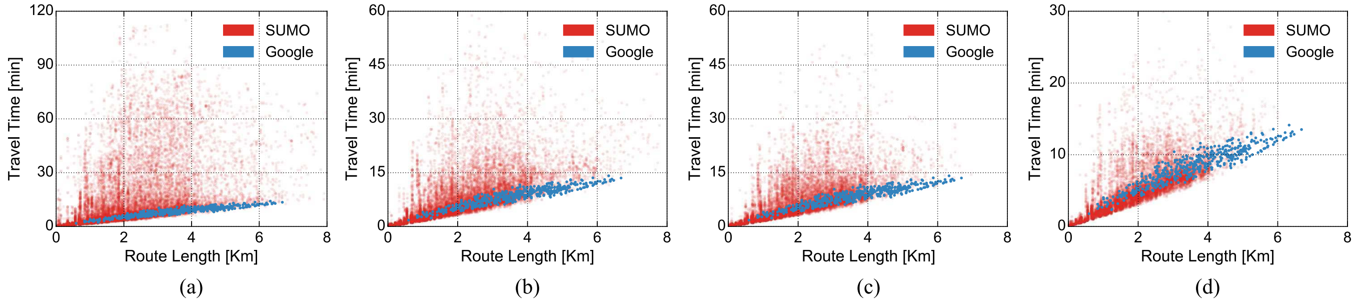


Fig. 11. Navigation-service-based validation. Travel time versus road lengths, as provided by Google Maps and recorded in simulation. Different plots refer to diverse road traffic loads. (a) 100% travel demand. (b) 90% travel demand. (c) 80% travel demand. (d) 70% travel demand.

assignment algorithm. To that end, we consider the complete travel demand provided by the iTetris project, as shown in Fig. 10(b), as well as the case where only 70% of the demand is injected in the road network, as shown in Fig. 10(c). In both scenarios, the presence of heavy and moderate background road traffic triggers path balancing through Gawron's algorithm iterations. We note that an increasing traffic load on the road network reduces the correlation, as it forces simulated vehicles to take longer routes to avoid congested road segments and intersections. Still, we remark that Gawron's algorithm does a good job of accommodating the demand without massively rerouting vehicles: Indeed, the correlation remains strong enough, at $r = 0.7$ and $r = 0.74$, respectively, which implies that comparable paths are traveled in simulation and suggested by the navigation service based on real-world data.

For the second validation phase, concerning travel times, we take a slightly different approach, which allows further demonstrating the generality of our methodology. We consider the 22 statistical areas of the city of Bologna, which are depicted in Fig. 10(d); they are the basis for common statistical analyses by public administrations and are typically available for most urban areas. We employ the centroids of the statistical areas as origin and destination points for queries on travel times to the Google Maps service; specifically, one query is issued for each combination of O–D centroids, generating a total of 462 statistical trips. We then compare the relationship between the travel time and the route length, as obtained by statistical trips from Google Maps and by trips simulated in SUMO using the iTetris travel demand.

Fig. 11 shows the travel time as a function of the route length, as obtained from Google Maps and in simulation. Plots refer to the usual different traffic demands described in Section IV-A3. We remark that the simulated travel times overlap for the most part with the values provided by the navigation service, under all road traffic loads. More precisely, the overlap increases as the demand is reduced. Indeed, the dispersion of simulation points toward the top of the plot is due to congestion, which is neglected in Google Maps data: thus, reducing the injected traffic leads to trips whose duration is closer to that anticipated by the navigation service. It is also worth noting that the minimum travel times recorded in simulation are coherent with the minimum values provided by the Google Maps service for any route length. Therefore, the speed profile of the vehicles in simulation well approximates the real-world ones.

In conclusion, all the results presented here confirm the quality of the Bologna Ringway dataset. In particular, both the road network conversion through the trigonometric NETCONVERT presented in Section III and Gawron's traffic assignment algorithm implemented by SUMO mimic real-world routing behaviors, which are inferred from navigation services.

VI. Related Work

Our work primarily concerns the microscopic-level simulation of road traffic for vehicular networking purposes. Several tools have been proposed in the past that are dedicated to this task. Many of them are commercial software that ensure a very high level of detail at the cost of configuration complexity: examples include PARAMICS [32], CORSIM [33], or VISSIM [34]. The economic cost and the steep learning curve of these tools have, however, favored the diffusion of simpler open-source simulators of vehicular mobility, which better fit the needs of a vehicular network simulation.

Popular tools adopted by the networking research community were originally based on simplistic stochastic models, such as those implemented by the IMPORTANT framework [14]. The latter were then replaced by a random travel demand over realistic road topologies [35]. Later on, microscopic models of driver's behavior, which were developed within transportation research, were included in such tools [36]. Since then, open-source road network simulators have become increasingly complex and currently allow accurate simulation of the movement of individual vehicles over real-world road networks: widely adopted software include SUMO [22], VanetMobiSim [37], or STRAW/SWANS [38]. Among them, SUMO is recognized as the current state of the art and is part of a number of federation efforts with network simulators, such as those carried out by iTetris [26], Veins [39], [40], or VSimRTI [41]. However, improvements to parts of SUMO are still possible, as in the case of the trigonometric NETCONVERT we present in Section III.

The availability of dependable simulation tools is critical to the generation of reliable synthetic datasets of vehicular mobility. However, it is not the only component required to that end. A major challenge lies today in the realistic modeling of the movement of large traffic flows across the road network. Without that piece of information, one has to resort to random trips [14], [35] or intuition-based traffic assignment [42], which are not representative of the actual movement of vehicles within urban regions. Clearly, a biased distribution of road traffic dramatically affects the topology of a vehicular network built on top of vehicle mobility and, thus, puts the significance of results at stake [43].

A very limited number of datasets are available today, which rely on O–D matrices that describe real-world traffic flows. Those describe vehicular mobility in Canton of Zurich [44], in Luxembourg [17], in the city of Cologne [18], and in suburban areas of Bologna [45]. Thus, the Bologna Ringway dataset we present in Section IV is a valuable contribution to the current set of dependable mobility traces available to the research community.

Finally, the validation of synthetic road traffic datasets is a last topic that is relevant to the work presented in this paper. This is a quite overlooked aspect in vehicular mobility simulation, and the only methodologies adopted to date within the networking community are limited to visual inspection against maps provided by live traffic services [18] and comparison against traffic count data that are, however, specific to a city and generally hard to retrieve [31]. The approach we present in Section V advances those attempts, since it is based on data that are publicly available for most urban areas worldwide and allows assessing the quality of the routing and travel times in the synthetic data.

VII. Conclusion

We have presented several contributions to the ongoing effort of enabling dependable vehicular network simulation through adoption of more realistic road traffic representations. Specifically, we have introduced an original Bologna Ringway dataset, describing the movement, during the morning rush hour, of more than 22 000 vehicles in a 25-km area that covers the center and outskirts of Bologna, Italy. It is our intention to make the dataset publicly available to the research community.

To generate the Bologna Ringway dataset, we employ an original version of the OSM-to-SUMO road network conversion tool, i.e., NETCONVERT, and show how it allows for a more reliable representation of road intersection and a better synchronization of traffic lights. These improvements lead to significant gains in terms of road network capacity, with reduced travel and waiting times in simulation.

Finally, we validated the routing paths and the travel times in the Bologna Ringway dataset, by means of a novel methodology, based on data retrieved from navigation services. Our approach demonstrates the quality of the proposed mobility dataset and is general enough to be reused for the validation of other road traffic traces.

Acknowledgment

The authors would like to thank G. Toso for his help in implementing and maintaining the trigonometric NETCONVERT. They would also like to acknowledge the Municipality of Bologna, Italy, for kindly providing us with the data used in this work.

References

[1] J. B. Kenney, "Standards and regulations," in *VANET—Vehicular Applications and Inter-Networking Technologies*, H. Hartenstein and K. P. Laberteaux, Eds. Hoboken, NJ, USA: Wiley, 2010.

[2] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. 802.11-2012, Mar. 2012.

[3] IEEE, "IEEE 1609.0-2013—IEEE Guide for Wireless Access in Vehicular Environments (WAVE)—Architecture," IEEE Std. 1609.0™-2013, Dec. 2013.

[4] *Communications Access for Land Mobiles (ITS-CALM-M5)*, OSI 21215 Std., Nov. 2010.

[5] *Intelligent Transportation Systems (ITS)*, ETSI Std. EN 302 665, Sep. 2010.

[6] Reuters, "Obama backs highway fund fix, touts 'talking' cars," *The New York Times*, New York, NY, USA, Jul. 2014.

[7] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, "LTE for vehicular networking: A survey," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 148–157, May 2013.

[8] J. Whitbeck, Y. Lopez, J. Leguay, V. Conan, and M. Dias de Amorim, "Relieving the wireless infrastructure: When opportunistic networks meet guaranteed delays," in *Proc. IEEE WoWMoM*, Lucca, Italy, Jun. 2011, pp. 1–10.

[9] R. Stanica, M. Fiore, and F. Malandrino, "Offloading floating car data," in *Proc. IEEE WoWMoM*, Madrid, Spain, Jun. 2013, pp. 1–9.

[10] C. Lochert, B. Scheuermann, and M. Mauve, "Information dissemination in VANETs," in *VANET—Vehicular Applications and Inter-Networking Technologies*, H. Hartenstein and K. Laberteaux, Eds. Hoboken, NJ, USA: Wiley, 2010.

[11] K. C. Lee, U. Lee, and M. Gerla, "Survey of routing protocols in vehicular ad hoc networks," in *Advances in Vehicular Ad-Hoc Networks: Developments and Challenges*, M. Watfa Ed. Hershey, PA, USA: Inf. Sci. Ref., IGI Global, 2010.

[12] Ann Arbor Safety Pilot. [Online]. Available: http://safetypilot.umtri.umich.edu/

[13] sim^{TD}. [Online]. Available: http://www.simtd.de

[14] F. Bai, N. Sadagopan, and A. Helmy, "The IMPORTANT framework for analyzing the impact of mobility on performance of routing protocols for ad hoc networks," *Elsevier Ad Hoc Networks*, vol. 1, pp. 383–403, 2003.

[15] M. Fiore and J. Härri, "The networking shape of vehicular mobility," in *Proc. ACM MobiHoc*, May 2008, pp. 261–272.

[16] W. Viriyasitavat, F. Bai, and O. K. Tonguz, "Dynamics of network connectivity in urban vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 3, pp. 515–533, Mar. 2011.

[17] Y. Pigné, G. Danoy, and P. Bouvry, "A vehicular mobility model based on real traffic counting data," in *Proc. Nets4Cars/Nets4Trains*, 2011, pp. 131–142.

[18] S. Uppoor, O. Trullols-Cruces, M. Fiore, and J. Barcelo-Ordinas, "Generation and analysis of a large-scale urban vehicular mobility dataset," *IEEE Trans. Mobile Comput.*, vol. 13, no. 5, pp. 1061–1075, May 2014.

[19] M. Gramaglia, O. Trullols-Cruces, D. Naboulsi, M. Fiore, and M. Calderon, "Vehicular networks on two Madrid highways," in *Proc. IEEE SECON*, Singapore, Jul. 2014, pp. 423–431.

[20] S. Joerer, C. Sommer, and F. Dressler, "Toward reproducibility and comparability of IVC simulation studies: A literature survey," *IEEE Commun. Mag.*, vol. 50, no. 10, pp. 82–88, Oct. 2012.

[21] OpenStreetMap. [Online]. Available: http://www.openstreetmap.org

[22] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO—Simulation of urban mobility," *Int. J. Adv. Syst. Meas.*, vol. 5, no. 3, pp. 128–138, 2012.

[23] D. Krajzewicz, G. Hertkorn, J. Ringel, and P. Wagner, "Preparation of digital maps for traffic simulation; Part 1: Approach and algorithms," in *Proc. Ind. Simul. Conf.*, 2005, pp. 285–290.

[24] D. Krajzewicz and J. Erdmann, "Road intersection model in SUMO," in *Proc. SUMO User Conf.*, 2013, pp. 3–17.

[25] Highway Capacity Manual. [Online]. Available: http://hcm.trb.org/

[26] M. Rondinone *et al.*, "ITETRIS: A modular simulation platform for the large scale evaluation of cooperative ITS applications," *Simul. Modell. Pract. Theory*, vol. 34, pp. 99–125, May 2013.

[27] Google Maps JavaScript API v3, "Directions Service." [Online]. Available: https://developers.google.com/maps/documentation/javascript/directions

[28] S. Krauß, P. Wagner, and C. Gawron, "Metastable states in a microscopic model of traffic flow," *Phys. Rev. E*, vol. 55, no. 304, pp. 55–97, May 1997.

[29] D. Krajzewicz, "Kombination von taktischen und strategischen Einflssen in einer mikroskopischen Verkehrsflusssimulation," in *Fahrermodellierung in Wissenschaft und Wirtschaft*, T. Jürgensohn and H. Kolrep, Eds. Berlin-Adlershof, Germany: VDI-Verlag, 2009, 104–115.

[30] C. Gawron, "An iterative algorithm to determine the dynamic user equilibrium in a traffic simulation model," *Int. J. Modern Phys. C*, vol. 9, no. 3, pp. 393–407, 1998.

[31] S. Uppoor and M. Fiore, "Characterizing pervasive vehicular access to the cellular RAN infrastructure: An urban case study," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2603–2614, Jun. 2015.

[32] Paramics: Microscopic Traffic Simulation. [Online]. Available: http://www.paramics-online.com

[33] CORSIM: Microscopic Traffic Simulation Model. [Online]. Available: http://mctrans.ce.ufl.edu/featured/tsis

[34] Ptv Simulation—VISSIM. [Online]. Available: http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/

[35] A. K. Saha and D. B. Johnson, "Modeling mobility for vehicular ad hoc networks," in *Proc. ACM VANET*, Philadelphia, PA, USA, Oct. 2004, pp. 91–92.

[36] S. Jaap, M. Bechler, and L. Wolf, "Evaluation of routing protocols for vehicular ad hoc networks in city traffic scenarios," in *Proc. IEEE ITSC*, Vienna, Austria, Sep. 2005.

[37] J. Härri, M. Fiore, F. Filali, and C. Bonnet, "Vehicular mobility simulation with VanetMobiSim," *Trans. Soc. Model. Simul.*, vol. 87, no. 4, pp. 275–300, Apr. 2011.

[38] D. Choffnes and F. Bustamante, "An integrated mobility and traffic model for vehicular wireless networks," in *Proc. ACM VANET*, 2005, pp. 68–79.

[39] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved IVC analysis," *IEEE Trans. Mobile Comput.*, vol. 10, no. 1, pp. 1536–1233, Jan. 2011.

[40] L. Bedogni *et al.*, "An interoperable architecture for mobile smart services over the Internet of energy," in *Proc. IEEE WoWMoM*, Madrid, Spain, Jun. 2013, pp. 1–6.

[41] B. Schuenemann, "V2X simulation runtime infrastructure VSimRTI: An assessment tool to design smart traffic management systems," *Comput. Netw.*, vol. 55, no. 14, pp. 3189–3198, Oct. 2011.

[42] C. Barberis and G. Malnati, "Epidemic information diffusion in realistic vehicular network mobility scenarios," in *Proc. IEEE ICUMT*, 2009, pp. 1–8.

[43] D. Naboulsi and M. Fiore, "On the instantaneous topology of a large-scale urban vehicular network: The Cologne case," in *Proc. ACM MobiHoc*, 2013, pp. 167–176.

[44] B. Raney *et al.*, "An agent-based microsimulation model of Swiss travel: First results," *Netw. Spatial Econ.*, vol. 3, no. 1, pp. 23–41, Jan. 2003.

[45] L. Bieker, D. Krajzewicz, A. Morra, C. Michelacci, and F. Cartolano, "Traffic simulation for all: A real world traffic scenario from the city of Bologna," in *Proc. SUMO*, 2014, pp. 47–60.

Luca Bedogni received the degree in computer science and the Ph.D. degree from the University of Bologna, Bologna, Italy, in 2008 and 2015.

He is currently a Research Assistant with the Department of Computer Science and Engineering, University of Bologna. His research interests include the study, modeling, and performance evaluation of future wireless networks operated in TV bands, cognitive wireless networks, and vehicular networking.

Marco Gramaglia received the M.Sc. degree in computer science from Politecnico di Torino, Torino, Italy, in 2008 and the M.Sc. and Ph.D. degrees in telematics engineering from Universidad Carlos III de Madrid, Madrid, Spain, in 2009 and 2012, respectively.

He is currently a Postdoctoral Research Fellow with the Institute of Electronics, Computer, and Telecommunications Engineering (IEIIT), National Research Council (CNR), Torino. Previously, he was a Researcher with Istituto Superiore Mario Boella, Torino. Between 2008 and 2012, he was a Research Assistant with IMDEA Networks Institute. His research interests are mobility aspects of wireless networking and vehicular ad hoc networks.

Andrea Vesco (S'06–M'09) received the M.Sc. degree in telecommunication engineering and the Ph.D. degree in computer and system engineering from Politecnico di Torino, Torino, Italy, in 2003 and 2009, respectively.

He has been a member of the Computer Networks Group with the Department of Control and Computer Engineering, Politecnico di Torino. In 2010, he joined the Networking Laboratory of the Istituto Superiore Mario Boella, Torino. His research interests were on quality of service (QoS) over packet switched networks and wireless access networks. Moreover, he carried out research activities on QoS over network on chip, in collaboration with the System-Level Design Group, Columbia University, New York, NY, USA. After ten years of experience and several publications in the networking field, he had the opportunity to focus his research studies on the subject of smart cities in 2013. His research effort is currently spent on the role of information and communications technology for the sustainable development of smart cities.

Marco Fiore (S'05–M'09) received the Ph.D. degree from Politecnico di Torino, Torino, Italy, in 2008 and the Habilitation á Diriger des Recherches degree from University of Lyon, Lyon, France, in 2014.

He is currently a Researcher with the Institute of Electronics, Computer, and Telecommunications Engineering (IEIIT), National Research Council (CNR), Torino, and an Associate Researcher with Inria, Le Chesnay, France. He has been an Associate Professor with Institut National des Sciences Appliquées de Lyon, Villeurbanne, France, and a Visiting Researcher with both Rice University, Houston, TX, USA, and Universitat Politecnica de Catalunya, Barcelona, Spain. His interests are in mobile traffic analysis and vehicular networking.

Jérôme Härri received the M.Sc. and the D.Sc. degrees in telecommunication from the Swiss Institute of Technology (EPFL), Lausanne, Switzerland.

He is currently a tenured Assistant Professor with the Mobile Communication Group, EURECOM, Sophia Antipolis, France, where he conducts research on wireless vehicular networks. Previously, he led the Traffic Telematics Junior Research Group of the Institute of Telematics, Karlsruhe Institute of Technology, Karlsruhe, Germany. He has authored and coauthored over 50 international journal and conference papers and is involved in various national and European research projects related to connected vehicles. His research interests are related to the optimization of the vehicular wireless channel usage, to the investigation of cooperative intelligent transportation systems strategies, and to the characterization of the mutual relationship between vehicular mobility and vehicular communication.

Francesco Ferrero received the M.Sc. degree in physics from Università di Torino, Torino, Italy.

He is currently the Head of the Smart City Strategic Program with Istituto Superiore Mario Boella, Torino, where a leads a multidisciplinary team of researchers conducting applied research in the field of economically, environmentally, and socially sustainable cities. He has recently edited a handbook of research on the argument (http://bit.ly/1DGZAPt) and is involved in several national and European research projects related to smart and sustainable cities and mobility.

Planning Local Energy Communities to Develop Low Carbon Urban and Suburban Areas

Giovanni Colombo, Francesco Ferrero, Giancarlo Pirani, Andrea Vesco

Istituto Superiore Mario Boella – Torino, Italy
E-mail: {colombo, fferrero, vesco}@ismb.it

Abstract—**This paper presents a model of a Local Energy Community (LEC) that is a set of users that share renewable sources and storage resources to satisfy energy demand while minimising carbon emissions and operating costs. The main contribution is the design of a planner that systematically assesses a target territory to cluster buildings into LECs while maximising an objective function. The planner leverages on a LEC simulation engine and the k-medoids clustering algorithm. The development demonstrate the feasibility of the approach and the results show the effectiveness of planner. It is worth noting that the planner paves the way for the implementation of a decision support system and the design of new business-oriented applications for smart urban and sub-urban areas in the energy domain.**

I. Introduction

The greenhouse effect is caused by the presence of greenhouse gases (GHG) in the atmosphere that hinder the escape of heat from the Earth's surface. The natural greenhouse effect adjusts of the planet's average temperature, while the anthropogenic greenhouse effect caused by human activities primarily the burning of fossil fuels, have intensified the natural greenhouse effect. Since there is a direct link between energy production and GHG emissions, the main producers and distributors are spending effort in developing the so called *Smart Grid* [1], [2] both directly, by using new technologies, and indirectly, by making renewable energy and efficiency programs more affordable and potentially larger. However, a different approach based on distributed micro generation of renewable energy and smart organisation of local users is also possible. The underlying idea is to aggregate local users into energy communities. In accordance to this paradigm a *Local Energy Community* (LEC) is a set of users that share renewable sources and storage resources to satisfy energy demand while minimising carbon emissions and operating costs. The main challenge is to design the architecture and the operations governing the LEC while satisfying the triple sustainability principle, *i.e.*, financial, social and environmental.

The *microgrid* [3], [4], a semi-autonomous grouping of renewable sources, storage resources and end-use sinks interconnected and operated, implements the LEC. Many architectures of microgrid have been addressed in literature [5] in order to better understand their behaviour and to devise optimal control mechanisms to strive upon the benefits of distributed renewable energy generation and shared storage resources. The microgrid can operate connected to the Grid or in islanded mode [6].

A novel framework for smart energy management based on the concept of quality-of-service in electricity is presented in [7]. Specifically, the resident electricity demand is classified into basic usage and quality usage. The basic usage is always guaranteed by the microgrid, while the quality usage is controlled based on the microgrid state. The controller minimizes the operation cost and maintains the outage probability of quality usage, below a target value, by scheduling electricity among renewable energy resources, energy storage systems, and the Grid.

The mismatch between harvested and consumed energy in individual homes is addressed in [8] by proposing an energy sharing scheme among nearby homes within the microgrid. The efficiency of the energy sharing system is ensured by a greedy matching algorithm and a transmission scheduling algorithm.

The control strategy of the storage resource is addressed in [9] to meet short and long-term energy needs while extending the life expectancy of the storage. The solution is given in terms of smart local prediction and local scheduling algorithms.

The problem of properly sizing the energy storage devices is analysed in [10], where the renewable resources are both photovoltaic panels and wind turbines. The solution proposed relies upon two mathematical models built for both the islanded and Grid connected modes of the microgrid.

The storage optimal design when wind and sun are exploited as renewable sources is examined in [11]. It suggests to use batteries and super-capacitors to cover fast power fluctuations without dramatically reducing the lifetime of batteries. A sizing method for the storage resource is also proposed.

All these works address either the single components or the architecture and the operations of the whole microgrid to increase the benefits of its members. These are the main lines of research when dealing with distributed and coordinated micro generation from renewable sources.

In a different way, this work addresses the organisation of a target territory into LECs taking into account the benefits for the community, the investors and the environment. We firstly design a reference model of the LEC based on a shared storage resource and a state machine that rules the energy

The work has been developed by the Smart City Strategic Program of the Istituto Superiore Mario Boella (ISMB) with the advice of Giancarlo Pirani (giancarlo.pirani@gmail.com).

978-1-4799-2449-3/14/$31.00 ©2014 IEEE

flows exchanged within the LEC and with the Grid. An engine of simulation is then developed to validate the model and assess the performance of the LEC. The main contribution is the design and implementation of a planner, based on the simulation engine and a cluster algorithm, that groups the buildings of a target territory into a given number of LECs while maximising an objective function.

II. LEC Model

A. Overview

The territory is divided into L LECs and each one is a set of heterogeneous buildings. In principle, each building has a renewable energy source and a storage. An interconnect enables the energy exchange among the entities within the LEC thus the availability of a shared storage resource. The LEC is also connected to the Grid.

An energy manager is responsible for satisfying the aggregated energy demand of the LEC by leveraging on different energy flows. In principle, the aggregated demand is satisfied by the renewable sources and by the shared storage when the renewable is not enough. In case the renewable production exceeds the demand, the unused portion of energy is stored into the shared storage. However, the manager may purchase the energy from the Grid when enough energy is not available within the LEC to satisfy the aggregated demand and it may also sell energy to the Grid. The manager controls the energy flows by means of a discrete control algorithm that takes an action every time slot t. The time slot t is referred to as the control period of the algorithm.

B. Energy Supply and Demand Model

A LEC is analytically described by the function of energy demand and renewable energy production. Let us define $r(t)$ the aggregated energy produced by the renewable sources within the LEC and $d(t)$ the aggregated energy demand during the control period t. Moreover $p(t)$ is the amount of energy purchased from the Grid and $s(t)$ is the amount of energy sold to the Grid.

Moreover, we assume that the energy production of Grid $G(t)$ is infinite with respect to the energy demand of all the LECs. More formally:

$$\sum_{l=1}^{L} d_l(t) < G(t)\ \forall t \tag{1}$$

where $d_l(t)$ is the energy demand of the l^{th} LEC.

C. Shared Storage Model

The shared storage is made of independent storages that can be recharged and discharged. We assume that the recharging and discharging operations of the shared storage can be performed simultaneously since the single storages can be scheduled and switched to be either charged or discharged. Moreover we assume that the storages are not leaky and we do not consider the loss in charging and discharging, since this amount is usually small as in [7].

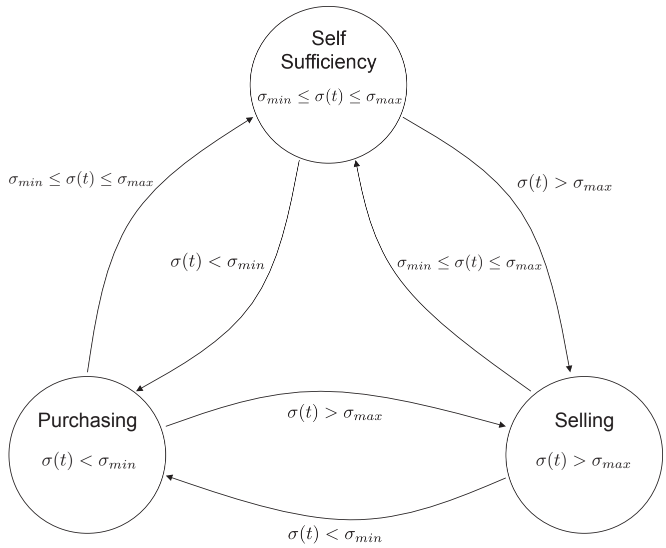


Fig. 1. The State Machine.

Let us define $\sigma(t)$ as a state variable describing the state-of-charge of the shared storage at time t and σ_{min} and σ_{max} two thresholds such that:

$$0 \leq \sigma(t) \leq F \tag{2}$$

$$0 < \sigma_{min} < \sigma_{max} < F \tag{3}$$

where F is the maximum storage capacity.

D. Energy Manager Model

An energy manager exploits a discrete algorithm to control energy streams in accordance to the state machine depicted in Fig. 1. The state machine comprises the *purchasing*, *self-sufficiency* and *selling* states and the transitions depend on the value of the state variable $\sigma(t)$.

When the system is in *purchasing* state the energy demand $d(t)$ is satisfied by purchasing energy from the Grid and the renewable energy $r(t)$ is entirely stored within the shared storage. It follows:

$$\begin{aligned} \sigma(t) &= \sigma(t-1) + r(t) \\ p(t) &= d(t) \\ s(t) &= 0 \end{aligned} \tag{4}$$

Moreover when the system is in *self-sufficiency* state the energy demand $d(t)$ is satisfied by the renewable sources and by draining the energy from the shared storage when necessary. In case the renewable energy production $r(t)$ exceeds the demand the unused energy is stored within the shared storage. It follows:

$$\begin{aligned} \sigma(t) &= \sigma(t-1) + r(t) - d(t) \\ p(t) &= s(t) = 0 \end{aligned} \tag{5}$$

Finally when the system is in *selling* state the energy demand $d(t)$ is satisfied by draining the energy from the shared storage and the renewable energy $r(t)$ is entirely sold to the Grid. It follows:

$$\begin{aligned} \sigma(t) &= \sigma(t-1) - d(t) \\ p(t) &= 0 \\ s(t) &= r(t) \end{aligned} \tag{6}$$

In accordance to (4) (5) and (6) the state variable $\sigma(t)$ may exceed the boundaries $[0, F]$. In that cases $\sigma(t)$ is set either to 0 or to F.

When the system is either in *self-sufficiency* or in *purchasing* state $\sigma(t)$ may get to F due to a peak generation of renewable energy $r(t)$ and a very low energy demand $d(t)$; in that cases a share of the generation may be dissipated. On the contrary when the system is either in *self-sufficiency* or in *selling* state $\sigma(t)$ may get down to zero due to poor generation of renewable energy $r(t)$ and a peak of energy demand $d(t)$; in that cases a share of the energy demand may be not satisfied. Let us define E_d the renewable energy generated but dissipated and E_{ns} the demand not satisfied.

Note that, in a real implementation, the energy manager operates the batteries, forming the shared storage, so that they are subject to complete cycles of charging and discharging not to decrease their lifetime.

III. Optimisation Problems

A. Formulation

The problem is to define the values of σ_{min} and σ_{max} that minimise the carbon emissions and maximise the energy balance during the LEC operation given the total capacity of the shared storage resource F and the aggregated curves of renewable energy production r and energy demand d. Our objective is to solve both optimisation problems ensuring that $E_d = E_{ns} = 0$.

Assuming that the carbon emission is related to the energy purchased from the Grid the problem of minimising the carbon emission can be written as follows:

$$\min_{\sigma_{min}, \sigma_{max}} co_2 \cdot \sum_t p(t) \tag{7}$$

Subject to

$$E_d = E_{ns} = 0$$

where co_2 is a constant that provides the Kg of CO_2 released on average per kWh taking into account the mix of fuels commonly used in energy generation.

At the same time it is important to maximise the energy balance to improve the economic sustainability. Defining the energy balance as

$$\sum_t \quad [s(t) - p(t)] \tag{8}$$

the optimisation problem can be written as follows:

$$\max_{\sigma_{min}, \sigma_{max}} \sum_t [\alpha \cdot s(t) - p(t)] \tag{9}$$

Subject to

$$E_d = E_{ns} = 0$$

where α is the ratio between the purchase and the selling price of 1 kWh by the Grid. Modelling the costs of energy in terms of α rather than in terms of absolute values lets our analysis be more general. In accordance to national regulations, in some cases there are economic incentives on the renewable production sold to the Grid. These incentives can also be taken into account to estimate α.

B. Solution

In accordance to the state machine in Fig. 1, σ_{min} is the threshold below which the energy demand is satisfied by purchasing from the Grid whereas σ_{max} is the threshold above which the renewable is sold to the Grid. Therefore to minimise the amount of energy purchased, hence the carbon emissions, σ_{min} must be as low as possible to anticipate the transition from the *purchasing* to the *self-sufficiency* and σ_{max} must be as high as possible in order to store enough energy to satisfy the demand and to delay the transition from the *self-sufficiency* to the *purchasing* state.

However to deterministically fulfil the energy demand

$$\begin{aligned} \sigma(t) &\geq 0 \\ \sigma(t-1) + r(t) - d(t) &\geq 0 \\ \sigma_{min} + r(t) - d(t) &\geq 0 \\ \sigma_{min} &\geq d(t) - r(t) \ \forall t \end{aligned}$$

The contribution of the term $[d(t) - r(t)]$ in the worst case is equal to $\max(d(t))$ hence

$$\sigma_{min} \geq \max(d(t)) \ \forall t \tag{10}$$

Moreover to assure not to dissipate renewable energy

$$\begin{aligned} \sigma(t) &\leq F \\ \sigma(t-1) + r(t) - d(t) &\leq F \\ \sigma_{max} + r(t) - d(t) &\leq F \\ \sigma_{max} &\leq F - [r(t) - d(t)] \ \forall t \end{aligned}$$

The contribution of the term $[r(t) - d(t)]$ in the worst case is equal to $\max(r(t))$ hence

$$\sigma_{max} \leq F - \max(r(t)) \ \forall t \tag{11}$$

Therefore

$$\sigma_{min} = \max(d(t)) \ \forall t \tag{12}$$

$$\sigma_{max} = F - \max(r(t)) \ \forall t \tag{13}$$

provide a good estimation of the solution to the problem of minimising the carbon emission.

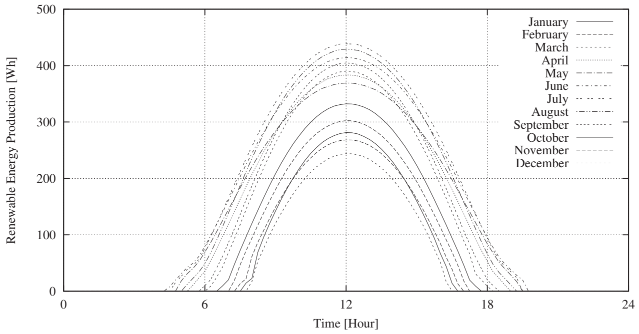


Fig. 2. Set of renewable energy production curves.

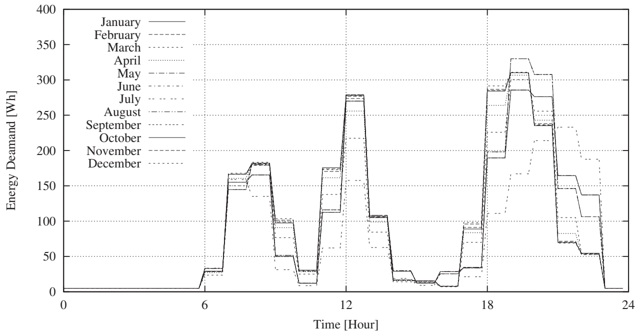


Fig. 3. Set of energy demand curves.

The interesting point is that (12) and (13) also provide the solution to the problem of maximising the energy balance. The following Section provides the validation of the above solution by means of simulation.

C. Validation

In order to validate the solutions given in Section III-B a simulator has been implemented. The simulator takes the aggregated curve of renewable energy generation and energy demand as input to assess the performance of the LEC as function of the shared storage resource. The simulator implements the operation of the supervisor in accordance to the state machine in Fig. 1.

1) Setup: the simulation scenario comprises 6 buildings aggregated into a single LEC as described in Section II-A. The input data for the simulation have been drawn from the real world and are relevant to the area of Turin (Northern Italy). Two types of data have been used: the former are relevant to the solar energy produced by the PhotoVoltaic (PV) system of each building, the latter refers to the energy demand of a typical household of the chosen area. Both series of data consists of values which refer to intervals of 15 minutes along the day with the approximation that the instantaneous power does not change during the whole 15 minutes' interval. Fig. 2 reports the PV-generated renewable energy during a typical day of each month. The energy values are derived from the produced power of a PV installation of 3 kWp of nominal power with assumptions of combined PV system losses equal to 27.1%. The solar radiation is retrieved by the PVGIS[1] database. Fig. 3 refers to the energy consumed by an average family in a typical day of each month. The data of PV-generated renewable energy and energy demand of the single days have been combined to generate a series of data lasting 365 days for each building. Then the series of each building have been aggregated to generate $r(t)$ and $d(t)$ describing the LEC as detailed in Section II-B.

A recent study carried out by ISPRA[2] pointed out that 0.5357 Kg of CO_2 are released on average per KWh taking into account the mix of fuels commonly used in Italy to produce energy. Therefore, given the aggregated energy demand $d(t)$ and assuming to satisfy it entirely by purchasing energy from the Grid, the total emissions of CO_2 would be equal to 9816 Kg.

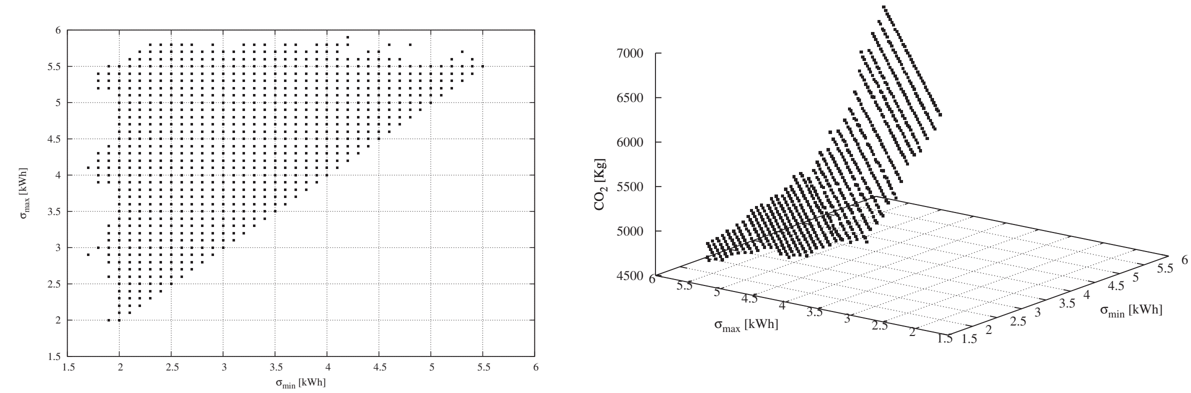
2) Results: in the first set of experiments the buildings in the LEC share a storage resource $F = 8$ kWh. Fig. 4(a) shows the couples of values (σ_{min}, σ_{max}) for which $E_d = E_{ns} = 0$ in simulation. Fig. 4(b) portrays, on a 3D graph, the CO_2 emission as function of the same values. As expected, the emissions decrease with σ_{min} and with the increase of σ_{max}. The absolute minimum equal to 4701 Kg occurs with $\sigma_{min} = 1.8$ kWh and $\sigma_{max} = 5.4$ kWh.

Our analytical solution, (12) and (13), suggests to set σ_{min} and σ_{max} to 2.0 and 5.3 kWh respectively. The total emissions are equal to 4745 Kg by setting these values. This result, very close to the absolute minimum, proves that our solution well approximates the optimal and leads to a reduction of 51.6%. The optimal solution reduces the emissions of 52.1%. The gap of 0.5% may be filled by deploying, in a real implementation, a tracking algorithm in order to measure and to refine the estimation of the maximum difference between d and r over time.

As pointed out in Section III-B the σ_{min} and σ_{max} minimising the carbon emissions also maximises the energy balance. Fig. 5 confirms our claim, in fact, the couple $\sigma_{min} = 1.8$ kWh and $\sigma_{max} = 5.4$ kWh also maximises the energy balance as far as $\alpha < 1$. When $\alpha \geq 1$ the balance is maximised with $\sigma_{max} = \sigma_{min} = 2.1$ kWh. In this condition the demand is almost entirely satisfied by purchasing from the Grid and the most of the renewable energy is sold to the Grid. This result suggests that exists a value of α for which the storage is not more required and it is more convenient to sell the renewable energy to the Grid than using it to satisfy the demand. However, since this condition is not likely, it is important to assess the energy balance achieved with (12) and (13) when $\alpha < 1$. Fig. 6 plots the optimal energy balance and the one achieved with the proposed analytical solution $\sigma_{min} = 2.0$ kWh and $\sigma_{max} = 5.3$ kWh. The maximum gap is very small and decreases as α increases.

[1]Photovoltaic Geographical Information System (PVGIS) http://sunbird.jrc.it/pvgis/

[2]The Italy Institute for Environmental Protection and Research (ISPRA) is a national public body, established in 2008, subject to the vigilance of the Ministry for Environment, Territory and Sea.



(a)

(b)

Fig. 4. The values of the couples (σ_{min}, σ_{max}) for which the energy demand is deterministically satisfied and the renewable energy produced never dissipated (a) and the relative CO_2 emissions (b). F=8 kWh in simulation.

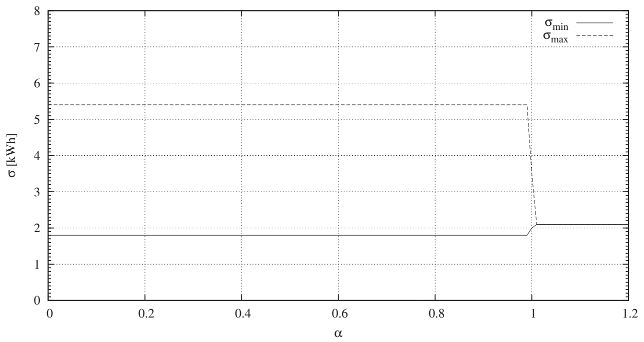


Fig. 5. σ_{min} and σ_{max} values maximising the energy balance as function of α.

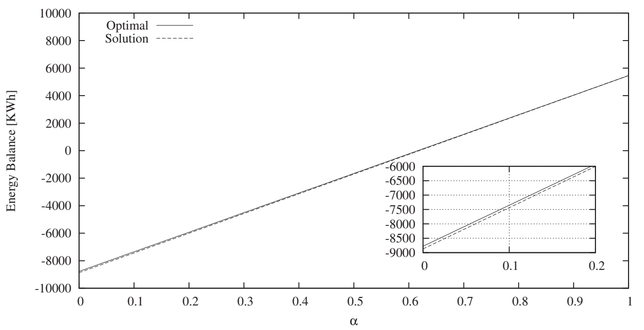


Fig. 6. The energy balance as function of α; the small picture shows the zoom for small α. For the sake of generality the balance is plotted in kWh; the y-axis can be easily multiplied by the price of the kWh purchased from the Grid to obtain the value in euros/dollars.

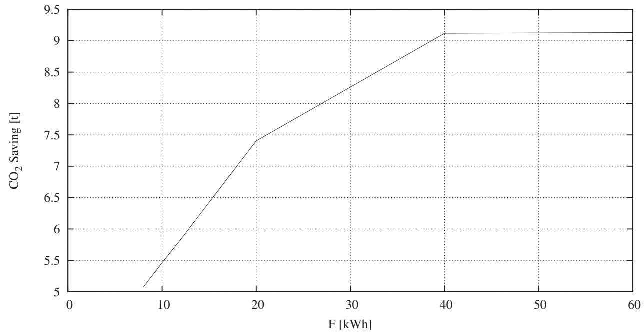


Fig. 7. CO_2 saving as function of shared storage capacity.

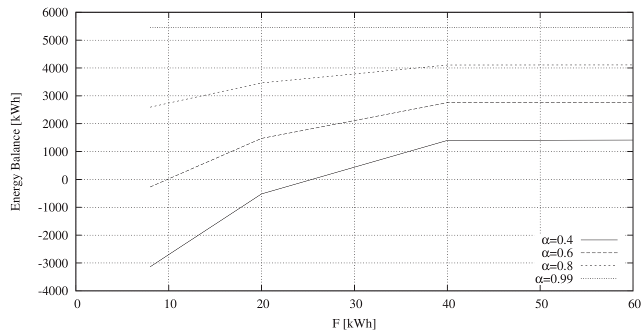


Fig. 8. Energy balance as function of shared storage capacity for different values of α.

In a second set of experiments the performance of the LEC is evaluated as a function of the shared storage capacity F. Given F the thresholds are set in accordance to (12) and (13) in all the experiments. Fig. 7 shows the amount of CO_2 saving per year as a function of shared storage capacity. The saving increases with the storage capacity until it reaches a maximum; from that point the saving is almost constant. The curves of energy balance in Fig. 8 evince the same trend. Moreover, the results in Fig. 8 again suggests that the prominence of the local shared storage decreases inversely to α.

IV. LEC Planner

A. Clustering Algorithm

The simulator of LEC has been exploited to assess the performance and to validate our solution to choose the right thresholds of the shared storage. However it is also intended as a simulation engine of a LEC planner. The planner evaluates a territory to cluster buildings into a given number of LECs. The planner systematically assesses different clusters of buildings to find the best solution in term of an objective function.

The planner leverages on the k-medoids algorithm [12]. In principle, this algorithm attempts to minimise the distance between points labeled to be in a cluster and a point designated as the center of that cluster. The number of starting medoids sets the number of clusters the algorithm attempts to create. Moreover, it may work with an arbitrary definition of distance, different by the simple geometric one.

In our application, the points are the buildings and the concept of distance is redefined as the performance of the corresponding cluster including that point. Thus, given L starting medoids each building i is aggregated to cluster $l \in [1, L]$ for which the performance c_l adding building i is the maximum. Let us define the cluster performance as it follows:

$$c_l = \sum_t [d_l(t) - p_l(t)] + [\alpha \cdot s_l(t) - p_l(t)] \qquad (14)$$

The performance is calculated by the simulation engine during the planning operations. In order to choose only the buildings that bring a positive contribution to the clusters the choice also subtends that $c_l \geq 0$.

The size of the interconnect is also taken under control by bounding the width of each cluster such that the average geometric distance among the points $I \leq x$. This choice lets the planner be more general from the topology viewpoint. In fact whatever the topology, from star (N connections) to full mesh ($N^2 - 1$ connections), the total size of the interconnect can be easily calculated.

After having calculated the first solution, the algorithm changes the starting medoids one at a time to evaluate new clusters and find an optimal solution on the territory. As a result, the planning operations subject to the objective function in (14) can be described as the following optimisation problem:

$$\max_{\text{clustering}} \sum_{l=1}^{L} c_l \qquad (15)$$

$$\text{Subject to}$$

$$\{F, \sigma_{min}, \sigma_{max}\},\ I \leq x$$

B. Planning Results

Fig. 9 shows the building positions over a real target territory near the area of Turin. In this experiment there are 55 buildings and each one is described in terms of renewable energy production and energy demand by using the data series described in Section III-C1; the data series last for 365 days. The data series of each building i is multiplied for a random variable uniformly distributed in $[1, 2]$ to consider a scenario where $\sum_t r_i(t) / \sum_t d_i(t)$ are different. The planner is configured to create 4 clusters with a width in terms of average distance among buildings $I < 100$ m, given the capacity F of the shared storage resource.

Fig. 10 shows the solution the planner converges to. The planner identifies the four clusters that maximise the overall performance. The cardinalities of the clusters are different and the size of the interconnect is bounded as configured. As expected not all buildings are clustered either because of the constraint on the interconnect or because they do not provide a benefit to the cluster in terms of performance, *i.e.*, in terms of the objective function.

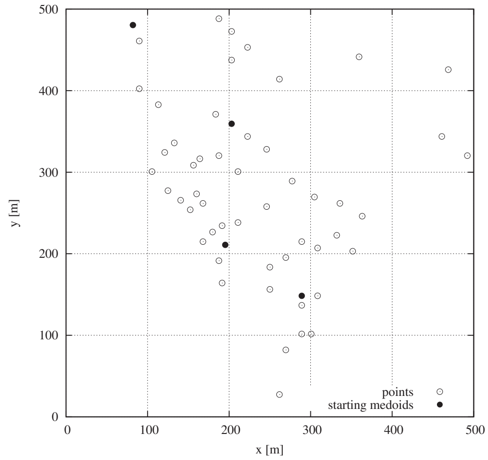


Fig. 9. The target territory in terms of building positions; full points represent the starting medoids.

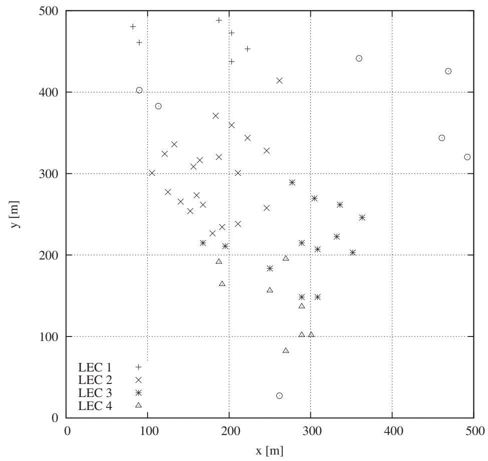


Fig. 10. The planning results.

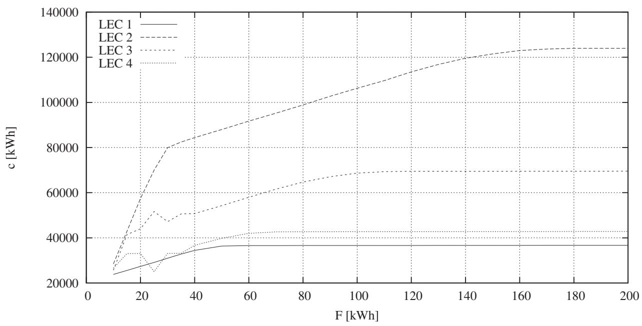


Fig. 11. Performance of the four clusters as function of F.

TABLE I
PLANNING RESULTS

F [kWh]	Performance [kWh]					Cardinality [#]				Interconnect [m]			
	L1	L2	L3	L4	TOT	L1	L2	L3	L4	L1	L2	L3	L4
10	23780	28450	25671	26566	104467	6	9	8	9	76.96	65.04	49.44	78.06
15	25563	42860	41408	32995	142826	6	13	13	11	76.96	87.63	98.38	86.51
20	27323	57383	44040	33069	161814	6	18	12	10	76.96	97.13	83.76	99.97
25	29115	69950	51638	25031	175735	6	21	15	6	76.96	99.03	90.28	75.57
30	30992	79941	47152	33102	191188	6	21	13	8	76.96	99.03	95.50	80.16
35	32796	82465	50577	33170	199008	6	21	14	7	76.96	99.03	92.48	74.63
40	34433	84410	50768	36695	206305	6	21	13	8	76.96	99.03	95.50	80.16
50	36372	88009	54256	39794	218425	6	21	13	8	76.96	99.03	95.50	80.16
60	36560	91708	57954	41997	228220	6	21	13	8	76.96	99.03	95.50	80.16
70	36571	95199	61567	42637	235973	6	21	13	8	76.96	99.03	95.50	80.16
80	36581	98891	64703	42683	242858	6	21	13	8	76.96	99.03	95.50	80.16
90	36591	102703	67065	42694	249052	6	21	13	8	76.96	99.03	95.50	80.16
100	36601	106266	68673	42703	254243	6	21	13	8	76.96	99.03	95.50	80.16

Table I shows the numerical results of the planning as function of the shared resource capacity F. The planner identifies the optimal solution in terms of aggregate performance for each value of F. It is worth noting that the cardinality of the clusters varies with F until the planner converges to a clustering solution. The convergence arises when $F = 40$ kWh, in fact, from that point on, the average distance among buildings within the same cluster does not change.

Fig. 11 shows the performance of clusters as function of F. The performance of each cluster increases towards a maximum value that is directly proportional to the cardinality of the cluster. Therefore, while the planner converges to a clustering solution the optimal capacity of the shared resource storage in each cluster is different. For example, F can be chosen as the minimum value that maximises the performance. In this scenario $F_{L1} = 50$ kWh, $F_{L2} = 160$ kWh, $F_{L3} = 100$ kWh and $F_{L4} = 60$ kWh. Under these settings the LECs lead to an overall CO_2 saving equal to 107.48 t and a total energy balance of 64.1 MWh over one year.

V. Conclusions and Future Work

This paper has presented the concept, and functional model of a Local Energy Community (LEC). The main contribution was the design of a planner that systematically assesses a target territory to cluster buildings into LECs while maximising an objective function. The planner leverages on a LEC simulation engine and the k-medoids clustering algorithm. The development has demonstrated the feasibility of the approach and the results showed the effectiveness of planner. The planner paves the way for the implementation of a decision support system and the design of new business-oriented applications in the local energy domain.

The future work will be devoted to devise new objective functions taking into consideration further technical details, costs of the infrastructure and financial aspects governing the investment.

References

[1] H. Farhangi, "The path of the Smart Grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, Jan 2010.

[2] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart Grid - The New and Improved Power Grid: A Survey," *IEEE Communications Surveys Tutorials*, vol. 14, no. 4, pp. 944–980, Oct 2012.

[3] R. Lasseter, "Microgrids," in *IEEE Power Engineering Society Winter Meeting*, vol. 1, Jan 2002, pp. 305–308.

[4] R. Lasseter and P. Paigi, "Microgrid: a conceptual solution," in *IEEE Power Electronics Specialists Conference*, vol. 6, Jun 2004, pp. 4285–4290.

[5] H. Jiayi, J. Chuanwen, and X. Rong, "A review on distributed energy resources and microgrid," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 9, pp. 2472 – 2483, Dec 2008.

[6] J. Peas Lopes, C. Moreira, and A. Madureira, "Defining control strategies for microgrids islanded operation," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 916–924, May 2006.

[7] Y. Huang, S. Mao, and R. M. Nelms, "Adaptive Electricity Scheduling in Microgrids," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 270–281, Jan 2014.

[8] T. Zhu, Z. Huang, A. Sharma, J. Su, D. Irwin, A. Mishra, D. Menasche, and P. Shenoy, "Sharing renewable energy in smart microgrids," in *ACM/IEEE International Conference on Cyber-Physical Systems*, Apr 2013, pp. 219–228.

[9] D. Tran and A. Khambadkone, "Energy Management for Lifetime Extension of Energy Storage System in Micro-Grid Applications," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1289–1296, Sep 2013.

[10] S. Chen, H. Gooi, and M. Q. Wang, "Sizing of Energy Storage for Microgrids," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 142–151, Mar 2012.

[11] A. Van Voorden, L. Elizondo, G. Paap, J. Verboomen, and L. Van der Sluis, "The Application of Super Capacitors to relieve Battery-storage systems in Autonomous Renewable Energy Systems," in *IEEE Power Tech*, Jul 2007, pp. 479–484.

[12] L. Kaufman and P.J. Rousseeuw, "Clustering by means of Medoids," in *Statistical Data Analysis Based on L1 Norm and Related Methods*, Y. Dodge, Ed. North-Holland: Birkhauser Basel, 1987.

Design for Sharing

Credits

Design for Sharing

final report/working paper

by Ann Light and Clodagh Miskelly

November 2014

supported by a pilot grant from the EPSRC Digital Economy *Sustainable Society Network+*

Principal Researcher: Clodagh Miskelly

Principal Investigator: Ann Light, Northumbria University

Editorial and Design Assistant: Deborah Mason

Advisors:

Robin Mansell, London School of Economics

Susan Steed

Contents

Foreword

Sustainability is often interpreted as actions that ordinary people must give up or constrain: buy less, consume less, waste less… there is a long list of things we should not do so that the planet's resources will stretch further. The negative message sits awkwardly with what we know of how to motivate behaviour.

A sustainable society is also one in which we choose positive behaviours that make us feel happier, more connected and more disposed to help others. *Sharing* is one such activity and it has the benefit that, as well as being a positive action that carries with it intrinsic appeal, it can also support a society that buys, consumes and wastes less as a by-product.

This report, part of the work of the Sustainable Society Network[1], offers insights from three linked studies into sharing practices and concludes with design recommendations for different kinds of sustainability based on sharing.

We look at three contexts: grass-roots initiatives in a south London locale within walking distance of a project researcher's home; sharing activities undertaken by researchers' Facebook 'friends'; and what a range of digital service producers are making to support sharing across the internet. In this way, we are able to compare sharing in ordinary life, as it happens among committed volunteers in small organisations; for digitally-active individuals; and in the hands of digital entrepreneurs.

We hold these contexts up to consideration against different understandings of sustainability to explore the impact of different kinds of design on factors such as social, environmental and economic viability and resourcefulness.

We do this as the 'sharing economy' grows. Hardly a day has gone by this year without some media reference to new digital services. So it seems an apt time to ask what kind of economy is being promoted by the sharing economy and how this tallies with what we found on the ground.

We hope that in exploring the way that sharing is taking place, how digital services are supporting it and what design can contribute to our future through considering different types of sustainability, we offer something for many types of reader. Whether you are running your own grassroots initiative, developing software or promoting sustainability policy, we are addressing you.

Ann Light

Professor of Design

Northumbria University

[1] Funded by the Engineering and Physical Sciences Research Council's Digital Economy programme

Executive summary

This report considers how finite resources, such as tools, produce, time and skills are shared.

We learnt that kinds of sharing in locally-developed initiatives are different in tone, scale, ambition and practice to those in the digital sharing economy. In the former, we see an emphasis on organising together to create shared spaces for collaborative use of resources and joint ownership of projects and places. In the latter, we notice the starting point is often sharer-to-sharer individualised activity, implying personal ownership, while many services monetise value, rather than encourage sharing as a social exchange mechanism. In other words, sharing economy business models feature significant elements of renting, leasing and hiring and many services focus on resource management, sometimes at the expense of community wellbeing; sometimes to provide new services at a larger scale (eg bike sharing schemes).

Overall, we studied design for sharing in three contexts: grass-roots initiatives developed by local change-makers in South London (p20); activities undertaken by 'friends' on Facebook (p44); and what a range of digital service producers are providing to support network sharing (p48). We looked at these in terms of innovation: what people set up locally and why; how digital entrepreneurs are approaching the opportunities that new sensing, managing and connecting technologies enable.

Our research allowed us to conclude with three design strategies for sharing, each promoting a different aspect of sustainability:

- design for optimal resource management (p62);
- design for improved societal relations (p63);
- design for economic (global, national, local, communal) vitality (p64).

Community cohesion is a necessary prerequisite for environmental and economic resilience and involves bringing these three strands together. The local initiatives had this balance at the heart of their work: an interest in environmental, social and economic sustainability, better resource management, increased access to resources and provision of more for less. Supporting this, digital systems can be applied to help manage trust, access, connection, the breaking of social boundaries and social inclusion, while preserving a sense of place. They can do so by providing collective solutions as well as the person-to-person brokering and vetting principally seen now.

The only argument for choosing to promote one sustainability strategy over others is to ensure that no facet of societal development is neglected at the expense of other, more dominant, ones. At the end of this report, we advocate policy that promotes design to support social cohesion, local community development and, with it, the model of local initiative-taking that makes places to live both pleasant and resilient. We do so for several reasons. **First**, some of the narratives of the sharing economy are conflating resource management activities with social cohesion goals, sometimes deliberately as a sweetener to market what is essentially a new form of business that monetises spare capacity; sometimes out of enthusiasm; and, in the media, seemingly because it makes for a better story. This shows the attractiveness of the end state, but does little to create it. **Second**, the market has been designed to take care of its own and social mechanisms need boosting in an economy where global enterprises can map and exploit all sources of revenue with the advent of total monitoring of (and micro-payments for) service provision. And, **third**, because in the narratives we tell ourselves about life in the future, only those where we address issues of inclusion and access to resources in a fundamental way seem sufficient to mobilise the human elements of a world facing the extinction of a majority of life forms, destruction of key natural materials and likely inundation.

We argue here that we need people with initiative and imagination and the resourcefulness to use these if we want to promote long-term societal sustainability and this has to be bred by creating the opportunities for it. Digital tools can be an important part of that, but targeting secondary markets and furthering commodification is only a small part of that potential. The economy of sharing (saving, managing, growing) is wider than the sharing economy and more useful in environmental and social terms. As ever, promoting vitality means providing opportunity that extends beyond those who have assets, confidence and skills, to support people who are already disadvantaged; and beyond individualised sharing tools, since they merely provide new resources for the resource-rich.

Coming at a point when digital technology is increasingly able to offer a convergence of mobile and real-time cloud-based 'big data' processing, sensors and social media, to sense, connect and control, we might speculate on a number of new types of engagement. We can make tools and things that know where they are and who is using them, when and for how long. We can imagine items that share themselves out and see prototypes in the car services that are centrally booked, but locally unlocked and returned. We can envisage time and space more evenly apportioned and sharing protocols managed autonomously by smart systems in our community resources. But inventiveness is best tempered by an understanding of contexts of use, group goals and personal interests. And sometimes technology may hinder rather than promote embedded social exchange. Our report is written to mitigate this possibility and provides many examples of alternative thinking.

Local study

Our seven change-makers in Brockley, South London tell stories of communal initiatives from within. Between them, they helped set up or develop **a microlibrary, a community-asset pub, a community garden, a digital whip-round service, a communal workshop space, a time bank and a conservation society**. Our study revealed seven tactics for success that can be shared (p46-47), just as their good ideas for shared spaces can be the basis of further innovation (p20-39):

One small thing - doing something that doesn't overstretch the capacity of the people involved;
Mobilisation – taking a broad cross-section of people with you as needs dictate;
Growing gains – giving the time it takes to build social systems around new resources;
Rootedness – aligning with the place and culture you are operating within;
Networks in the making – using existing networks to embed and form partnerships for delivery;
Scale and tempo – linking interests at an individual and small group level with the wider scope of the project;
Stepping into the weave – offering help to other organisations and so creating a culture of support.

Digital networked tools

The analysis of sharing economy services (p50-57) raises many considerations for anyone planning digital sharing resources: **How is the cultural dimension of people's experience handled?** What provision is made for people for whom austerity is already a reality? **How is the risk of loss, harm or embarrassment handled? Is there dependency on existing infrastructure for success?** Does the service encourage us to trust each other or trust the site and its procedures (and the wider legal and economic frameworks of society)? **What model of social relations does it embrace in trying to win a critical mass of users?** How are relations with other agencies established and signalled? **What understandings of trust appeal in different contexts?** Does the service augment local capacity or replace it? **How does the rhetoric of the site position the service?** How is sharing resources without encountering other users different from schemes which bring people together?

While sharing economy digital initiatives may be a stepping stone to a more embedded use of technology, with connected devices and internets of things, they also act as powerful agents in our understanding of what it is to share and how we do it, so current distinctions in interpretation matter. It will be important not to neglect more collectivist models in supporting growth of both wellbeing and innovation as part of tackling the uncertain futures ahead.

Philippe, Brockley:

I know some of my neighbours on the street: I know their jobs, I know what they do, but I don't know what skills they've got. And I can't easily find out. I could knock on their doors and ask. But it's not natural to do that. You know, just door knocking is not always helpful…

So there I am in my house and I want to do a piece of writing on digital software and there's a guy down the road who knows all about it and can help me with it, but I don't know. So we're not using the street's assets. When you pool the street's assets, you can get the stuff done because, yeah, I can't do it, but he can.

But for that to happen you have to have a mechanism, a connecting mechanism. Of course, the Brits are very reserved, as well, and so you're not supposed to talk to your neighbours… (I'm not British, I'm French, but I've become Brit by force of habit.) So people don't communicate and I think there's also fear that if I connect with you, you might invade my life a bit, so I've got to put some boundaries there.

The value of sharing means that you have access to new resources and they are people resources as well as material resources - so that, for me, sustainability is linked to the combination and you can start being creative. But you have to trust people and break down the barriers.

There's a block of flats in Paris where people realised that they're all buying broadband. They installed broadband for the whole building, and, with repeaters, saved tons of money. In fact, the money they're saving has allowed them to pay someone to run a crèche based in the building while the parents are at work. It is brilliant.

Yeah, so that's the value of sharing – it becomes very powerful, I think. It has a transforming aspect.

With thanks to all our participants…

Design for Sharing

In recent times we have seen a revival in the language and practice of sharing. This has emerged in the wake of a global recession, with the continued growth of networked and distributed digital tools which enable new ways of managing peer-to-peer relations and the use of individual or shared property. These tools have provided opportunities for file sharing and new forms of virtual exchange, but also offer the means to manage the other things we share: those which are not endlessly replicable, but are part of a limited resource that we wish to stretch further.

This report seeks to bring together many of the complex factors that could and should inform the design of social systems, from the politics, to understandings of sustainability, to on-the-ground lived experience. It presents analysis of two contrasting types of initiative to support sharing: putting the work of local change-makers in the context of markets and new commercial services, and vice versa. Sharing has become a by-word for social, economic and environmental wellbeing and resilience. What does that mean for designing structures to support it?

What is sharing?

The notion of 'sharing' has diverse interpretations, increasingly influenced by the ready duplication of digital code. But the sharing featured in this report is of another kind: that of a local place-based practice which redistributes finite resources, such as time, skills and physical items. This kind of sharing, by definition, cannot be managed alone. Yet, even within the context of one small urban neighbourhood, we found no easy way to define what we were looking at. We saw systems operating at different scales, affected by the purposes and interests of a range of participants and fully embedded in ordinary life. So this section is broken into a review of theoretical, local and new 'sharing economy' definitions to help us delineate the space we are discussing.

In the literature

Belk has written extensively on how we share and what it means. 'Sharing is an alternative to the private ownership that is emphasized in both marketplace exchange and gift giving. In sharing, two or more people may enjoy the benefits (or costs) that flow from possessing a thing… Sharing, as used here, includes voluntary lending, pooling and allocation of resources, and authorized use of public property, but not contractual renting, leasing, or unauthorized use of property by theft or trespass. We can share not only places and things, but also people and animals (to the extent they are ours to share), as well as our ideas, values, and time…. [it is] the act and process of distributing what is ours to others for their use as well as the act and process of receiving something from others for our use. The receipt of shared goods may be for an indefinite or prescribed period of time. And our share may be for our exclusive use or for use by us as well as others.' (Belk 2007 p127). Belk leaves out simple coincidences, such as that we may "share" a common language, place of birth, or set of experiences, since they do not depend on volitional sharing, and we follow suit here.

Belk asserts that there are things that one can share without losing them: he cites 'a song, a joke, a story, our bodies, things we put up on our Web sites, or music files shared on the Web' (Belk 2007 p132). We leave endlessly reproducible types of sharing out of our study here, being principally concerned with the intersection of finite resources and sharing practices. We do include intangible but finite qualities, such as time, enthusiasm and responsibility.

Belk also observes that much economic activity avoids feelings of commitment, whereas sharing promotes it: 'Even though notions of customer relationship management might have it otherwise, in economic theory commodity transactions are balanced with no lingering indebtedness and no residual feelings of friendship. … If we conceive of a continuum involving these constructs, commodity exchange lies at one end and sharing at the other, with gift giving somewhere in the middle.' (Belk 2007 p127).

This is important, as it hints at something intrinsically social to sharing that differs from other forms of transaction. In *Sharing Nicely* (Benkler 2004), Benkler contrasts the rich world of social exchange with the formalisation of markets. He juxtaposes the legal enforcement and crisp, precisely defined interactions of markets with the more contextually rich, less streamlined activity of social exchange mechanisms, which rely on 'tacit, learned, and culturally reproduced capacities to read and interpret social settings… communicating information with great subtlety and nuance.' (Benkler 2004 p315-18). This requires 'tremendous investment, acculturation, and maintenance, every bit as much as markets' (Benkler 2004 p317) along with enforcing mechanisms, which he calls 'social punishing'. These richer, less formalised processes lend themselves to situations where value cannot be expressed simply or formally, where there is vagueness, where resources are spread thinly across a number of people and so on. Instead, 'market transactions systematically require a greater degree of precise information about the content of actions, goods, and obligations, and greater precision of monitoring and enforcement on a *per-transaction* basis than do social exchange systems' (Benkler 2004 p317).

In the local community

The seven local change-makers[2] we interviewed in Brockley had no immediate definition of 'sharing', though each welcomed it as a positive activity, even as essential to making an area feel welcoming. Sharing was seen as part of looking outward to your community, trusting and caring; there was a strong emphasis on the social value, seen in these examples:

'Giving or pooling skills or resources and not necessarily expecting a return for it … I guess when I think of it in terms of what I've done, I view it as me having shared my skills and my time in return for no financial payment but in return for having an interesting neighbourhood and a really good pub on my doorstep and it having been a really useful experience for me because I've learnt a lot in the process and got to meet new people …it's about exchange for things other than cash.' (Tessa, community-owned pub)

'So far as the sharing aspect is concerned, that's the thing that's surprised me, that over the last month not one single person has vandalised it and I think that's because of the goodwill and the notion that every single person, when you give a book technically you're making yourself a bit worse off because you're giving away some of your possessions, but the overall cumulative thing is that everyone becomes enriched by it: lots of people making a very small sacrifice.' (Sebastian, microlibrary)

'The value of sharing is people connecting. It's a social value. I think it goes beyond "I've got a spare drill, you can use that" There are probably several drills in the street that are barely used, so why not share what you have and save money helping each other out? In sharing my drill with you, I'm connecting with you and if I'm connecting with you, I've got potentially a sense of identity with a community of people or a neighbourhood. It also means that, being connected, I have access to information and we can mobilise and we can do stuff together. It means that we're not buying into the consumer society.' (Philippe, time bank)

'I think sharing is a good word. It's not a jargon word. If you start talking about "community cohesion" and so on people switch off. It's a fairly mainstream word… In a sense, it's a word that actually challenges everything that's wrong with our society.' (Clare, local conservation society)

'It brings people together. It makes people happier. It is nice when you walk out on a street and you recognise people … when you know them to talk to and you know their names. You feel more part of a community. I'm convinced people feel happier and more secure and I know that bad things still happen - people get burgled - but it's a much more comfortable feeling. You feel, yeah, part of something.' (Jane, community garden)

[2] There is no good single word that sums up this mix of enterprise and public service across all sectors, since 'social entrepreneur' is now linked with running a not-for-profit, 'social architect' is overstating it, etc.

Sharing is about doing with others and the sum may be greater than the parts. Local sharing involves both practices for sharing the tangible – space, tools, etc. – and sharing of time and skills or labour. But this is crucially built on less tangible sharing of care, responsibility, and vision and values, in some cases memories and history, and certainly trust. These are key factors in what motivates and maintains sharing. We look more closely at these and other factors in the case studies on p20.

It is clear that sharing sits within a set of collaborative and exchange practices and is rarely entirely distinguished from other related practices on the ground – either in description or execution. There is a range of old and new terminology in use such as: barter, gift, exchange, swap, recycle, repair, repurpose, etc., merging in the weave of practices that make up engaged living. In other words, it is possible to find many different practices rolled together in embracing sharing. Some rely on dealing with existing waste, such as foraging. Some are about coming together, as much for the experience as the resourcing. There are also stories of trust abused and of inadequate systems of management and tension between the different goals and interests of people in a neighbourhood.

In the sharing economy

At its outset, the sharing economy took its reference point from new forms of collaborative and open source production made possible by digital networks. In *Sharing* (2011), Aigreen suggests that over the past fifteen years, file sharing of digital cultural works between individuals has been at the centre of debates on the future of culture itself. This burst of sharing activity, enabled by the virtuality of code, digital copying technologies and the Internet, began a new consideration of sharing and its impacts.

However, the emphasis has changed in the last few years, focussing more on consumption than production and co-production. Bellotti et al (2014) note that 'Commercial peer-to-peer service exchange businesses, such as AirBnB, Lyft and TaskRabbit, are expanding rapidly, but their non-profit counterparts are lagging behind'. There is a booming market in new services that are using the language of sharing, collaboration and peer-to-peer exchange for positioning. With help from the media and innovation policy-makers, they are knowingly defining a field to operate in, generating debate about what is and is not sharing. Recognising this, other terms are being used too, such as 'collaborative economy' and 'peer economy'. We review several digital services that promote collaborative behaviour and we explore their self-presentation in detail, below (p48). Here we look more generally at this new sector supported by digital tools and how it is being defined.

Innovation charity NESTA is working hard for clarity in this area and favours the term 'collaborative economy'. Stokes et al (2014 p13) identify five collaborative economy traits:
1. Enabled by internet technologies
2. Connecting distributed networks of people and/or assets
3. Making use of the idling capacity of tangible and intangible assets
4. Encouraging meaningful interactions and trust
5. Embracing openness, inclusivity and the commons

They acknowledge that the final two traits are sometimes more aspirational than evident, relating them to drivers identified through their survey of collaborative economy organisations (p21):
• Creating value out of idle assets
• Rebuilding social capital
• Creating new economic relationships
• Environmental sustainability

(Our own analysis of the gap between rhetoric and service offering can be found on p50)

Rachel Botsman, who popularised the term 'collaborative consumption' (Botsman & Rogers 2010) has repeatedly tried to define relations between the various names in use. 'You may have noticed the terms "sharing economy," "peer economy," "collaborative economy," and "collaborative consumption" being used synonymously. … So, do I think these terms have different meanings? Yes. Are there common core ideas that explain the overlap? Absolutely.' (Botsman 2013). Of these, she describes the 'sharing economy' as an economic model based on sharing underutilized assets (from spaces to skills to stuff) for monetary or non-monetary benefits, largely talked about in relation to peer-to-peer marketplaces, but equally possible for business to customer models. She gives the example of a ridesharing platform that offers a means by which ordinary drivers 'can earn extra money by giving rides to people who need them'. The common themes across all of the new 'economies', she argues, are Distributed Power, Disruptive Drivers (technological innovation in the shape of social networks, digital payments, online identity systems and mobile devices; shifting values as a connected society rethinks ownership; economic realities and environmental pressures) and Innovative and Efficient Asset Utilization. (Botsman 2013). More recently, she and Helen Goulden of NESTA have identified common themes of the collaborative economy as 'access instead of ownership; collective efforts instead of solitary endeavours, [….relying] on vast, distributed networks of people and goods – and the technologies that can build and maintain such networks […and] a public willingness to place trust in people they have never met.' (Goulden & Botsman, in Stokes et al 2014).

However, many commentators have questioned whether there is now any sharing in the activity of the sharing economy, despite the collaborative tag. 'Traditionally, in order to make renting viable the goods available for rent had to be stockpiled … Thanks to digital technology, it's now feasible to do what Zipcar does and disperse the cars throughout the city. Since the cars are dispersed, they're more convenient. But none of this is sharing. My neighbor and I share a snow shovel because we share some stairs that need to be shoveled when it snows and we share responsibility for doing the work. If I owned the stairs and charged him a small fee every time he walked in or out of the house, that would be the opposite of sharing.' (Yglesias 2013) The debate around these practices has raised concerns about 'sharewashing': moments where the terminology of sharing is being used to promote practices which are new modes of selling (Troncoso 2014). It has also been suggested that some services promote de-unionisation, poor hourly rates of pay, casualization of the workforce and tax minimisation - practices that replace social models of old with more commercially oriented ones (eg Asher-Schapiro 2014 or http://uberpeople.net/threads/driving-in-la-since-the-latest-pay-cut.2388/).

Makwana (2013), discussing values in the sharing economy, suggests that 'by attaching too much emphasis on self-interest and personal gain in relation to the concept of sharing, the altruistic aspects of sharing could be undermined and the more benevolent motivations of those who share could be increasingly ignored.' He notes a tendency in the mainstream media to consider the sharing economy almost wholly in financial terms. Makwana's comments come in the context of work by *Common Cause* (Crompton 2010), which argues that using extrinsic motivators (such as financial reward) to promote charitable giving and concern for greater-than-self issues, like the environment, is actually counter-productive. This view has been endorsed by research into how children develop sharing behaviours and a perception of themselves as moral agents (Warneken & Tomasello 2008) and Steed's analysis of the relationship between money and giving (2013).

Returning to Benkler (2004), we can see a clear distinction between the mechanisms of secondary markets and social sharing systems, which he calls *alternative* transactional frameworks possible for use in disposing of excess capacity. In other words, selling on spare capacity (finding a secondary market for your unused room, car or used goods) is not the same as sharing it. If there is no doubt in Benkler's work that these are contrasting economic systems, it is interesting to consider what work the crossover of terms and ideas is doing as organisations attempt to position themselves in a world of networked technology and connectedness. In Stokes et al's report (2014), there is a wide spread of goals, motives and structures, all co-existing in a space still struggling for definition.

The Politics of Sharing

Most earthly resources are already historically apportioned. And although sharing is largely being regarded as a social, economic and environmental good in the world at present, there is a social justice dimension to sharing schemes that the feel-good factor of making an individual contribution can mask. Much of the sharing economy rests on effective use of spare capacity, assuming an ownership model. In much of the rhetoric of service provision, *access not ownership* is extolled. There are other models involving shared responsibility for ownership and maintenance which are not being discussed as prominently. While we do not look closely at the notion of the commons here, we recognise that some structures created by the change-makers in Brockley are common-pool resources, unlike those advocated in the sharing economy. In our study of a local area, much of what was being developed did not belong to one party. Instead, structures for managing shared responsibility have been agreed as part of establishing, maintaining and developing the resource. In the community garden, the workshop, the community pub and the time bank, ownership is distributed across a group, which is a *commons* model (Rowe 2013) with its own design challenges (see Ostrom 1990 for principles of long-enduring common-pool resources, such as clearly defined scope and procedures and inclusion strategies). Some of the simpler arrangements work on an ad-hoc basis, with no need for complex agreements: the sample on Facebook included those who had bought tools with neighbours, such as a commonly-owned ladder, or who shared a nanny. But no system involving multiple parties is ever free of some form of arrangement and talk of sharing foregrounds both the distribution and management of resources.

Ownership matters

With ownership comes power, whether at the level of deciding what to share when and with whom, or in the sense of owning a service that is mediating in the effective use of resources. Private and public sector ownership shapes what and how we share, and what we see as sharing. At one end of the scale, mutual trusts and municipalities are formalised sharing structures that have developed elaborate control mechanisms, such as voting, periods of election and contribution through taxation. Elected bodies and their officers manage resources collectively on our behalf, which removes the need to share some of the most difficult things at a local level, such as waste disposal and the upkeep of roads. It is only recently in the UK that utilities (such as the electricity and water supply) were privatised, shifting shared ownership of these resources by the whole population of a country to shared ownership amongst those who pay for shares: an example of the different kinds of sharing conferred by ownership. What is now understood as sharing is more often chances for communal redistribution not addressed by state and big business, although municipal services emanate from the same impulse to co-manage and maximise resources in a given area. Still highly organised, but in the hands of the immediate communities of an area, are the shared ownership models, such as cooperatives and community assets, now supported by the Localism Act (2011).

In North American culture, Belk suggests that 'Materialism, possessive individualism, and the conviction that self-identity must be developed by extension into possessions are all factors that inhibit sharing' (Belk 2007 p136). Identification with material assets is prevalent in the UK too and, Sennett suggests, is heavily class-based (2013). Conversely, says Belk, when we feel a shared identity with others—whether in our neighbourhood, group, city, state, or nation—we also feel a common sense of moral obligation toward them and this wider perception of identity supports sharing across family and other social boundaries. Nonetheless, identity is related to trust, raising issues of belonging, inclusion and difference in geographical areas of greater diversity. Ownership, even of shared resources, may stay in the hands of some at the expense of others.

Can bonds built by sharing reinforce existing patterns of ownership to the detriment of wider circles of inclusion? Possibly, though, in a US-based study on neighbourhood cohesion, Hipp and Perrin (2006) found no weakening of wider ties through the development of stronger local ones, while the ‘sharing of resources among residents allows all “boats to rise together” and keeps any households from suffering unduly’. They looked at both neighbourhood and city level cohesion, finding, predictably, that respondents expressed a higher average degree of cohesion with their neighbourhood than with the larger community, though specifically *not* to the detriment of wider relations.

Social justice matters

Considering assets and their distribution, Coote & Goodwin note that even something as ostensibly egalitarian as sharing can ‘privilege some people over others (for example, where better-off parents share a car pool to ferry children to improving after-school activities)’ (2010 p4). They list reasons that individuals and groups may be excluded or disempowered: because of how much discretionary time they have, where they come from, where they live, or their state of health. In a comment that is pertinent to our study of Brockley, they observe ‘Some neighbourhoods seem to be awash with activities that enrich and strengthen social connections, while some appear beset by divisions or distrust, or have less opportunity for social exchange, because there are no meeting places, or populations are transient, or fear of violence keeps people indoors.’ (Coote & Goodwin 2010 p4-5).

With individualised sharing systems comes a danger that well-resourced people are able to lend and borrow as a lifestyle choice while stigma attaches to forms of exchange in less affluent contexts. This doubly penalises people who lack resources, embarrassed by the attention suddenly given to *making do* practices, even as they could benefit most from both the economic and social aspects that sharing systems confer on a neighbourhood. State management of scales of contribution, in the form of graded local taxation, takes this inequity into account in ways that many local schemes do not. Even an assumption that everyone has equal time to give to a project or confidence to participate may act to the prejudice of some participants, while time-banking schemes have long suffered from a perception that some jobs are not equal to others, even if they take as long to do. For instance, does an hour of gardening equate to an hour of legal advice?

A detailed study of mutual aiding in Leicester looked at support given to kin and non-kin in two areas of the same English town, comparing a relatively affluent ward with one that features high on the Index of Multiple Deprivation (White 2009). White is one of few commentators to look at the issue of *receiving* resources and the pressure this brings, not least if you have fewer resources to offer in return, citing examples of people who did not accept offers of help despite signs of need.

But in his study of mutual aiding (including sharing, lending and borrowing) he found that far from being an austerity measure or linked to one social echelon or another, ‘unpaid reciprocal activity forms a significant part of daily coping strategies’ across the board. In other words, he says, it is not, as many have suggested, a last-resort measure in households with limited resources. Mutual aid was apparent in both the ward of Saffron [‘deprived’] and West Knighton [‘affluent’] (White 2009 p463). He finds complex motivations for this, but his evidence allows him to attribute much of the eagerness to share as coming from a desire to help other people, feel needed and create a sense of belonging: ‘this over-riding need for many respondents to feel wanted by others, resonated deeply in both affluent and deprived households’ (White 2009 p467). He argues that this ‘social embeddedness’ is a major reason that mutual aiding has been able to resist commodification in contemporary society.

What happens when access to the means of participation is mediated by networked tools? As NESTA puts it: ‘The role of digital technology in facilitating activities within the collaborative economy raises questions not only about who is (and is not) participating, but, more specifically, whether technology is either enabling more people to collaborate, or a smaller number of people to increase their collaborative activity.’ (Stokes et al p21). In other words, another dimension of access to resources is access to the mechanism distributing them. Meanwhile, a recent study of Airbnb (www.airbnb.com) shows a different form of discriminatory impact in using current digital tools. Edelman & Luca (2014) combined pictures of all New York City landlords on Airbnb with their rental prices and controlled the study with other information about quality of the rentals (eg location, size, etc). They show that non-black hosts charge approximately 12% more than black hosts for the equivalent rental, 'suggesting an important unintended consequence of a seemingly-routine mechanism for building trust [by using pictures]'.

Production matters

Sharing, in being based on a non-financial model, is part of what Goodwin has named 'the core economy'. The notion of the core economy was introduced to balance an emphasis on financial exchange in economic theory by stressing the role of community relationships and resources and the work of the family (Coote & Goodwin 2010). Co-production advocate Cahn suggests that something like 40-50% of productive economic activity takes place outside of the market and is not measured by traditional indicators. He quotes futurist Alvin Toffler using this idea to shake up 'the CEO's of Fortune 500 companies: "How productive do you think your work force would be if it was not toilet trained?" …No society has the money to buy, at market prices, what it takes to raise children, make a neighborhood safe, care for the elderly, make democracy work or address systemic injustices.' (Cahn, n/d). Instead, such work is (co)produced in domestic and social arenas of life, with different organisational structures and criteria, as Benkler and others note.

Sharing practices have a strong relationship to co-production systems. First presented by Ostrom as a model of public service delivery in 1978 (Ostrom et al 1978), she uses the transition to co-production in choosing a condominial system for a block of flats in Brazil to show how sharing space can involve active participation: 'teams first set up a series of neighbourhood meetings where a general overview of the process, opportunities, and costs of a condominial system is presented. Then, meetings are held in each block where detailed discussions centre on the choices that residents will have to make', noting significantly that 'Block meetings are called off if half of the households on a block are not in attendance to ensure that there is wide availability of relevant information and good discussion among those living on a block.' (Ostrom 1996 p1074-5).

Supporters suggest a co-production approach can result in a 'novel solution to a social problem that is more effective, efficient, sustainable, or just than existing solutions, and for which the value created accrues primarily to society as a whole rather than private individuals' (Phills et al 2008 p34). Evans (1996) writes that '[m]arket advocates see it as hopelessly muddying the logic of individual incentives and rational resource allocation.' Much of this hinges on the value given to social process. Co-production stresses the role of citizens and communities and their active stake in producing solutions (Stephens et al 2008). This relationship fosters 'self-esteem, personal aspiration and a sense of purpose' (Stephens et al 2008 p17) while providing experiential learning (Boyle et al 2010) and thus developing communal resources.

We might see some of the recent 'sharing economy' digital tools promoting secondary markets as closer to co-production of services, than *sharing* per se. Linked back to Coote & Goodwin's (2010) comments about resourcing, we can also see that success in the monetary economy privileges the sharing of material resources in both financial and core economies, because it gives access to a greater number and variety of assets. And, last, we can note that the production of new means of sharing and new spaces and opportunities to do so feeds straight into the core economy, providing value in terms of capacity and learning for the community.

So, why design for sharing?

The market is usually allowed to decide what technology flourishes (eg Brynjolfsson & McAfee 2014). But there are warnings, such as Benkler's (2004) that letting financial mechanisms crush sharing initiatives will be counter-productive: 'If indeed we live in an economic system made up of price-based, hierarchy-based, and sharing-based modalities of production, if it is true that optimizing our institutional system for price-based production undermines productivity in the sharing modality, and if it is true that our communications, computation, and information sectors are undergoing technological changes that improve the efficiency of social sharing, then we are making systematically mistaken policy choices not on the peripheries of our economies and societies, but at their very engines.' (Benkler 2004 p281). And there are pleas to develop a different set of economic criteria for policy decisions: 'We want to move from an economy based on scarcity of economic resources to one based on an abundance of human resources. We also want to move beyond a deficit model of need where we simply pay attention to problems that require fixing, to a more rounded and positive approach, where we consider what we need to lead a good and satisfying life – not only what makes life possible, but also what makes life delightful.…to promote sustainable social justice and well-being for all' (Coote & Goodwin 2010 p4).

So there are the economic and the social arguments for looking at the forms of exchange in our societies. There is increasingly another argument too. The urgency of environmental issues draws our attention to the management of finite resources, and the potential of digital tools to help us work with them effectively. At its crudest, this can be seen in terms of simple resource management. But sustainability has social, environmental and economic elements, and some would add cultural too (Mortberg & Stuedahl 2012). While sharing extends the capacity of households and communities to function effectively, demanding fewer resources and spreading them further, it also generates social and cultural connection, such as knowledge and friendship, extending networks. Gauntlett concludes an analysis of happiness by suggesting that for activities to be meaningful, it is 'especially valuable if they are not contained at the individual level but involve some form of sharing, cooperation or contribution to other people's well-being' (Gauntlett 2011 p126). These activities increase the sense of interdependence and collaborative spirit that is the basis of both fulfilment and resilience in crisis, going to the heart of what makes society sustainable. The emphasis, in sharing practices, on positive messages, intrinsic rewards and on-going gratification is motivational in making local changes to address peak oil and climate change globally (Heise 2008), contrasting with the more finite regime of technologies that focus on monitoring and immediate energy reduction. Sharing presents people with a positive activity, offering engagement that has social benefit and builds cohesion in the face of stress. This means it emphasises resourcefulness as well as the immediate environmental value of using fewer resources in the short-term.

We might also ask what it is about the UK in the early 21st century that means we need to rediscover these values. Writing in 2008 about North American culture, Belk comments that the pessimistic view would suggest Western societies are headed towards less, rather than more, sharing. He cites the enclosure of intellectual property through rights management hampering the sharing of scientific discovery. He notes new trends in the family: 'What were once the family radio, the family car, and the family television are increasingly privatized property of individuals within households. And the family meal is becoming a quaint memory.' (Belk 2007 p127). Sennett's work *Together* identifies several major changes in how Britons and Americans live that undermine our ability to cooperate and interest in doing so (Sennett 2013): structural inequalities, short-termism (with the casualization of working patterns) and a resultant loss of 'character' as people withdraw rather than meet social challenges. If cooperation is a craft, as he argues, then we are not working hard enough to develop it.

Another series of factors driving change in the UK are specifically political. In 2010, the Big Society was launched by the incoming coalition government, with the rhetoric of a more engaged citizenship and sweeping cuts in third sector, arts and public sector budgets. 'Austerity culture' became fashionable (Bramall 2013) and a spur for design, reflected in initiatives such as frugal design (http://www.nesta.org.uk/frugal_innovations), but was also criticised heavily as increasing inequality (Slay & Penny 2013). Into this mix, the Localism Act (2011) made it possible for well-organised and well-connected community groups to finance the joint purchase of major buildings and other assets that were previously unavailable to them as a collaborative venture. The legal framework had been modified slightly in favour of shared ownership.

Hardship; top-down promotion of DIY values; and bottom-up growth in anti-consumerist sentiments have produced a turn to alternative economies and means of exchange. In the *new materialism* advocated by the New Economics Foundation, an economy of 'better, not more' (Simms 2012) includes encouragement to: 'look at all your things, think about what your friends might need or could benefit from, and share at least one thing a week' (Simms & Potts 2012).

By evaluating how sharing works with and through the use of digital technology; which aspects are better managed more traditionally; and how different materialities in different social and economic contexts demand different design approaches, we can discover the potential for new light-weight digital interventions that not only connect people, but also physically-bounded resources. We can use technologies that will not worsen the environmental footprint of the digital sector, or at least provide a trade-off in improved environmental outcomes. We have an obligation - in this area of technology development even more than others - not to waste resources by innovating for innovation's sake. And we are warned not to conflate means with ends in our rush for solutions. For instance, Hollis, on Shareable (Hollis 2014), specifically criticises technological 'solutionism' as part of the problem. At the heart of necessary adaptability to bounce back from adversity is not some technological innovation, he argues, but 'trust - the essence of the understanding that the places where we live and the lives that we lead are shared experiences, rather than tradeable properties'. This kind of trust, he says, is nurtured in equality. 'It is not, as some thinkers will tell you - such as Francis Fukuyama or Robert Putnam - based upon some social transaction or the reward for one kind of participation or other, but rather the dissolution of the difference between us and them.' (Hollis 2014).

In sum, sharing is a positively-orientated sustainable behaviour which has persisted for reasons of both need and life-choice, offering the potential to reflect many different aspects of sustainability. By understanding more about willingness and opportunity to share, we address sustainability as an interwoven social, economic and environmental good. This begs the questions: how do we design social, technical and business systems to best support sharing? And are there subtleties to what we emphasise that affect the kind of sustainability being supported?

Design factors

In addition to reasons why we might promote practices of sharing in their various forms, we can also consider design issues more closely. In doing so, we notice that even where business models, social support networks and types of transaction are not changing, there is change in domestic patterns and, with these, shifts in what might be or is shared. Car sharing schemes are on the increase, but launderettes have almost disappeared in much of the UK. A launderette chain shares the business model of Zipcar: a company which provides locally-situated resources available for people to use at a small fee and for an allotted amount of time – offering this at a point when household washing machines were an unnecessary luxury. But a launderette and a car sharing service have totally different technical requirements. As the objects that we share change, so too must our designs.

And we are working at a time when the evolution in technology means the potential for new ideas and opportunities. The Digital Social Innovation report (Bria et al 2014) introduces developments in hardware and software across five areas:

(i) New ways of making; including 3D printing, maker movements, etc;
(ii) Participatory mechanisms and open democracy, featuring new projects pioneering direct democracy and citizens' participation;
(iii) The sharing economy 'that includes crypto digital currencies and platforms, crowdfunding methods, exchanges and new economic models';
(iv) Awareness networks enabling sustainable behaviours and lifestyles, such as citizen science projects using sensors and collective feedback;
(v) Open access and information commons and mesh networks projects (Bria et al 2014 pii).

In looking at what support technology can offer, we turn to Janelle Orsi's analysis of managing different complexities of sharing practice (Orsi 2009): as sharing becomes more institutionalised, the time and resources needed are greater and more people benefit (though more intangible value may diminish). There is also a need for greater infrastructure and much of this, nowadays, is digital.

- **First Degree: Requires Cooperation + Minimal Planning: '**At the most basic level, sharing arrangements require little planning, time, or money.' She gives examples of sharing with one other person, such as borrowing and lending goods; dog walking exchange and harvesting and sharing fruit from neighbourhood trees.
- **Second Degree: Requires Cooperation + More Extensive Planning:** This entails administrative work and, likely, a written agreement among sharers, **'**Sharing ownership of a car with a neighbor, for example, takes shared transportation to this second level' and other examples include sharing garden space for food cultivation and a neighborhood home repair group.
- **Third Degree: Requires Cooperation + Extensive Planning + Infrastructure:** The need for tools, like an online calendar for scheduling, grows. 'As a result of creating infrastructure, third degree sharing arrangements often have an identity independent of their individual members ….even when there is complete turnover, the sharing arrangement remains and becomes a lasting community institution.' An example might be cooperatives that facilitate sharing of resources.
- **Fourth Degree: Requires Cooperation + Extensive Planning + Infrastructure + Community-Wide Restructuring and Mobilization.** Examples here might be expansion of public library systems to include lending of tools, equipment, and other goods and the planning of neighbourhoods and design of housing to facilitate extensive common areas and community interaction.
(summarised from Orsi 2009)

As well as networks enabling simple brokering between people, managing payments and sharing information, which the internet does well, the transition to networked objects (aka Internet of Things) increases the scope of tools and provides the opportunity for remote access and things that have their own functionality (see Rowland et al 2015). 'Everything is being connected: cars, cargo containers, street lighting and parking spaces …Automatic data interchange between connected devices, or machine-to-machine communication (M2M), is already firmly established in nearly all industries,' says Kiessling (2013), suggesting that networking of objects is an important precondition for sharing because it allows us to track things and people.

Friends of the Earth's *Sharing Cities* (Agyeman 2013) sets out to redefine what future 'smart cities' could mean: harnessing smart technology to an agenda of sharing and solidarity (rather than one of 'competition, enclosure, deepening inequality and division'), a theme also to be found among the many policy and design ideas for sharing in urban contexts in *Policies for Shareable Cities* (Orsi 2013).

Elsewhere, social media - blogs, social networking sites and so on - make available the means to communicate about and quickly promote opportunities and hyperlocal activities without recourse to more formal media mechanisms.

In reviewing the argument for upgrading our designs for sharing, we must also sound a note of caution. People interested in the development of new technologies may be particularly ill-equipped to judge how little some people like using digital tools, especially in their leisure time, and how many people still do not have easy access to them. As new technologies are created, divisions in usage will intensify.

Informal design processes

Our study features the work of seven local change-makers. All have shown remarkable initiative, though only three are paid for their efforts and these three are as much advocates, architects and initiators as the other four. In this regard they are not unique, but join a legion of other motivated lead citizens, activists and agents of change who bring into being new social structures and means of engagement because they can see a shape and value to them. They are not urban planners or designers, whose work it is to do so. They are people who have felt sufficient passion to act. Hester terms these actions 'labours of love'; projects initiated by community-based innovators which have significant impact on urban spaces and which are 'born of personal creative necessity, thrive where there are scarce resources, and produce flexible environments that are lovingly human' (Hester 1984). If we want to understand design for sharing, we must balance a study of technological opportunity with analysis of the practice of people enabling sharing and building infrastructure on the ground. These are the people who are taking sharing from the level of serendipitous encounters to a more organised and mobilising series of engagements. Some are using digital technology, as does Olivia of Patchwork Present in our sample (p28), and some require no digital tools at all.

Our earlier studies (Light & Miskelly 2008, Light Miskelly & Thompson 2008, Light & Akama 2014) explored informal processes that, nonetheless, in a deliberate and thought-out fashion, lead to new infrastructures and change the social and physical characteristics of the local environment. We observed a fluid and responsive design process among these practitioners. We asked how practices which take place outside the usual context of design (what we called 'designing-in-the-wild') develop without recourse to the formal tools that planners, architects and designers use.

Dorst (2008) argues that the collaboration, context, interaction and learning which inform design practice are often ignored. These ancillary aspects become central when working in informal contexts with social change, when the designer is also advocate, facilitator and, in Louridas' (1999) words, bricoleur. This speaks of a crafting that is often at the mercy of contingency, making do by reinterpreting and being inventive with the tools to hand. 'The work is often characterised by passionate dream, sacrificial struggle, allies, a campaign of education and visible results, followed by a period of transfer, adaptation and institutionalisation of power which is necessary but painful for these individuals but which ensures a community-owned project' (Light & Miskelly 2008). We revisit our analysis in the context of this new project, use it to help understand what the change-makers of Brockley have achieved and to site their work in the wider movements affecting 21st century living.

Three Studies of Sharing

Sharing is an age-old practice, but personal and communal initiative can promote opportunity and willingness to share. We explore that in three ways: grass-roots initiatives developed by local change-makers in a south London locale; activities undertaken by 'friends' on Facebook; and what a range of digital service producers are making to support networked sharing.

Context

Sharing of resources happens in a place. Social relations, power structures, physical distance, local geography and the nature of what is being shared all impact on what is possible and desirable. This is a challenge for designing support structures since contexts vary and solutions may not scale or work outside a particular location. We address contextual features by conducting three studies, each with a different relation to place. The first was a highly located study of communal initiatives, in a small part of south London, where we deliberately used social networks on the ground and geography to shape the study and decide who we interviewed. The second applied the same principle - of using our own networks as a basis for the research - but situated the study in the digital realm, using Facebook to ask a widely flung set of individuals questions about sharing practices and so ensuring answers came from digitally active types. The last study analysed a range of digital services intended to promote sharing in some form, looking at how they pitch themselves, how they negotiate where sharing takes place from a disembodied position and what their business model involves.

Methods

Interviews in a Locale

We conducted seven detailed interviews in situ with people who could tell the story of local communal initiatives from within and explain their motivation in getting involved in making change. The seven initiatives were a microlibrary, community-asset pub, community garden, digital whip-round service, communal workshop space, time bank and conservation society.

This list was drawn up after initial research into community action in a particular locale and who the key 'movers and shakers' were in setting up collective activity there. While this was not an exhaustive list of active local people or facilities in the area, our research was governed by what was energetic and relevant at the time of exploration and what would give a good overview and spread of issues, relying on the network and judgment of the principal researcher in role as participant-observer. It will be seen that the sample initiatives are varied and represent a good range of collective activity possible in the UK, concentrated in one small area of London. In our analysis, we consider both the individual initiatives and also the reason for the concentration in this part of London, noting that place and other contextual features have a bearing here too.

A constraint we imposed was that only those initiatives taking place within a small area were considered: we extended our locus of inclusion to just beyond the Brockley ward in the London Borough of Lewisham, defined by the principal researcher's reach. We settled on a reach of no more than *the area that she can walk from her house in 30 minutes*, this being a likely limit on most people's willingness to travel to collect items from others or participate in communal activity.

The interviews were informal, but each participant was asked about definitions of sharing and use of information and communication technologies (ICT) as closing questions. A typical opening, which demonstrates the engaged and informal nature of the interviews, is as follows:

"It's a very informal interview approach that I use so it's more like a guided conversation. I have a few questions to guide it but I'm happy to take whatever direction you want, to what you think is of interest and how you see sharing. I haven't got a tight definition of it, but to start off, it would good

to know a little bit about how long you've lived here and what kinds of things you've been involved in in the local area."

Facebook questions:

There were two rounds of questions posted from the account of *Design for Sharing* (defined as a community interest page) to the combined 'friendship' group of the two researchers.
1) 'We would really like to know what, why and how you share.'
2) (after lapse of 2 months, asked to all original respondents and more widely)
'…may we ask one more thing of you? How do you manage your sharing activities? Do you use any customised software such as websites (ouishare, ecomodo, landshare, etc?), any computer supported sharing systems (like zipcar's booking system), any more standard digital tools (email, Facebook, Twitter, texting) and/or any other communication and management systems (eg phoning, making lists on envelopes)?'

This gave us access to a group of digitally adept commentators who could answer as individuals. The group hardly overlapped with the local people of Brockley, spanning the world and bringing in commentators from Europe, southern Africa and Australia. We saw this as a way of getting beyond the approach of the kinds of project we were reviewing, to find out more about how people share, what they use and why. This would provide some contrasting understandings and practices. Obviously, this was skewed to our networks, but no more so than the use of Brockley and a local researcher who only walked 30 minutes to find her interviewees. This provides an online correlate to taking the radius of a researcher's walking journeys in Lewisham, providing an informed test group to discuss issues of sharing and technology online. We see this relative intimacy as a strength and note that our networks are well informed about social sustainability, collaborative consumption, etc.

Review of digital resources

All the material featured in this section, unless otherwise noted, is drawn from the featured sites between July 2013 and April 2014. For a glimpse of the digital world at that time, we refer readers to the Way Back Machine (http://archive.org/web/). Here again, the sample was reached through a study of available sites and services at the time of our study. We sought to catch a glimpse of the state of the 'sharing economy' online during this time, sampling for a good range of approaches, services and business models. We undertook what we might call an 'expert review'[3], where we looked at both the existing literature and a number of initiatives to provide us with our wide-scale context. From this we drew out some of the common factors in 'shared economy' projects and organisations and how different types of sharing affect and are affected by these factors: for example, trust, loss, commercialisation, collaborative consumption and maximisation of resources to avoid waste. There is always a danger of circularity when choosing resources because they are in the sharing economy and then using them to define behaviour of sharing economy services. We guard against this by corroborating their status using media reports of the time.

Approach to Analysis

In each of these contexts, we examined practice, organisation and the types of facilitating tools used – whether these were bespoke digital, 'every-day' digital or analogue. We looked at place and locatedness, purpose, definitions of sharing and founding narratives. In the case studies, where we present local change-makers' work in establishing the service or infrastructure, we call this 'set-up'. In considering digital practice and tools, we looked wider than tools explicitly for sharing to consider sharing via tool use.

[3] The term 'expert review' derives from the practice of conducting analysis of a website by assessing its features and how they will function in a particular context, a core part of usability, design and marketing work.

Brockley Case Studies

The studies

Case studies are used to frame the seven interviews collected as part of the research. These take sections from the transcribed interviews, combine this with locally available information and researcher observations and present the initiative featured. Each case is given a theme:
One small thing - microlibrary
Mobilisation – community-asset pub
Growing gains – community garden
Rootedness – digital whip-round service
Networks in the making – communal shed
Scale and tempo – time bank
Stepping into the weave – conservation society
and all of them follow a schema: set-up; organisation; financial model; digital technology; sharing; special theme.

We have left the names of local places in the text of the case studies, only clarifying the social significance where relevant. Interviewees are referred to by their (real) first names and all the accounts have been checked back with them and are approved for public distribution in this form.

The area

We chose the small area of London in the London Borough of Lewisham known as Brockley for a number of reasons. One of the authors lives nearby and has good connections in the area, which facilitated a short study. The area has a social and economic mix: in one street, residents include students, private rental, owner-occupied houses and flats, social housing and a couple of extremely wealthy individuals. There is a mix of age, occupation and income. There are families of all ages, couples, houses of multiple occupation and lone occupants.

There is a lot of activity in the area: an annual fair and annual arts festival, a thriving weekend market and a wide range of voluntary activities organised through diverse civil society groups. This includes a number of environmentally-focused groups and the area is notable for having elected the only Green (ie only non-labour) councillor in the borough of Lewisham. The area has a thriving hyperlocal news site (frequently mentioned as a model example of such sites) which was useful in finding out about what is going on locally. And there is an interesting range and variety of sharing practices and related activities, both recently initiated and longstanding.

The following seven personal accounts come from people who have chosen to act on their environment and for their neighbours. They have been honest about their challenges. We present these accounts to show what can be done, but also why these individuals acted and, indeed, what support might be needed to encourage others to act, by showing vulnerabilities as well as successes.

Our last account does not feature a single change-maker, but the fabric of the locale. It scales outwards to consider how far the seven accounts refer to other work in the neighbourhood to make sense of their own. In doing so, it highlights some of the relations that make these initiatives sustainable as part of a local culture, rather than as self-standing activities.

One small thing…. the microlibrary

The Microlibrary sits on a busy main road in a former BT phone box, lit and full of books. Every phone box is a Grade 2 listed building and so Sebastian's first desire to paint it sparkly gold was thwarted by planning restrictions. It is a good space for a single browser: a K2, of which only 1500 were ever made, before size and costs were cut. There are books in Polish as well as English and none of the original books remain, regularly taken out, traded in and replaced by local residents. Apart from a short note explaining what to do, the microlibrary just stands there.

Set-up

Sebastian explains: 'There was a sign in the window saying that, as part of the kiosk exchange scheme, does any local resident want to turn it into a library or a toy store, something like that. And the telephone box is looking a bit sort of sorry for itself - the windows were broken - so I thought okay, I would give it a go and it took quite a while to track down the people that actually left the thing in the window.

'It's a sort of project that I thought I could do well, because it's not an ongoing project with no boundaries. I have a tendency to turn everything into a grand scale project, but I thought, with this, how can I do that? That's what appealed to me; the fact that it's limited, that I'm not getting involved in an endless additional work thing.'

Through the Brockley Central blog, Sebastian found Clare at Brockley Society and then negotiated permission. British Telecom will sell the kiosks for £1, but only to local charities, community groups and local authorities, not to individuals, because they are worth so much.

He visited a similar venture near his mother in Tisbury, out in the countryside. This one is overflowing with books. Dubiously, he started ripping out the guts and working on shelving. 'The scariest bit was the first day when I've got the ladder over my shoulder and walked up the hill. That's when I thought to myself "Oh, what have I got myself involved with? Why am I doing this? Given myself another project to do…" I'm always complaining about not having enough time.'

Almost immediately, some wood was stolen from outside the box, boding badly for an unattended facility. Sebastian resolves to do all the creative work at home and install the library in one go: 'I thought it's almost like a social experiment as much as anything else. If that's happened to the bits of wood, what's going to happen when I put books in here? It was quite a disheartening moment. I was very nervous about it. I mean my wife's had to put up with a load of grief. I've been sawing in the kitchen - because I didn't want to move things up there incrementally and have bits of brackets there which will be stolen. I wanted it to go from nothing to finished overnight.'

Organisation

No formal organisation supports the microlibrary; none seems needed in the short-term. The simple rules of borrowing or exchange are largely adhered to, but not policed, and bulk deposits balance any gradual erosion. Notably, it has survived in good shape as a communal resource, despite initial misgivings, and although Sebastian has now moved out of the area, other local people tidy it and do the minimal maintenance.

'I was there painting it and the people that walked past were like something out of *Breaking Bad* and I thought why am I doing this stupid thing? It's going to get trashed. The fact that it hasn't been trashed makes me feel very happy, that it shows that we're not horrible people. When everyone decries how society is becoming worse and worse, you think, well there's all those books that people could take and if someone wanted to they could go to it and take all of them and no one has. Not one single person out of all these supposed ne'er do well's that walk past…'

Financial model

It has cost Sebastian about £500-600 of personal money for materials and the loss of a week's work. And the electricity supply is still a mystery, so that may, in time, come to cost a local organisation.

He also gave thought to the value of the contents of the library in setting it up. 'I had a friend from work saying "I've made some books; they're handmade books. Do you want them?". I said "No that's too good. I don't know how it's going to work; the books could get thrown over the pavement; people could sort of smash it up, you know." I wouldn't want anyone to give me anything too good. It would just be horrible. And so I said don't give me any first editions or, you know, really valuable books.'

He also dismissed stocking CDs or DVDs after a friend pointed out they have a significant resale value. 'It has to work on a totally non-money basis. Someone said to me "Would it be okay to sell some of the better books on Ebay or Amazon to get money to buy more books?" and I said "No, the only currency it works in is books, not money." If people thought that the books they gave were sold, then people might think that someone's making a profit out of it and that would destroy the goodwill, which is worth more than anything. It's obvious when anyone walks past that no one's making any money out of it and that is the crucial thing.'

Digital technology

Documentation was an integral part of making the library. 'My wife has set up this Facebook thing that's getting about 8 new 'likes' every day. It's got about 260 'likes' [at time of interview] and it just makes people smile, which is the most wonderful thing really, yeah. I took a picture of every stage, so on the Facebook thing, the pictures are all arranged in chronological order and you can see as various things were cut away and added … you can see where I've chopped off the communications bit but I've left the electric.'

Sharing

As noted, Sebastian sees the money-free model operating in the library as encouraging the goodwill that protects it: 'when you give a book you're making yourself a bit worse off because you're giving away some of your possessions, but the overall cumulative thing is that everyone becomes enriched by it.'

One small thing…

'It's a very constrained little thing. It's not like an allotment that will always be there; you always need to do something to it. I thought: it's very physically contained, it's just a box …I've made music and spent years making albums that no one's ever listened to but, with that, I thought… maybe that's it…doing something that the public respond to. If I do something like this, the public respond to it…' He considered a range of enhancements, such as music and a message about his purpose in setting up the library, rejecting these as destroying the simplicity and broad appeal. He chose something containable. 'It's made me an incredible train spotter with phone boxes…'

Mobilisation … The Ivy House Community Pub

The Ivy House pub serves Nunhead, near Brockley. It is an interesting venue: it survived the Second World War when the rest of the street was bombed; it has a musical history, presenting artists like Elvis Costello and Jeff Beck over the years; and it now carries a Grade 2 listing for its architecture. It also was the first pub to be listed as an Asset of Community Value and the first building in the UK to be bought for the community under the provisions of the Localism Act, invoked so shortly after the Act came into force that the application had to be written on borrowed forms. Tessa helped lead the purchase and sits on the management committee.

Set-up

The Ivy House, as was, closed in April 2012. Tessa recalls: 'The tenants at the time were given a week's notice. I live around the corner and this was very much my local - I had my wedding party here a few years ago - so I knew the building well and I popped in on my way home from work to try and find out what was happening and there were about 80 people also trying to find out. My husband called me at work and said I've heard there's a rumour that the Ivy House is closing in a week and I popped in. There was no meeting or anything.

'Nick, the bar manager, was so sick of all the local residents saying "What's going on? What's going on?" that he started introducing us to each other and a group of us ended up sat there chatting about what we could do to save it. That group of people who were sat there became the steering group of the campaign and are now the management committee. We didn't know each other before or anything. I think we just got very lucky that around the table there was a really good mixture of skills that proved to be really useful. That's basically how the initial campaign group came together. It was purely by chance really.

'At that time, the Campaign for Real Ale had already submitted an application to have the building Grade 2 listed with English Heritage, so we lobbied our local councillors and our MP to get that expedited and pushed through before the building closed, to protect the interior, and it became Grade 2 listed two days before it closed. Then, after it closed, we started thinking about what it could offer the local community, what kind of pub it could be if it was in community ownership, but at that time the relevant parts of the Localism Act weren't in force so there wasn't actually anything else we could do. But we started writing a business plan and we were still talking to other pub operators - if anyone else had fancied to take it on and run it as a really good local pub, we would have supported them. We just thought it's what the neighbourhood needs and it would be a real shame if it disappeared.'

Things changed in the autumn, when the Localism Act came into force and, with help from the Peckham Society (because at the time the group was constituted as a private limited company and could not use the relevant bits of the legislation), they were able to start raising the money needed to transfer the property into community hands. Two major funders pitched in, yielding £1M to buy the property. Next, the group launched a share option locally and raised another £142,000 in community shares from 371 shareholders. 'We had a target of £100K and so we were really pleased. There was a lot of interest and support for the property and, I think, generally around here there's a lot of interest in community ownership of things.' Most people who bought shares were individuals from the area, though some businesses, including breweries that support independents, and some more remote supporters also bought in.

Organisation

Tessa is now part of a committee that oversees the management of the pub, which opened again in August 2013. She makes the point that it is co-operatively owned, but not co-operatively run. 'It's not like shareholders work behind the bar or anything like that. It's run in the same way any other pub will be run, by a manager. He's got an assistant manager; he's got a head chef; and, obviously, all the bar staff as well, and he does all the day-to-day running of the pub, although we are still much more involved than I thought we would be at this stage. We are slowly able to ease off a bit, but there's been, inevitably, lots of teething problems with the building and all that kind of thing to sort out.

'One of the reasons why we have been quite heavily involved is that we felt we sold the pub as a project to people who bought shares in it. …So we wanted to remain involved in the initial set-up, to establish the framework and ethos - things like London suppliers. It's not an exclusive thing but we have tried to use local suppliers wherever we can. We've had a lot of support from other local businesses so it feels quite natural to support them in return. Among our share-holders are some London breweries. It's also just easier in many ways because if you're using London suppliers, if you need another keg of something, then some London breweries will just put it in a taxi and send it to you...it buys you added flexibility.'

The committee have a big commitment and would welcome another couple of members to spread the load. 'Simply because there still is quite a lot to do, but it's actually having the time to stand back and find those people and carve out defined roles for them so that they're coming in and they know what they're doing. It does worry me, because I can't keep doing this forever, not at this level anyway. So it's about developing strategies that allow us to step back.'

Financial model

The owners are a community benefit society, which has limited liability so nobody is personally liable for the debts of the organisation. Nonetheless, running at a profit is part of the plan. 'We're a not-for-profit organisation but you've still got to make a profit. It's just that we have to use our profit in a different way; it's not okay to run something at a loss. A lot of people ask if we're going to pay dividends and the answer is "possibly".' The society is legally able to pay returns of up to 5% above base rate. 'But there has to be the money there to do it. And also, most people invested not expecting a financial return particularly, but because they wanted to be part of the whole project to keep it open.'

'In terms of keeping it going and keeping people coming in, I think that we just have to make it a really good pub and, if it's a really good pub, people will come and use it. …because it's community-owned, but also because it's doing a good job. I think if it was community-owned and it was rubbish, people would support us, but not for...they'd give you, like, a few months, but not indefinitely. We're slowly figuring out what people around here want from their pub and it's slowly developing into that and that's why people are coming back.'

Staff training is taken seriously and they are paid a 'London living wage', higher than normal bar work. 'We thought what is going to set this place apart in terms of being a really, really good pub is the quality of the products, but also the staff and the standard of service you get when you come in, so we're quite happy to invest in that. But we do expect quite a lot in return.'

Raising the money was not as difficult as the team expected, though with Christmas 2012 in the midst of their campaign, the whole thing was managed in only 4 and ½ months, instead of the full six. The moratorium provided by the Localism Act didn't feel long. But, principally, Tessa found the campaign intense because she was also juggling a busy job and a 3-year-old son.

Digital technology

The pub uses the usual accounting software and so on. 'We don't do anything particularly innovative with technology. Social media has obviously been really important, particularly during the share issue, when we were trying to promote that. We're on a Perfect Pint app, which is a real ale app. …We do what any other business would do really in terms of using technology.'

Sharing

'We've had a lot of helpful volunteers - all kinds of areas and particularly with things like sign writing, just bits and pieces …leafleting. Is that donated or is it sharing? I suppose it's sharing because it's the community-owned building, so you're sharing your skills and expertise with the community, which is, I guess, essentially what the management committee does as well. I guess it's about exchange, isn't it? It's about exchange for things other than cash.'

Mobilisation

Taking over the pub required persistence and expert co-ordination from its committed local team, though from the meeting of strangers in the pub a week before closure, to the last-minute English Heritage listing, to the hasty application of the Localism Act, the journey to ownership also reveals the role of luck. 'You go into it not knowing if all your efforts are going to pay off or not, but I would rather do it and it come to nothing than be sitting around in 5 years' time thinking "Oh, I wish I'd tried …blah blah blah". It's all very well sitting around complaining about things, but if you're not prepared to try and do something then... I've never done anything like this before but I just thought there was that one chance to try and save it and... just have a go.

It also required a responsive council: 'It was great that Southwark were able to process the application so quickly… it was days as opposed to weeks. It was maybe 10 days or something, because it was so urgent and the risk of it being sold at auction and the interior just being ripped out… So we asked them to expedite it, which they did - it was great.'

As to why a group of strangers were able to make this happen: 'I think it was the sense of anger that a lot of people felt over having no control over the development of the area in which they live. It's a really lovely building and we didn't want it to be lost. It would have left a massive physical gap in the neighbourhood.'

The pub now reflects the effort that went into saving it. 'It's just much, much busier. That's been the nice thing about the pub since it's reopened, our customer base is so broad. When we had our public meeting in April [2013] to publicise the share issue and explain to the community what our plans were for the building, the nice thing about that meeting was seeing the massive range of people and age ranges as well. And you see that in the pub too - one of its great strengths is it doesn't cater to any particular demographic and it's very representative of the local area. It's really, really mixed.

'We advise lots of other groups. It's very difficult when people just say "We want to save our pub. What shall we do?" I guess the first thing you've got to do is figure out if it's worth saving and if the community is behind it and if there's enough interest in it to make it worth all the time and effort, because, if the community isn't behind it, it sounds brutal, but there's not much point.'

Growing Gains… The Breakspears Mews Community Garden

Breakspears Mews now hosts a community garden. The big houses of the conservation area look over it on one side, while on the other are council flats: a meeting of different communities. Till recently, it was a run-down fly-tipping area where dodgy car repair businesses plied their trade, breaching environmental health laws and creating a haven for other forms of destructive behaviour. Jane led the transition from burning tyres and paint fumes to the flourishing vegetables of the raised beds.

Set-up

Jane has lived overlooking the mews for 20 years. 'It became a terrible problem with fly tipping, noise and comings and goings all night. It had become a bane...a thorn in everybody's life. Along with several neighbours, I complained to the council and sent in photographic evidence - to the planning enforcement initially and then environmental health got involved to a degree. Speaking to one of the planning officers, he said it would be really nice if it was a community garden, so I thought it's not just me who thinks like that. It would be really nice to have a garden down the mews, everything demolished and just a quiet garden everybody could use.

When another local resident equally disturbed by the activities formed the Breakspears Mews Action Group, Jane was able to help by sharing her records and previous correspondence of the problems. At the Brockley Society, a sub-committee came forward to survey those affected and discuss the mews' future. The years passed. The garages were demolished and the rubbish cleared, and cleared again. Jane recalls that it was cleared any number of times. In March 2011, an open day at the site elicited growing interest in the idea of a garden. 'That's what gave the push and so we formed a little sub-committee again of immediate locals to see whether we could get a lease.' £1000 from a local fund paid for the planning permission to change use from garages to community garden left some money to buy materials. But there remained the issue of liability: 'If anybody fly-tipped up there in the future, the responsibility was on us to clear it and we didn't have control of the gates or the landlords out there with the unreasonable tenants who were tipping oil and car parts everywhere.' Eventually, Brockley Society insurance was extended to cover that eventuality.

Additional funding from the Council was obtained and the lease was signed in March 2013, ready for the growing season. Jane borrowed facilities at the set-building company where she works to prepare wood for the beds, so raised beds were ready in kit form. Lorries of soil were booked to arrive as the council finished clearing the rubbish so that, as Jane puts it, 'the garage people couldn't park vehicles there and block the access'. In fact, it poured with rain all day and the site was still being cleared well after the soil turned up, but, by the next day, it was ready for the volunteers who were giving over their first weekend to line beds with geo-textiles and the back breaking work of filling with shingle for drainage and the soil.

Jane was selective about what volunteers were asked to do. 'You realise that people have limitations of experience and I'm not used to that because I've always worked with practical people, so it was a big learning curve for me. But that is partly also why I did get the majority of the structure pre-built. If it was novices with hand tools, this could go on for a year just making the beds up - you'd lose people.' Instead, it was done quickly and by the open day, the beds were finished, planted and glorious.

Organisation

There is a committee which brings together hands-on people with some who take a more background role. At an AGM shortly after inauguration, the vote was for miniplots as well as communal beds. About a meter square, the idea was that people could try their hand. More experienced gardeners opposed this on practical grounds and a few pulled out at this point.

An immediate problem was that not everyone sprang into action: some plots remaining empty for six weeks or so when the growing season was at its best. Others came with impractical plans. Some let their patches get out of shape. 'This is why some of us wanted to do it communally, so you'd have regular sessions of sowing, weeding, watering, feeding, cropping. You'd just go through the season and learn about a wider range of vegetables; you'd know what they look like as seedlings, how to thin out, what to prune when and you'd recognise weeds.'

The committee had agreed that people with plots must come every seven weeks. 'Some didn't come that often and their beds would get weedy and, of course, several of us just weeded these beds, because what we didn't want was to look a total mess for the ribbon cutting open day!'

After the first year, the next AGM voted to move to communal beds throughout the garden. This also caused controversy. Meanwhile, Jane is still working at the spread of people who use the garden. She and two others are the current backbone and they are looking at how open days are planned to make them as inclusive as possible. The garden is now opening two evenings a week, occasionally Saturdays, but generally Sunday morning or afternoon. 'We are still not attracting many growers from the council flats which is shame as I think some of them will have knowledge of growing other crops that we don't know. Some of them will have experience of gardening and we're not getting that knowledge coming in.'

Digital Technology

Jane identifies communication as an issue for the community gardeners. She finds email slow as a way of keeping people informed on what needs doing when. She doesn't use Twitter, but recognises its potential, also seeing the merit of group texting. There remains a tension between getting on with the job and communicating about it: 'Like last summer, you'd be out there till it was dark watering. You couldn't then come in and communicate to people because you needed to eat and go to bed. I've never been involved in something where, in a way, communication is actually paramount to how successful it is. I don't really know enough about speedy communication.' This year the pressure has been relieved by one of the other key gardeners sending out regular emails to the mailing list advertising the opening times for the forthcoming week, and since sessions are now communal, jobs are getting done because there's always a 'session' leader to direct peoples' energy.

Sharing

'Sharing is hard. Some people are very giving and others are quite selfish. I've been both pleasantly surprised and occasionally shocked. Sometimes I am disappointed that those who come along to help with the heavy work don't reappear for the finer planting and reaping the rewards but then I feel they must enjoy the physical work on the construction side rather than testing their green fingers. Overall after experiencing two summers I would say that sharing works and you do come away after an open session with a feeling of satisfaction.'

Growing gains

Speaking of the transformation the team has achieved from no-go area to flourishing garden, Jane says 'I think you just become intolerant of things. You just think: actually I don't have to put up with this; we could do something about it. And then you've got to find the right people to do things with. It's taking on the challenge and sharing the responsibilities.'

The work of the gardeners is for the long haul. Just as it took years to convince the council and sort out the terms of occupancy, so the garden is going to take some time to bed in and attract users from all parts of the community. 'There is much to learn on all sides - none of those currently involved has any previous experience of working in community ventures. I think ultimately it comes down to a few people having a driving vision and doing an awful lot of work.'

Rootedness… Patchwork Present

Patchwork Present is, according to Olivia's grandmother, a whip-round. It is an online business based in a Brockley shop, which supports groups of people buying a single collective present. An item, such as a bicycle, or a series of elements, like those making up a honeymoon, are divided into small, manageably-priced bundles, shown in a patchwork image on the website, under the slogan: "Get friends and family together to fund one gift that's really wanted – piece by piece". Then each contributor pays for a piece and the site levies a fraction of the cost – a form of crowdfunding.

Set-up

After some years working in marketing and another stretch at an environmental charity, the idea for the service came to Olivia as she planned her own honeymoon. 'I had a real issue with asking for money, partly because it just feels sort of rightly taboo that, if you are celebrating love, it should be about love and meaning - and gift giving should have some element of ritual to it and it shouldn't just be a financial transaction. So I just created this website, effectively created a shop for myself, selling my honeymoon. What made it real was the imagery that I chose, that I'd already done a bit of research to really understand where we wanted to go and then given each experience a rough price. Our friends and family could come to the site and choose which piece of the honeymoon they wanted to buy. So they could choose a physical thing like two beers in a bar or a night in a hotel or a donkey ride or a diving trip.

'What also led me to get so excited about starting Patchwork Present was I could start a business that I totally believe in, that is really, really useful to people and it doesn't sell people anything. It's a site than enables you to collect money globally and spend it locally. The roots of the business are quite personal, but the principles are about being resourceful, using our money wisely to invest in things that are wanted and needed and not buying each other a ton of crap that we don't need, that ends up in landfill. I never set out thinking "Oh, I'm going to be a business woman!" really. I just had this idea and it sort of seemed to work so well for us as a small family and the friends I have here in my community that it felt like it would be replicable.'

Olivia first looked for an office in Shoreditch (London's *Silicon Roundabout*) 'to get a ridiculously overpriced studio, tiny warehouse horrible unit for a ton of money' when the business outgrew her kitchen table, but settled on a local space. 'This shop had been empty for 7 months and I kept walking past and I just thought "Hold on a minute, we should just have a shop. Why shouldn't we have a shop?" And actually that made it cheaper in terms of the rent.' An added bonus is that Lewisham Council waives business rates for some categories of small business.

Organisation

There are three staff in the shop: Olivia specialising in marketing, a finance and a social media expert: 'But, obviously, because there's only three of us, we all clean the loo and we all answer emails and we all help people and we all just do whatever needs doing really.' Then there is support at the agency that built the site: 'Ideally we'd want to have an in-house team and we did actually start looking to recruit but, in the early days, it just seems to make much more sense, financially, to have a retainer with the agency who built it because you don't want somebody just sitting here kind of twiddling their thumbs and then paying them a lot of money for it.'

Financial model

Olivia didn't have the money to fund the project. 'It was nice because we raised the money according to our patchwork principles. The banks didn't want to lend me any money, especially as I said I wasn't going to risk my flat and put it up for equity. I have absolute faith and belief in my idea and myself and what it could be, but I'm not prepared to risk my children's security on it and when I went looking for investment I said exactly that to everybody who invested. I said "Don't give me any money you can't afford to lose". I ended up getting the investment from 25 individual investors who all have 1% of the business, which is great because also it means that I retain control. Lots of people have said finding money is the most difficult thing and clearly it is for most businesses, but actually it was one of the most enjoyable bits of the journey.' The company raised £250,000 in 2 and ½ weeks.

The site takes 3% of each contribution: 'If a friend contributes £10 towards a bottle of wine in a bar, we take 30p.' Olivia likens it to other crowdfunding platforms, such as Kickstarter and Just Giving, which charge more, or Taskrabbit with a 20% service levy. She points out that an advertising-supported model would not work, nor would affiliation to any retailers. 'So we'll see, because, actually, in terms of competition, our commission is quite low. That will be one of the things we will have to look at - meaning that the sharing economy only really works when you reach scale. So there won't really be any money to be made until we reach real scale, because getting 30p every day, you know...'

She explains her reasoning on setting this figure: 'Most of the decisions, I take from a user's point of view and just intuitively what feels right - like if you're sending a friend 100 quid and we take £3, it's kind of okay. If you start going "OK, let's take £10", it's like "Well, nice idea, but…." So it's just trying to get that balance between what's being paid for and the benefit of paying for a service that makes that financial transaction much more fun and interesting and real and personal but where you still get value, where you still get the money. So, we'll see. 3% is where it's at. What's quite nice is on the average wedding gift, it works out at the same price as a piece of wrapping paper or, on the average birthday present, it's like the price of a stamp.'

Digital technology

Development began in an ad-hoc way with friends, but it became clear that sporadic activity was not going to deliver as needed so Olivia moved to an agency, who later invested in the business. 'I didn't have the necessary project management skills to manage a very diverse group of people working a few hours here and there. I had absolutely no experience in tech whatsoever. In fact, you know, I'd always choose a pencil over a laptop any day.'

The result allows for customisation, both in choosing images, prices, etc and in then allowing you to send a personal "thank you" to the individuals who bought each piece. 'So you can send pictures of you drinking your beer on the beach to all the people who bought the beer and, you know, hotel night to all the people that bought the hotel.'

Olivia is impressed by the effort that people put in. There are ready-made patchworks for most occasions that anyone can use, but some people go far further. 'We've had a couple getting married who have not only made a patchwork themselves, they've staged all the things they want to do and then photographed themselves doing it and then uploaded them to the patchwork. There's certainly an effort involved in our site, but I think people enjoy it.'

Sharing

'People have always just had a whip-round in their own family and streets and communities and friends to get gifts or to help family members with the one thing that they really want, so that's not new. I suppose in a recession we do that more. We're forced to be much more resourceful, much more collective, much more collaborative - it's like this buzzword at the moment about "collaborative consumption". So this is just another platform that lets people help each other out, and help each other out around a specific thing, which is not skill swapping or any of the other things that you would share and trade and help, it's just "It's my son's birthday, you all love him, this is what he really wants, can you help me to get it for him?" or whatever.'

Rootedness

Unlike many online services, Olivia's company has a very specific presence in the locale in which she operates, including a shop front. 'We've built a site that we hope to be global, but you can only ever build a site for an audience that you know and understand and that audience is just me, my mates, all of us, all of our friends, our friends of friends, extended family, community and I just think if we can engage with people in Brockley and people in Brockley like and know and understand what we're doing, then that is replicable. You can't start off with a tiny small business and go "Right, hello world!". You don't have the resource to do that.'

'Then also there's the fact that it's got a shop window. You're talking hundreds of thousands of pounds a year to be able to put an advert on a billboard and we've got a window that basically acts as one every day with a logo above the door and an open door and a window that changes.

'But the other thing, I think, is that it's almost deliberately because we're online, we need to be on the street. Online businesses are so removed. I mean, I can't think of a business that I shop from online that I can call up and say hello, let alone go and knock on their door. You have contact forms - you don't have phone numbers, you don't have addresses and it's really hard to look somebody in the eye and say "I've got a problem" or "Can you help me?". So, clearly, if you live in New York, you're not going to pop around but you can call us and you can email us and if you do happen to live nearby and you have some issue with the site, then we're here and I just hope we can always do that, even when we grow, because it's important.'

Being open to the street has had its negative side. The shop was burgled and laptops stolen out of the hands of the staff. But it also has its compensations. 'When we had our laptops stolen from our very hands, I lost the fight. I'm really not much of a fighter. I did try, but I lost. So I said something on Twitter like "Tech start-up just had all our machines stolen, trying to build a start-up with a pencil and a piece of paper. Has anyone got any string?" And a woman came from Crofton Park with a ball of string. I mean, partly as a joke, but because I'd asked for it. She literally knocked on the door and she went "I've got some string". And I thought: sharing is a way of connecting and it's also a way of connecting without standing like a weirdo in a pub saying "I'd like to all be friends". You can do something and by doing something you make a connection and you all feel the better for it. There's the practical thing, but I think the emotional connection for community is really important.

'…So, definitely no proof of success to give you yet, but looking good.'

Networks in the making … Mensheds

Mensheds started in Australia, specifically for older men (http://www.mensheds.org.au/). They give a forum to address mental health and other wellbeing concerns, providing shared tools and support in the shape of carpentry and repair workshops. The sheds offer space to stand side-by-side and use craft skills, attempting to overcome the social isolation that haunts unemployed and retired men. In Australia, many suburbs have a shed. In Britain, they are a new undertaking, with a few in London and Cheshire, set up by Age UK. The shed developed locally is supported by the Sainsbury Trusts; one of three (Hoxton, Lewisham, Tower Hamlets) for which Alys has secured funding. Alys is also setting up a UK Men's Shed Association as part of the grant.

Set-up

Alys was working in an old people's day-care centre when she realised there was nothing for the men, so she pursued the idea till she arrived at Mensheds. 'My interest is older people and depression and that side of mental health. I think it makes people feel good if they're doing something for somebody else. You feel valued and the quality of your mental health goes up.'

At the time of interview, three sheds were all at different stages of readiness with the best support in the Lewisham area: 'I think maybe it's because I know the right places to go, because it's where I've lived for the majority of my life. Though I've worked in Hoxton and Tower Hamlets for a long time, I don't know where I'd look for information about the local area.' However, the Lewisham shed still needs work to meet health and safety and fire requirements and become a carpentry workshop. 'So I'm going down the route of it becoming a short-term reminiscence storytelling place, with a view to getting the group together before we build a stand-alone shed.'

Organisation

Alys is bringing many different interests together: funders, her employer, volunteers to help in the sheds, targeted users, the organisations feeding referrals to her and those with tools to donate. Referrals come from charities like Mind, from occupational therapists and people who have heard about it through the local council magazine. She has to consider sustainability - the funding runs out in late 2015. She has to ensure compliance with a range of protective legal frameworks, especially since some of those operating machinery can be classed as vulnerable adults. Officials are nervous because there are big tools involved.

She intends the UK Men's Shed Association to have a similar role to the Australian one: helping people deal with the paperwork and the admin side of it to make it more appealing to start sheds: 'In Australia, the Australian Men's Shed Association is big now. It's been going for many years. The [health and safety] recommendations are very much set out by the Australian Men's Shed Association. They've got this model and it's very easy to set up a shed there.'

Financial model

The sheds are not intended as factories where users make things to sell: 'Then it becomes like going back to work again and people don't enjoy it as much. They don't *have to* go along to the shed, they *can* go along.' She points to the failing health of many potential users: 'The *older* older men, they've got the experience, they've done it in the past, and they know how to do it -and they want to tell you the stories of how to do it, but they're scared of it. They want to, but they don't have the strength, which is why they can come and use the shed and not feel that they have to actually make things in the shed. There is a different dynamic with them.'

Meanwhile, another significant part of the shed concept is that tools are donated and she is working on getting their maintenance gifted. 'I don't know if that's because of where I've advertised, but we've had about 7 or 8 people donate masses of tools and paint and resources and that's from people that have maybe lost a partner that used them or they've become too old to use them themselves or they're downsizing and they need to get rid of a lot of things.'

The initial set-up grant was for £100,000 to cover staff salaries and set-up of the sheds. After that, the three sheds must fend for themselves, including rent and any paid help. There are other cost implications, such as getting people in to support people with support needs, training, etc. Even volunteers must be rationed as they need lunch. 'It's quite a big thing at the moment, men's health, so there are lots of pockets of funding that you can get money from. At the moment, it's not too difficult to keep going.'

Digital technology

With so many stakeholders, Alys is using a range of communication media. 'I obviously did a generic email out to all the local NHS and social services telling them about the project.' She says that volunteers and people who want to come to the shed have responded to news posted on the local blog, Brockley Central, but even more through the local council magazine.

Keeping interested people informed is problematic: 'I might send emails out but then you get the whole thing "Oh I can't open the attachment", "I don't know what I have to do" … For me, I find that really difficult because I'm from an age where just emails is a lot easier and a lot quicker - I'm only 2 days a week and I need to talk to 40 men. I appreciate that older men don't use Facebook or Twitter. I don't think it would be something that I'd use to attract people to each individual shed, because that doesn't seem to be where the men that I need to get through to - the isolated, the lonely men - would look. But people are starting to become aware of it and it's the future in terms of what's going on, so you need to be on board. I'm arguing about the importance of it, not necessarily to collect new members, but in terms of promotion.'

Alys also suggested that the Lewisham shed got a 3D printer, not least to encourage inter-generational work. It was considered too much a novelty item, but she is persisting. 'I've bought them all a laptop because my aim is for them, again in the long term, to develop the skills to be able to be able to sell products on Ebay if they wish, maybe set up their own blogs.'

Sharing

The kind of sharing that Alys has promoted so far has been of tools, resources and contacts. She describes a volunteer: 'He's a retired carpenter, he's worked in the building trade for a long time and he's got the skills to give, but he hasn't got a job at the moment and he'd like to get a job so he wants to share his skills and, in turn, use the CV.' Beyond this, Mensheds are intrinsically about sharing space and working with communal resources to promote a sense of wellbeing. 'When you're young, it's sharing toys or sharing your food and things like that. As you get older, it becomes less about objects and more about sharing your time, sharing resources that you have… just helping people really.'

Networks in the making

Alys has connected with the local health services and housing schemes to make a network for referrals in the long term, and she mentions the Brockley Central blog more than once as a lynchpin in building relationships on the ground. But the same feature that makes Mensheds work as a way of tackling loneliness also draws in volunteers: 'It's not your usual kind of volunteer work; it's a carpentry thing. People can use their skills and I think men and women like that concept compared to, say, maybe befriending… It's actually going to do something rather than having to go and talk to somebody.

'So, you might maybe sharpen tools or do the bicycle maintenance in Lewisham and then we can have the carpentry, so we can all share and use each other's shed in a different way. Then the aim is to open it up nationally, to have a shed in each kind of location…so that's the end goal, but we need to actually have a shed set up before I can do anything about that at the moment.'

Scale and rhythm… The Rushey Green Time Bank

The time bank operates alongside a GP practice. People give an hour of their time to someone and, in turn, can claim an hour from another person in the scheme. Philippe started working there in 2007, when there were 120 members, but the figure is far higher now. As chief executive, he is leading the shift towards a distributed model, with five hubs across Lewisham. The service has won awards for its work in community health and influenced the growth of others around the country.

Set-up

Inspired by Edgar Cahn's work on co-production, the practice that set up timebanking in Rushey Green saw it as a remedy for issues, such as motivation and esteem, that it could not treat easily: a kind of social prescribing. 'It was about GPs being aware that it's not always a physical element, it's often depression, loneliness… those who felt lonely for a variety of reasons - could be bereavement, could be job loss - and they don't necessarily need a medical treatment, but they need to be part of a community.' In 1998, the King's Fund supported a pilot and the first group of patients engaged in the scheme then convinced others of its merits.

Organisation

The purpose of timebanking is reciprocity. 'You earn time credits by doing something for someone else and then if you want something done by someone else or you want to do something with someone else, you use your time credits. Co-production is fundamental to what we do: you've got to do something, rather than being served.'

Philippe is part of the team who facilitate the exchanges, collecting the hours and totting up the statistics. They see the impact of the work: 'We're using timebanking to turn passive recipients of services into active people. I can't prove it and they don't realise it because they're involved in community, but actually their health hasn't deteriorated so fast. We can see, just by the spirit in the people and the way they behave after they've been there for a couple of years. You see that transformation, so we know something's happening.

'We run events that they want to go to and if they want to go, they've got to have credits. The thing that attracts people usually is outings. They want to go out and some people haven't done anything to earn credits. They'll dish out the money and say "I'll pay cash". I say "No, we're not taking your cash. That's bypassing the reciprocity." So they say "OK, then what do I do?" Then they can help at the surgery and earn time credits. Then what happens, their behaviour starts to change and they become more active.' Ideally, members are brought in as a group with several others, he says, so that they have to mix and so feel the benefit of socialising and working collaboratively.

At Rushey Green, Philippe has just introduced a new model where they are the hub and resource centre for many smaller hubs. 'The idea is to have little clusters of timebankers in little areas, part of the bigger family but very much local. We have *timebanking*, just not Rushey Green Time Bank and Crofton Park Time Bank… we have Lewisham Local, a group of hubs that will grow.' The smaller hubs are sustainable in the sense that people own them and run them, supported by the centre who trains them up. 'You need a critical mass of people for it to start to become vibrant.

'You have to put in some sort of rules and code of conduct so people abide by the principles of being good and honest to be part of that. And we control that here, because that has to be managed somewhere.' The centre can also coordinate applications for funding, so nearby units do not compete with each other. 'Funders are getting requests every day for the same thing. When it comes from a borough-wide organisation, it is simpler and attractive for the funder.'

Financial model

There are costs to timebanking: some time banks manage with just a book of contacts, but Rushey Green works to support people with mental health issues and so they match people carefully: 'We don't send people out to strangers...we go with them and introduce people and we solicit feedback to make sure that the relationship is good. We know our people; what they can and can't do.'

However, there is pressure to cut back: 'So you get the message "Get rid of all this", but you've got the other message from the statutory sector saying "You've got to manage your risk; you can't just send people out." So time banks like these need funding. We're not funded by the NHS. We need funding because the brokerage is very important and people can't work for free. Some time banks have closed because of the lack of funding. We're still here after all these years.' The new model of devolved groups, who can do their own facilitation after training, is a way of keeping the brokerage that Philippe sees as critical to success and allowing the service to grow, but keeping down costs.

'I don't advocate starting time banks as organisations; you can do timebanking within a group of people or something that already exists. Don't set up a new organisation because you need insurance, Companies House, Charity Commission, staff, and you're in for lots of money.'

Digital technology

The time bank has been introducing a digital platform called *Time and Talents*, allowing its users to assemble groups based on location and link to bigger networks. Members can then post requests for help and these can be shown on a map. The map can show exactly where members live, with full address, or it can give a general impression of where someone is based – a subtlety of design that Philippe welcomes for data protection and safeguarding members.

However, Philippe sees a drawback with using a digital tool that cannot be easily surmounted. 'Say that "Julie" needs some help on Saturday and it's not pleasant - the job's not pleasant. It's not, you know, do some shopping on Saturday. You put it on the digital thing. Okay, I'm there on my little phone, right, and I'm looking; there's a whole list. There's Julie, but there's also another job. Someone needs to go to Ikea. So, what I do? I turn off Julie and I select Ikea: click, done. I haven't even seen it, okay? Whereas, at the time bank here, we've heard about Julie… You're coming in the time bank and we say "Julie on the 7th floor, could you help her?" You're on the spot now. And the compassion and stuff starts to work. Because you're unable to switch off. You could well say "I can't do it" but in telling us "no" you're going to have to manage your body language...and actually that's a massive, massive difference.'

Sharing

'The value of sharing in the community? I think it is a trust thing. I don't think you share because you get something back out of it. When I first lend my drill, it's an act of kindness, I don't think it will come back broken. If it is broken, it means I have to make some decisions.' But Philippe sees the demands that individualised sharing can make: 'If you spend all your time sharing: …you're sharing great ideas, you know, transition towns, green stuff, things like that, but you're neglecting your children or partner, something is wrong. And I think that we need to talk about that. The way society is today, it's more work, less money available, so we definitely need to rethink our lifestyle and prepare the next generation to think about sharing or doing things differently - so I think there's going to be a resurgence of sharings: sharing the washing machine downstairs in a block of flats, sharing the central heating, like it used to be.'

Scale and rhythm

Philippe talks about economies of scale, but also the scale at which timebanking is meaningful: 'There was one person; I think it was in Ladywell… "I've got a car", she says. I say to her "It's too far". "I've got a car; I can come over here in my car." I was thinking: well, you can't park around here and people, here, they haven't got a car to go over to your place or to your area. If they want to reciprocate, it's not going to work. So she joined, but we never saw her. I knew this would happen. You've got to be very local in such a way that you bump into each other in the shops; where there's an event you can be there; when you need help it's around the corner - that's how it works. It has to be very local: you know more about what's going on in people's lives; you see them more often, stuff like that.'

As well as the dynamic this creates, he points to another factor that keeps the service at personal level: 'If you say "Oh it's great, I'm a new timebanker, I can get plumbing done for free or cheap" it starts to put a value, because the other plumber is £60 an hour. The angle we're looking for is not the cheap plumber; we're looking for the relationships people are building through the exchange. Part of the reason why time banks are tax exempt here is because there's an agreement that it is a tool; it has no cash value.

'We don't have plumbers; we have people who have skills and time to share and it's between them if they share between themselves. Where it doesn't work is when people ring us and say they want a service. As a member, even though you may be a driver, you may say "Can't do it next week". It's not a regular volunteer service where the volunteer turns up every Wednesday. So even if you are available on Wednesday with your car, you can say "no" because you're a member, you're not a volunteer. You say "Actually, I don't feel like doing it today." That's why we use "dip in dip out": you get involved when you want to and with who you want to.'

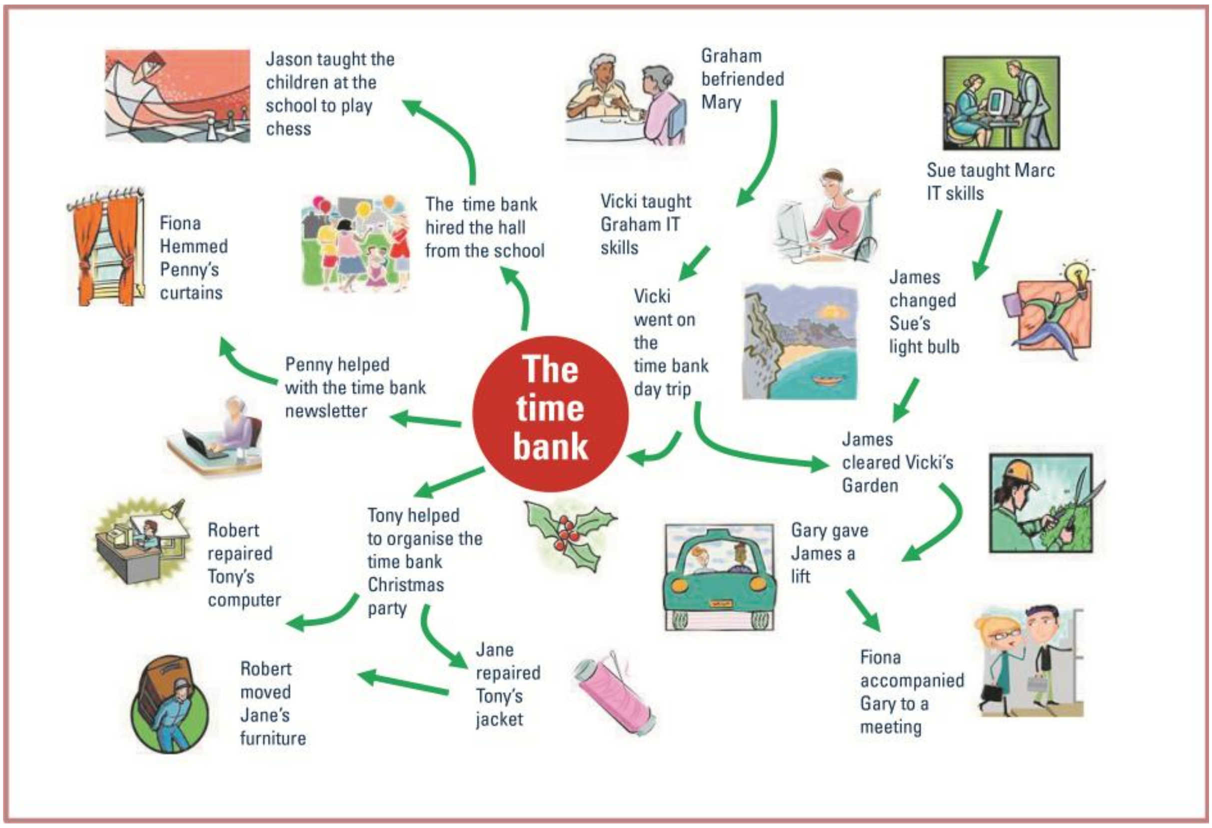


Figure 1: Illustration showing how the time bank works, courtesy of Rushey Green Time Bank

Stepping into the Weave – Brockley Society

Brockley Society is chaired by Clare, who took on the role in 2010. Its popular printed newsletter is delivered three times a year, free, to 4,000 houses in the conservation area. It has run a Midsummer Fayre since its inception in 1974 and supports other local ventures, such as the community garden in Breakspears Mews, the co-ordinated front garden sale in early autumn and a group of tree wardens. It was set up to represent the interests of people living in the Brockley Conservation Area and to monitor planning issues and now includes the surrounding areas.

Set-up

'We stand on the shoulders of people who started Brockley Society as the Conservation Society. It's what makes living in this area worthwhile, if you've got all this, you know.' Clare has lived in the area for nearly two decades and has helped with the society since 2004, as it evolved into the backbone of community-building in the area. She gives an example of how the society grows: 'We've got tree wardens, who suddenly appeared 3 years ago: one appeared at our AGM and said "I'd like to look at trees in Brockley" and that's taken off somehow.'

Organisation

The society runs the annual fair on Hilly Fields, the park central to the area. Along with planning issues and the newsletter, this is Clare's core work as chair. However, there were 116 stalls in 2013 and Clare sees a need to make changes. 'In a way the fair is becoming too big. There's a question which faces quite a lot of voluntary organisations: what happens when you get too big and you have to reorganise? At the moment, it's far too dependent on me and it's been dependent on the chair for a long time.' She points to the fact that she is also personal signatory on the lease for the community garden, with four others. The only area she has little responsibility for is the work of the tree wardens who are very active, but self-organising. Nonetheless, change is beginning. A new member with conference organisation experience helped plan the fair last time and produced outlines for team roles and made a pack up for each group of volunteers on the day, such as stewards, which can now be used annually to give each group a bit more independence.

Financial model

Brockley Society does not apply for funds. 'You can make money on a fair. And that's what we do. It tops up our income every year so that we can spend money.' The fair contributes about £6000 to the society's finances annually and the newsletter is self-financing through regular adverts.

Digital technology

In response to requests from the fair organisers of the time, the website now hosts the forms to sign up to it but not much else. Clare notes that the society has a complex website with nothing on it except the forms, but this is an effect of understaffing. 'I'm aware that we can appeal for things on there and we can put information on there. I do not want a Twitter account and I don't want a Facebook account myself, personally, because I have too many people contacting me. I get something like, dozens of emails every day and I've got 1200 people on my circular. So I don't want it, but I perfectly understand why it's a good thing. Somebody is doing the Twitter stuff – she's actually doing it from New York. She's gone to New York for 2 years. And Facebook…if somebody was going to do it, fine, and if somebody wants to do a map for next year's front garden sale, fine, but only if somebody volunteers, otherwise we'll just stick with a list.

'You can't let the technology determine what you do, but you do have to use it. There are lots of people, and certainly young people, who aren't very predominant in anything that we organise around here but if we ever want to reach them, of course, those things are really important.'

She notes that the Brockley Central blog is efficiently run and compensates for the limits of the society website.

And in the last year, the society has instituted a new free online accounting system. 'The fair bookings can be entered on this programme and you press the button and the invoice goes straight off to the people. It's automated the fair applications much more than we had done up till then and it means that next year, if the same person applies for a stall, their account is already there on the thing, you can just re-enter it and press the button, then off it goes.'

Sharing

Clare gives an example of what she means by sharing. The main tree warden observed to Clare that, in his road, there are people who have really good relations with their friends and family, but do not want to know when it comes to anything to do with the community: they don't care if there is a tree broken down in the street outside their door. 'I think that's true of quite a lot of people…they're friendly to you on the street when you meet them but their lives have got that sort of enclosure, whereas sharing is really saying that, even if you're just one new person moving into another area, you would like to know other people.'

Stepping into the weave

Brockley Society sits between many other organisations which work in the area. 'You try and keep an open link with all of them, even when you think they're doing things that are stupid or when some of the people around you think that.'

She cites many instances of cooperation. The society has shared gazebos with the Friends of Hilly Fields and borrowed a shed from Blackheath Conservatoire. She works closely with Brockley Cross Action Group, which campaigns around the Brockley Station area, looking at planning applications together. She is talking to the newly re-established St John's Society and sharing advice. Brockley Market is another local organisation with links. There is the Brockley Social Club that the society is working to help revive. And there are connections to St Andrew's Church, which is busy in the community and runs two clubs for old people. The local housing association has worked with the society on recent street parties. Voluntary Action Lewisham, which coordinates volunteering, sends out society notices with their newsletter, as does the Pensioners Forum in Lewisham. 'We have an old people's party in the social club in February called the "mid-winter warmer" and we try and get that promoted.'

Transition Towns works closely with the tree wardens. The wardens have developed a good relationship with the people in the council, Greenscene, who do the planting of street trees, and the tree officer in the planning department. 'There was quite a lot of tension there because there seemed to be too many trees being allowed to be cut down and we arranged a meeting to talk about what the constraints are and what they can do and can't do and our tree officers were there as well and that has really established good relationships with all of them.

'There's a general feeling that everything the council does is wrong … you can never get the council to do anything. But you find, when you get underneath the surface, that there are people who care about the job that they're doing …these are people who love the buildings, they know the buildings, they do their best, they're working under pressure, they're being cut like everybody else and you're not going to get any further by telling them they're wrong all the time. You're only going to get further by trying to make a relationship with them and then try and tease that out. The existence of Brockley Society, I think, is quite important because it's a long-standing fairly solid organisation.'

Nothing happens in isolation…

Evident in Clare's account of the Brockley Society's work are the interrelationships in the area, which is something she fosters as society chair. But they appear through all the interviews: sometimes taken for granted; sometimes given to explain why it is that Brockley is a place that the person speaking could take the action they describe.

It is not an effect of how the research was conducted that we see this web of association, as might happen if Clodagh (the principal researcher) had chosen to mix exclusively in a set of friends. Although there was an element of snowballing (asking for referrals from a current set of interviewees to widen the pool), it is clearly more endemic than that; the character of the place and the tendencies of its inhabitants reinforce each other. This richness is, thus, a reason for exploring the particular locale and the initiatives taken there. So, while it is not surprising that we should find this strong weave, this does not lessen the value in asking how it operates and what it provides.

Lewisham, the municipal area that includes Brockley, is recognised as a poor borough, with a share of emptying local centres and boarded-up shops. But some parts respond differently to others. Olivia compares Brockley to another part of the borough: 'There's the Brockley society and the Brockley transition group and there's lots of grass roots things that pop up…even the fairs and there's just tons of stuff going on and I just think people are quite active. It's a very active community. You can't fake communities, you can't like just *create* them. Like eco-wise, Brockley is probably pretty good. It's not just about people's earnings and people's incomes that make it relevant for Patchwork Present, it's about a sort of level of creativity that a community has.' (Olivia, Patchwork Present)

The interviewees are aware of how this social fabric is constructed, shown most clearly in another extract from Clodagh's interview with Olivia, where the idea of sharing as a pretext for social engagement, as well as precursor to it, is strongly alluded to:

C: I don't know if you've seen it on Brockley Central, but there's three women who are trying to make friends with people locally and they turn up places with a pink umbrella - I am going to email them – and they've been in my local a couple of times. It's just such a cute idea.

O: And that is interesting because they're doing it without anything to share other than their lovely selves, which is great, but not what usually happens. Usually it's: "Anyone interested in knitting?" or "Anyone can contribute some veg?" or "Anyone want to help with this?"

This is not the only conversation about others in the area. Almost all the conversations include some reference to other local projects, because of admiration, because there is a link or because the interview has wandered that way. People are aware of each other; no one is working in isolation. Borrowing goes on between projects and the effort to find materials, support and funds is often shared, even when occasionally the need for further people resources stretches capacity and patience. Not least, as Olivia suggests, the area has developed a character that includes such initiatives, so that no one is surprised when another grass-roots effort takes root locally, such as the farmers' market, new foraging walk, or shop converted to be a gallery and workshop space. But more importantly for those doing the initiating, the willingness of others to support enterprise of this kind, be it lending out Lewisham College sheds or the big barbecue owned by Brockley Society, is part of what makes such experiments possible. Not only are there many communal assets; there is the local *relational* asset of having a rich sharing culture to draw on and to use for support. Oldest of the projects featured here, and decisive in the development of this mood, is Brockley Society, which formed when the council threatened the houses that back onto Breakspears Mews with compulsory purchase of their long gardens to build another row of homes. However, that is not enough to explain why the society flourished and expanded its remit after the immediate threat was over. Sebastian (microlibrary) and Jane (community garden) both pay tribute to Brockley Society's role in helping them achieve their goals.

It is possible to watch cross-fertilisation of projects and ideas in action, as this exchange between researcher Clodagh and Alys (Mensheds) shows. Alys has been talking about what the people using the communal workshops which she is setting up might make and do:

C: I'm smiling because you're talking about repairing tools and all this week I've been trying to find somebody to sharpen gardening tools.
A: Oh really?
C: Yeah. Nobody around here.
A: Oh really? Not around...
C: Yeah, I don't – there's a real kind of...
A: Oh, that's good to know. So we need to get a...
C: Because that's probably near enough to go to. It's actually for [X]. He was asking because he thinks I know these things. All of his gardening tools are blunt…
A: That's not good. Well, if there's nothing around here, then that's...
C: And, yeah, I think there's lots of things like that, actually, that just need to exist.
A: We could do it. Yeah. That's good to know. Yeah, we'll start …once we're set up and running, we'll get a sharpener and we'll… yeah.

In a more formal interview, this might be regarded as a digression. But seen as two local residents considering the way that people help each other out, we can view it instead as an example of the spirit that informs Brockley's change-makers: collaborative solution-seeking. The tentative nature of the exchanges and the gradual forming of the idea reveal the navigations that accompany such serendipitous encounters and something of the tact that accompanies many of them.

What our case studies reveal…

The first thing that strikes us, as researchers, about the activities featured here is the sheer hard work, determination, ingenuity and luck involved. These are design processes of the kind that Light & Miskelly (2008) describe, where contingency and passion play a major part. Only the time bank and the society are long-standing and now in the hands of someone other than their founder, though both Philippe (timebanking p33) and Clare (local conservation society p36) are responsible for considerable initiative of their own in developing their role. From the manual dexterity of Sebastian's microlibrary (p21) to Olivia's inspiration for Patchwork Present (p28), from the persistence that brought Jane to launching the community garden (p26) to the rapid response team saving Tessa's local pub (p23), the people interviewed have all made significant contributions to their community. Their involvement was sometimes almost random, inspired by frustration, a chance glance in a phonebox window, being somewhere at the right time. Their methods were determined as each eventuality showed itself, so there were steep learning curves involved. Some of the outputs were a surprise at times. Alys (sheds p31) gives an almost poignant account of balancing time spent on the red tape needed to open a workshop with keeping several sets of people happy so that there will be members and volunteers when the sheds appear. Uniting all of the endeavours is a desire to build spaces – some virtual, like the crowdfunding service and the time bank; some physical like the garden, library, shed and pub – where people come together and collaborate. This goal of shared space for activity takes emphasis away from what one person can give another - or even reciprocal exchange - and puts the focus on what happens when people gather together to share resources.

Doing things digitally

As part of analysis, we divided the studies to highlight how each initiative started, its organisation, economics, tools and our interviewees' concerns related to each. We now ask how the initiatives have used digital technology at each point and what can we learn from this.

Set-up

The stories of setting up each resource vary. Only one is a digital resource in and of itself. Patchwork Present has required all the usual elements of website design, from a payment mechanism to a striking interface with appropriate functionality. Olivia's account stresses the search for the right arrangement that works for her business in a relationship where, if the site is not adequate or functional, she has no product, so she is tightly involved with the software firm supporting her.

This is the only story of major technology use for set-up, although Philippe is developing his devolved time bank model with support from the *Time and Talents* free software he mentions (see p48 for more detail). In using this, he is using a custom, but highly configurable, application.

If we look at the other four accounts of starting up, there are many references to organising partners and members through email. There is the importance of social media to the Ivy House share launch and we see Sebastian documenting his progress on Facebook so that others can follow his example.

Major challenges include the difference in technology usage patterns between the founders and others. Jane has now delegated emailing so that she can get on with running the garden. Clare is aware that the society's website needs work, but hopes for the right volunteer since her hands are full with the fair. Alys's world includes Twitter and 3D printing, but her target users may not have smart phones or a computer and certainly do not use Facebook, making it laborious to reach them.

Organisation

The library, in its simplicity, may have a digital following, but it does not need one. If locals, in passing, see that it is messy, they pop in and tidy it. Proximity is the determining factor, not use of the right communication mechanism.

The importance of proximity and the integration of tools into existing activities form the backbone of the interviews. Much of the talk is of face-to-face meetings and almost all the software in use is regarded as typical – such as pub management and accounting software – rather than purposed for the collaborative aspects of ongoing management. After requests for easy access to fair booking details, Brockley Society uses its website to host an interactive form and now has a database of stallholders. The Ivy House lists its music events and menu on its websites, as other pubs do. There is nothing truly bespoke about any of this except the time bank software.

Nonetheless, the teams are coping with several management and distribution challenges, not least since they involve bringing together different parties. The sheds and the time bank, particularly, involve coordination of many agencies in ongoing fashion; Alys coping with physical tools and shared rented space, as well as people and care agencies; Philippe running an office at the GP practice. Tessa of the community pub's committee tells the story of several waves of collaboration: first an intense phase to secure the pub involving different statutory bodies, then broad recruitment for a share issue (including customers and suppliers), before the current activity of coordinating a new service, still setting policy and recruiting staff. The tensions of a journey toward a good structure for collaborative activity are caught in Jane's account of the garden's AGMs where land is divided into small individual plots and then returned to a shared model over the first two years of existence.

Parting with something expendable is easier than sharing since it has no trust of care attached to it. And there are informal ways of managing that, such as putting it on the doorstep with a note saying 'Take Me'. By organising a front garden sale, coordinated across the whole ward and promoted on digital media, Brockley Society both sets a deadline on the domestic coordination needed to divest a household of unnecessary things and provides a bigger pool of people hoping to get a bargain.

Only one respondent attended to the potentially serious problems with sharing resources, noting that sharing undertaken on an individual level can have drawbacks, such as the loss of property or damage to valued possessions. He also stressed that most households are already in a sharing system that means *shared* resources and attention are in fact being redistributed to share with others - and that needs local negotiation. Communal systems of sharing, like joint ownership of resources across a block of dwellings, overcome these issues. However, the growing pains of the communal garden reveal that shared ownership can bring its own problems, especially with a maintenance-intensive activity. We can all imagine the situation where someone fills their bags with produce at the end of a growing season having done very little to deserve it. Short of deliberate rationing, this is something that the 'social punishing' of which Benkler talks (2004) has to control, over time, or the gardeners have to endure. There is potential for technological solutions here, but they would have to be very sensitively managed. Instead, we are reminded of Philippe's comments about how easy it is to evade awkward tasks when considering reciprocation mediated by a screen. Mediation changes the dynamics and, even face to face, sharing fairly can already be hard.

Financial Model

Economic sustainability is arrived at in different ways too. Some of the initiatives are not wholly independent of awards and funding. Others, like the microlibrary, need no resources to run except books left by local people, after an initial outlay from Sebastian's own pocket. Patchwork Present seeks to be commercially successful. Alys, setting up sheds, and Philippe, chief executive of Rushey Green Time Bank, are both drawing salaries to support their work, but can only do what they do if part of their time is spent fund-raising. Powering on with local energy, Brockley Society is self-funding, showing that a margin can be made if someone is there to organise it. However, there is a trade-off between raising funds and having time for anything else, especially in a voluntary capacity.

No one had anything to say about how digital technology was making them more sustainable economically except Philippe, who was using *Time and Talents* to help in his devolution plans.

Sharing

Apart from Patchwork Present, which behaves like a sharing economy site and was defined as such by its founder, no one referred to the sharing economy or its bespoke services and software at all. These seem to exist in a parallel universe that only Olivia connects to (though she clearly also relishes the Brockley universe and takes pains to be embedded in it). In fact, if we turn round the question and ask how the sharing economy relates to Brockley, Olivia becomes a very interesting entrepreneur, since she is present in a local shop as well as being global in intention. Her personal journey brings marketing experience and green and localist values together and she has a sensibility for neighbourhood that she shares with other interviewees, despite her more commercial agenda and crowd-funding model. We see something similar in the narratives on other sites (p48).

Digital Technology for supporting these practices

People are not using or aware of the digital developments in sharing sites. Since we interviewed them as change-makers, this is not particularly significant. Nonetheless, none of them were using the likes of Streetbank, Peerby or Ecomodo (for details, see p51) as a side activity or an ally.

As in supporting any activity, a finely-tuned digital tool is not to be dismissed out of hand. But in considering digital practice and tools, we need to look more widely than tools explicitly for sharing to see how sharing happens through general tool use. And we can ask if there are tools that might provide support for the less individualised forms of sharing that involve the set-up, organisation and sustainability of shared spaces, such as we are seeing in Brockley. We are looking at activities that exist at the second and third layer of Janelle Orsi's guide to coordination and infrastructure (2009) and we are seeing many ad-hoc ways of dealing with the responsibilities this brings.

Possibly there is an appetite for more systematic management of people and physical resources to be met with mechanisms for fair distribution. There are also other kinds of software, such as those which seek to interest people in collaborating in the first place (eg Heitlinger's watering can, which recounts stories, recipes and growing advice for herbs in a communal garden, 2014) and are not concerned with efficiency and management. These might support local enterprises in a more entertaining way.

What we need to ask carefully in looking at our interviewees' practices of making do is what will make the technology work for each context, activity and combination of actors? But, yet more importantly, we must ask the higher order question: what makes the sharing work? And technology may be helpful, irrelevant or hindering when we answer that question.

Wider themes in creating sharing systems

Local understandings of what constitutes sharing helped to bound our definition and approach to the study. We can see the extent of concern for the wider community in all the accounts. We can also see how intolerable conditions drove some activities, whereas merely a sense of personal challenge drove others. Almost everyone made some reference to how much they were able to learn about themselves and their neighbourhood as part of intervening and also to the relationships that were initiated – and reminds us the proposed merits of co-production and working in the core economy (see p13). All of this speaks to a different concept from pragmatic sharing of stuff because we are not using it, though a desire to consider both the local environment and the state of global resources is evident in many accounts. Tackling waste and overconsumption is handled through offering a positive alternative and speaks to concepts of civic duty, community spirit and so on, as well as efficiency. A strong theme is negotiating with the other systems in place, be that local authorities, housing associations and/or fellow voluntary groups. The specific drivers of this behaviour included:

- Giving something up in order to be rewarded in other, better ways
- Exchange for things other than cash
- Fixing something for the benefit of everyone

- Giving something back
- Experiencing the sense of having made a contribution
- Pooling time and expertise, skills and resources
- Being part of something bigger.

If we consider services which are specifically about promoting individualised sharing in the context of communal activities in Brockley, we notice something odd about them. Tools for sharing remind us, as a starting point, that we have assets that are lendable. Conversely, in everyday life, this awareness rises in chatting with a friend or neighbour or in spotting a request for help on social media. It is not usual behaviour to promote our assets without a context; the action could look like showing off our good fortune. The time bank asks that people consider assets ahead of their use, but this comes in the context of highly localised group of potential users who have existing contact with the medical services and is being used as a tool to support mental health. The community garden came about through the desire to do something positive with a no-go area… all the initiatives in Brockley can be seen to have an agenda that extends past promoting any simple behaviour; they all have a long-term structure to them. Much of the online activity of the sharing economy is divorced from these contexts that make a collaborative behaviour suddenly relevant and helpful.

Last, we sum up the seven characteristics that we pulled out of the studies:
One small thing - doing something that doesn't overstretch the capacity of the people involved
Mobilisation – taking a broad cross-section of people with you as needs dictate
Growing gains – giving the time it takes to build social systems around new resources
Rootedness – aligning with the place and culture you are operating within
Networks in the making – using existing networks to embed and form partnerships for delivery
Scale and tempo – linking interests at an individual and small group level with the wider scope of the project
Stepping into the weave – offering help to other organisations and so creating a culture of support.

The Facebook inquiry

25 people provided comments on the Facebook page 'Design for Sharing' set up in October 2013. 19 people responded to the original question regarding sharing practice: *'We would really like to know what, why and how you share.'* 8 people of the 19 responders provided answers to the follow up question about tools (Feb 2014) with further people joining at this stage. The second question specifically asked about custom sharing services: '*How do you manage your sharing activities? Do you use any customised software such as websites (ouishare, ecomodo, landshare, etc?), any computer supported sharing systems (like zipcar's booking system), any more standard digital tools (email, Facebook, Twitter, texting) and/or any other communication and management systems (eg phoning, making lists on envelopes)?'*

All respondents are linked to one or both of the two researchers behind this report, but form a tiny subset of our combined 'friends' and their friends, who could also see the page. They are not representative of any group, practice or nation. What they do give is a geographically-broad perspective (Dundee, Manchester, Newcastle, Bristol, Sheffield, London, South Africa, Australia, Netherlands, Denmark), offered by individuals who casually practice sharing and also use Facebook, showing a digital competence and/or interest that we do not assume in our interviews in Brockley.

What was shared and how

The box (Box 1) gives a taste of what people shared. If we look at the examples that follow (Box 2), it is evident that sharing happens in diverse ways, influenced by convenience, efficiency and cost-saving, but also intended as a means to fit in with the lives of others, enjoy others' company and achieve things together. It depends on stage of life (childcare), part of world (mielie plants vs courgettes), relative affluence (petrol vs dog sitting) and a host of cultural and circumstantial variables. It is not taken so much as a deliberate act to share than as a solution to a collective problem or an opportunity that happens to present itself.

We are also reminded of the difference between sharing because it is pleasant and efficient and sharing out of basic need. 'In my village we share anything we have, eg. When there is no sugar in my place I ask from my neighbours. We share just to show humanity.' Being able to borrow a cup of sugar, then, is seemingly a universal emblem of neighbourliness, far more than just a solution to running out of essentials.

Box 1: summary of answers to the 'what' of question 1:

Tools (inc Large ladder, wheelbarrow, chainsaw, bike pump,)	9
Time/Skills/Effort (expertise, design, computers, diy, gardening)	8
Food (often surplus or from shared gardening activities)	6
Gardening (either as a reciprocal arrangement or in a communal space)	4
Child care (ie nanny, babysitting, etc)	4
Child transport (shared school runs/for activities)	3
Pet care	2
Books	2
Unwanted stuff	2
Other [Bike, Washer/drier, Boat, Ocado delivery, Petrol, Lifts]	1

Box 2: Several answers to Question 1 about what people share and why, showing range

Q1: I grow mielies, spinach and tomatoes on my sidewalks for passers-by to take. There are 80+ types of edible and useful plants in my garden, which I share with my gardener and his family, teaching him about their uses. The eggs from my chickens get shared with my tenants, family, friends and staff. I teach people how to transform the waste and rubbish in their environments into craft items for own use and for potentially selling. I host open days at my craft studio for the members of my neighborhood to learn new skills, meet each-other and learn about permaculture. No doubt there are other ways in which to share, thanks for giving us this food for thought. (South Africa)

Q1: In my village we share anything we have, eg. When there is no sugar in my place I ask from my neighbours. We share just to show humanity...we call this''Ubuntu''.......''One in all and all for one'' (South Africa)

DFS 21/2/2014

Q1: I share washer+dryer with the two other flats in the house. We also share the garden and a lot of tools like the drill. (Denmark)

Q1: We mainly share time/skills and produce. The allotment often produces a glut of something, so we share whatever we can't use. This year, a neighbour did the same with courgettes, so Sarah baked courgette and chocolate cakes and gave one to the neighbour in return. We've shared our web/interaction design skills in the past, and I'm often asked by friends to help troubleshoot their computer problems. In return, one of our friends has helped out by providing manual (landscaping) skills. That sounds more like a reciprocal agreement than it is - it's far less formal than that. (Yorkshire)

Q1: Childcare. I have had an after school agreement in the past with a good friend, and also babysit for various friends under an informal reciprocal agreement. Plus - some amazing ideas here that I might try to implement (Yorkshire)

Q1: We share school runs - it saves time and petrol. We share garden stuff, like our apple press because - why not? It makes sense that loads of other people should use it when apples are in season. We share boat and equipment transport at the canoe club; it’s part of being a member of a community etc etc. Knowledge, skills and expertise freely given and taken in many ways. (Midlands)

Q1: We share driving kids to a regular weekly activity. Saves on petrol, time and allows you to get to know your children's friends. (Australia)

Q1: I've used Streetbank.com in the past, but sharing seems hard to do when everyone can afford most of what they need, or people don't trust each other to look after their stuff. Re sharing skills/time, it always seems to work out better when I'm doing paid freelance work; arrangements seem to evaporate when it's voluntary.

Q1: A chainsaw, with a friend, why - only occasional use so cost, how- whoever needs it goes and gets it from where it was last!

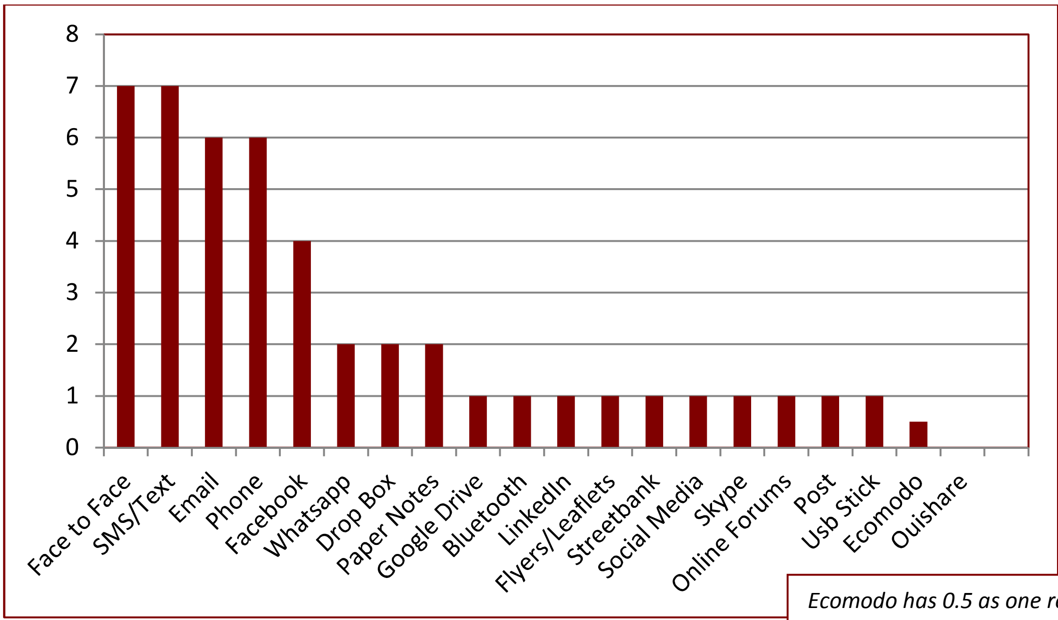
Q1: In the rural village, where i have more material things than others, people give me time and talent such as helping me with the car, looking after the dogs, hand washing when i am in a hurry, explaining, etc., i give things like food, petrol, lifts, busfares etc., Its not easy to answer why - its just the way life proceeds and it would be strange not to. Its done on as needed basis, no one thinks it strange to be asked to lend a hand; though people are very embarrassed indeed if they are in a situation when they have to ask for money, will put it off and ask only in private. (South Africa)

Q1: We share a nanny with some friends. To help her manage both of the kids we jointly bought a double baby buggy. Oh yes, and we share the nanny because it's half the price for each family if we do so! (and easier for her as the kids entertain each other). (London)

Use of Tools

The box below (Box 3) gives an overview of what tools were mentioned in accounts when people were specifically asked about how they share.

Box 3: summary to answers for question 2 - tools mentioned as used



Ecomodo has 0.5 as one respondent had signed up to it, but never found anything for which to use it.

Below we give pairings of answers that include both what is shared and how. It will be immediately obvious that despite the high levels of computer competence that our respondents possess, they do not use them much for social living. Email, face to face and phone feature extensively. Streetbank (unprompted) and Ecomodo (alluded to as an example) are the only cited custom tools and get one mention each.

Some respondents also stress the role of trust in their exchanges. Sharing is not parachuted in to make an area more acceptable to live in; rather it emerges from good relations and a lively interest in things going on locally and happens where networks are already established.

Q1: My neighbourhood [in Sheffield] has a mailing list where we post info of common interest or ask for suggestions/help. We use the list to borrow from each other (e.g. dehumidifier, long ladder etc.). We also share a shredder for garden waste and we buy (good) food as a group so we pay less. We have considered installing solar panels on the roofs of most houses but it is still too expensive even as many. We also share a "walking bus" to school with children trooped in pairs and only 3 parents to walk to school. And we leave plants or other things on the street for others to pick up: I left flowers and got a rosemary!

Q2: The highest technology is a mailing list. Most is done on paper, e.g. leaflets and street chat. However, a tool with an inventory of things (who can share a ladder, a shredder etc.) or suggestions (names of reliable plumbers, roofers etc.) would be good to have. A person in the road started a monthly Food Bank and he started with fliers through the letterbox.

Q1: We share dog care, child care, garden chores/tools/shed, plants, left over food and labour / extra hands / DIY with our neighbours on either side of our terraced house because life's easier and more pleasant when we help each other out. We do this as an when but at least once or twice a week (more in summer when we all live out in the back gardens which are all very similar because we landscaped them together a few years back.

Q2: It's mostly very local so almost always face to face (over the fence, in the street, down the local caff) or the Amii Stewart method (knock on wood ;)), now and again text. Do a fair bit of getting and giving digi resource shares via FB Linkedin etc (abandoned Twitter was a social media too far for me) but that's with my "distributed work network" - non work stuff, it's phone, text and f2f.

Q1: I try and pass stuff on rather than share - because I have a lot of stuff and it is too complicated to organise sharing it, etc. If it is something I don't use, then that's fine. I am loathe to lend stuff that I want to use myself as I forget to get it back. I lent a cordless drill to a friend/neighbour for about a year without realising.

Q2: My sharing is intermittent and I might write it down somewhere then forget where I put the piece of paper. My partner keeps a list of what he's lent out. That would be on paper too. The thing we lend most is probably books. (Bristol)

Q1: Some friends and I work cooperatively together to restore and cultivate a walled garden [in Yorkshire]. When the owner finalises the lease, we will share paying the rent. Meanwhile we have been working there 18 months, restoring it from a bramble covered wilderness to a productive kitchen garden. All the crops are shared in common. The first year was rent free. We do this for enjoyment, food, exercise and fellowship. It is very rewarding, and there is pleasure in sharing everything. Also when we have a surplus of crops we share them with our friends and acquaintances, because we do not want to waste them, and again it gives people pleasure.

Q2: We manage our sharing activities through discussion and trust. It seems to work. We text, phone and email each other.

Q1: We share babysitting - taking turns to look after each other's children. Why? It's free or nearly free (crisps and wine) and it's a person we know and trust, and who the children know and trust. Babysitting is a direct exchange - an evening for an evening, a daytime slot for a daytime slot. Arranged by text message, or in person with a family in our road. We also share Ocado deliveries with my parents, who are around the corner. One delivery charge, and only one delivery to wait in for. It's also useful to add just a few things to someone else's delivery when things run out between big shops.

Q2: Just SMS, email and phone. For babysitting we ask using one of these methods or face to face, plus normally confirm by SMS on the day or day before. For Ocado, if my mum or I book a delivery, we'd text or email the other with day and time. The person adding to the list will email a list of their groceries so it's easier to sort through the bags. No special software. I'm on Ecomodo but have never found what I need - would use it though. (London)

A Review of Networked Tools[4]

In this study, we have already seen *Patchwork Present* as an example of digital sharing – and Olivia identifies the site as part of the 'sharing economy' (p28). We have also been introduced to *Time and Talents* by Philippe at Rushey Green Time Bank, a tool that is neither specifically targeted at individual sharers, nor offers a fixed model of interaction, but allows organisers to customise what is seen by local people (p33). We can call networked tools that are made available to and configurable by organisers *Sharing Support Tools.* We look here at *Time and Talents* as an example, but there are quite a few other software platforms for timebanking, such as host to many Latin American pilots Cylos (http://www.cyclos.org) and Community Forge (http://communityforge.net), support for the Bangla Pesa (http://koru.or.ke/bangla) as well as many European community currency pilots.

Time and Talents: https://www.hourworld.org/index.htm

hOurworld, a US cooperative, has been distributing Time and Talents (T&T) web software for several years, based on the development work of timebanking members and using 'open innovation' to customise it for local conditions. It is available free. 'We implement the best ideas of our members and coordinators. This broad based co-production model of networking and development affords us all, the best tools for sustaining our local communities… We believe in supporting local diversity while maintaining a central cooperative organization to provide accessible support, training, and communications. We are documenting and teaching these models.' The result is highly configurable software to manage the particular service delivery model peculiar to each group using it and a support system to make it manageable and ensure learning is recycled. Bellotti et al (2014) analysed the structure of T&T, with other timebanking services, in their research into use of timebanking as a prelude to building a mobile interface with the organisation. They are also contemplating automatically detecting service activities and proactively proposing to smartphone users that they might want to record hours to help the time bank demonstrate its value (Bellotti et al 2013) among other experiments to see what contributes to the success of a timebank.

This tool demonstrates how one might support those promoting local sharing. Most tools receiving attention at the moment have a different ambition. We identify three additional paradigms: *Evangelists*, *Co-ordination Sites* and *Transport Sharing Tools*, exploring what they claim to be doing, how they do it and their understanding and use of place (and locatedness more generally).

Evangelists

Since we are looking at sharing practices, we pay especial attention to the organisations that have sprung up to promote these. We have noted that the terms collaborative consumption, collaborative economy and sharing economy are appearing in the media and also in the world of social and economic policy (eg https://www.gov.uk/government/news/innovation-in-giving-fund-makes-further-awards-to-boost-social-action). The following organisations are prominent in the promotion of a sharing economy. We look at how they present themselves, showing both the differences within the sector and some common tendencies that contrast with what we have seen in Brockley.

Ouishare: http://ouishare.net/en/about/story

Ouishare emerged out of a small group of Paris-based enthusiasts who shared regular pot luck suppers to discuss sharing and collaborative consumption. We use it as a benchmark, since it has been quite successful at mobilisation.

---

[4] Where no specific reference is given for a direct quotation, it comes from the site at time of review (Dec 2013-April 2014) – see p18 for more details of method.

It narrates its beginnings, from the Paris meet-ups onwards, as part of its marketing: 'From this collective energy, one core idea emerged: we wanted to better understand the change underway and accelerate the transition to a more meaningful society by empowering the pioneers of the collaborative economy from across the globe' and thus 'accelerate the shift toward a more collaborative economy'. The network is now international, supporting and promoting online and local activities related to this mission. It includes 'entrepreneurs and social innovators, non-profit and business leaders, grassroots activists and public officials, researchers, designers and more'. There is also a multilingual online magazine, an annual festival, a job board, and Ouikit - an experiment to develop resource-sharing systems within the OuiShare community and beyond. In terms of managing scale[5], Ouishare works through *Connectors* of different types: local (which here is city, region or country, unlike our study), global (communities of interest), knowledge and project.

Ouishare makes no distinction between collaborative consumption and 'the sharing economy', describing 'the seamless circulation of products and services among individuals through sharing, swapping, trading, renting, borrowing or giving, fostering access over ownership and reducing waste.' These activities sit alongside 'Open Knowledge, Open and Horizontal Governance, Crowdfunding and Person-to-Person Banking, Open Design and Manufacturing.' The language used draws on the world of digital development. Terms like 'permanent beta' are presented as concepts and an expressed value is MPRL (Meet People in Real Life) which suggests that the starting point for activity is online and distributed. Throughout, the site talks extensively of online activities and includes little of related practices that are not on digital platforms or involve co-working spaces. Central to this definition is access to and acceptance of rich technologies and flexible working styles. As with Shareable, below, the contributions on the site reflect a wider set of understandings and interests than the blurb for the site itself.

Shareable: www.shareable.net

Shareable describes itself as a 'nonprofit news, action and connection hub for the sharing transformation' with a mission to 'empower everyone to share for a more joyous, resilient, and equitable world'. Its tone differs from Ouishare in that it talks about movements and acknowledges political and social structures, varying widely in the content it features as a hub with many different authors. However, it does have explanatory pages which point to a particular view of the world and how it is changing: supporting 'a movement of movements emerging from the grassroots up to solve today's biggest challenges, which old, top-down institutions are failing to address', a move away from centralisation to distributed power and activity. Like Ouishare, it is international in scale and rallies its readers with strong rhetoric: 'New and resurgent solutions are democratizing how we produce, consume, govern, and solve social problems. The maker movement, collaborative consumption, the solidarity economy, open source software, transition towns, open government, and social enterprise are just a few of the movements showing a way forward based on sharing.' Sharing is at the heart of the project, as the name would suggest, and: 'The sharing transformation shows that it's possible to govern ourselves, build a green economy that serves everyone, and create meaningful lives together. It also shows that we can solve the world's biggest challenges - like poverty and global warming - by unleashing the power of collaboration.'

Collaborative consumption: www.collaborativeconsumption.com

The Collaborative Consumption website places a strong emphasis on network technologies as catalyst and game changer and, like Shareable, an emphasis on transformation on a global scale: 'the comprehensive online resource for collaborative consumption worldwide and network for the global community'. The site builds on the book *What's Mine is Yours* (2010) and other work by Rachel Botsman and offers a mix of curated content, directory features and blogs that draw together

[5] The Facebook page for London publicises the occasional event, or posting about free food (though on a London scale so not a very efficient way to deal with surplus) and co-working days which tend to present sharing as something for social entrepreneurs and freelancers.

information from a range of sources about different collaborative consumption endeavours and related issues. Occasionally there is material about the difficulty of definition in this area, such as a restatement of Botsman's thoughts on the overlaps between the collaborative and sharing economies (quoted on p9).

Peers: http://www.peers.org

The Peers network is a lobbying membership organisation which specifically addresses the needs of the sharing economy, intent on growing a new commercial market and solving the problems thrown up in its novelty, such as public liability when private individuals become hoteliers, taxi-drivers and chefs. The site argues that: '[t]he sharing economy is helping us pay the bills, work flexible hours, meet new people or spend more time with our families.'

The Evangelical Tendency

These sites all describe and align with new socio-economic structures. They focus on information sharing and supporting the nascent movements that they identify with, seeing an important transition afoot to new forms of social and economic engagement and speaking of the values associated with them. The Peers network goes further, taking a position on existing frameworks and campaigning for changes to laws and regulations. They have deliberately set their scale of operations as global as part of embracing networked change, though offering more local activities.

The potential of digital, networked and distributed technology is key to the development of the changes they see and to how each understands and promotes sharing and other aspects of a collaborative economy. All use quality of life issues to justify their work, though there is little concession to the different socio-economic and cultural contexts in which people live and seek to operate. Though a need for small-scale and local activities may be recognised, the world here tends to be presented as homogeneous and global.

Beyond these commonalities, we might see the language as heavy on optimistic rhetoric, citing cultures of openness and collaboration that are so far small-scale but attracting considerable interest. The sites themselves do not present critique, contextualise with any form of history or evidence the transitions they are keen to promote, though the content on Shareable goes far further into the philosophy of the new movements in individual contributions.

A further element of these sites is how they construct the anticipated user as a well-resourced and highly flexible individual and do not seem to notice the extent to which ownership predominates still – ownership of the tools which enable this kind of sharing as well as ownership of things to share. We are reminded of Stokes et al (2014)'s comments that the economy's stated traits of 'encouraging meaningful interactions and trust', and 'embracing openness, inclusivity and the commons' are sometimes more aspirational than evident.

The questions these sites raise are general in nature: How is the cultural dimension of people's experience handled in the collaborative economy? What provision is made for people for whom austerity is already a reality? What protections are lost in moving to a more 'open' culture of exchange? What might be learnt by situating these narratives of transformation in wider cultural trends and social movements of the past? How is technology expected to be transformative?

Coordination sites

Another distinct group of sites might be termed *co-ordination sites*, in that they enable peer-to-peer sharing by using computation to handle functions such as matching locations and interests, verifying identities, handling any financial aspects and so potentially connecting up people who live near each other but do not know each other. They mediate in setting up the meeting of people. There are a great many examples of different tools and sites of this kind. Here we focus on some high profile UK-active examples that show the range as well as the similarities in design and function.

Ecomodo: http://ecomodo.com

Ecomodo, 'the marketplace of good returns', is a UK-wide online brokering site that enables people with proximity to share - primarily through lending and borrowing, but also with the possibility of coordinating the sharing or providing of skills and space. It emphasises the research and design that has gone into developing the site and has sophisticated functionality, for instance, offering an insurance option for lenders so that they can reclaim the value of the item if it is damaged or not returned. Users can set up sharing circles to keep their activity within bounds of their existing trust. It allows for the hiring out of items as well as free distribution and this is the business model of the social enterprise behind its construction: a percentage is levied on any financial exchange. There are a range of benefits and motivations on the site: save money, reduce waste, reduce production, reduce debt, build local connection, change public attitudes and eventually shift industrial practices. There is also a statement of ethical position: 'We commit to donate a proportion of our profits to the Ecomodo Foundation which will directly support community and environmental causes. We choose suppliers and partners with strong ethical and environmental policies, For example: RSA (The Royal Sun Alliance) who provides carbon neutral insurance. We use recycled materials and we minimise travel. We consider inclusion and accessibility. We encourage stakeholder involvement. We practice equal opportunities.' In posting this, it presents a congruent message of care for the environment and social causes.

There is attention given to processes to encourage people who are strangers to share willingly, presenting a low-risk service experience. The site reassures users that it is easy to use, with many protections that do not occur in everyday sharing. Specifically, 'Ecomodo make it easy for Lenders and Borrowers to act with confidence' by supporting: a find and browse interface; introduction and negotiation between lender and borrower; simple dispute resolution; registration of a loan agreement; holding of hire fees and deposits and paying out funds on completion; reputation building through feedback; and provision of insurance for higher value items. Ecomodo does not assume people start from a position of trust.

Like any site, a vibrant user group is needed to encourage participation, because without anything to borrow or lend nearby, people can lose interest and revert to other ways of getting what they need.[6] Examples of things to share include tents, awnings and golf clubs, sensible in that they are robust, substantial (so worth going to some effort to locate) and not much used, but also presenting a model of a high-end, leisured user. At time of press, there were 2040 items to borrow on the site.

Streetbank: http://www.streetbank.com

Streetbank, 'share things with your neighbours', has a styling that more completely masks its government funding and a stronger emphasis on forging community links than Ecomodo. It presents its story as a simple journey of discovery: 'Hello. We are Sam, Nic, Tess, Kate and Alex. A while back we got interested in the idea of lending. Sam had had a good experience with his next-door neighbours. They had been lending stuff to him – small stuff mainly (like a cup of sugar), but it got bigger (like a ladder) and in time he found he was actually hanging out with them just for fun. As Sam chatted to friends he started to think it could be more than just lending, like people giving things away and helping each other too.' It also addresses the potential barriers to engaging with others, but handles it differently: 'By joining Streetbank you are saying to your neighbours, I'm interested in my community and I'm happy to lend a few of my things and help out in a few small ways. Knocking on a strangers' door is quite a bold thing to do. Streetbank reduces the barrier.'

It has an offers/requests model, specifying distance: 1 mile, 5 miles, wider for rural areas[7]. Most people do not include profile information; of the few that do, they expressly mention wanting to

[6] In the Brockley area, we found 6 users, all of whom were offering one or two objects for hire, and little evidence of interaction.

[7] In Brockley, there were around 200 members with about 100 things offered, at time of study (April 2014), with no interaction - of the 8 requests current, most had no response.

build local connections. Active users were treating the site like Freecycle (p53), divesting themselves of unwanted goods. At time of press, there were 56,598 neighbours sharing 56,064 things, showing a significant rise on the figures for April 2014.

Peerby: https://peerby.com

Peerby is a sharing service that was developed in the Netherlands and has considerable take-up there (Reimerink 2014), now launched in UK: 'Peerby's unique approach lies in its proactive approach: you don't just post the item you're looking for; the site asks your 100 closest Peerby neighbors if they have the item. This way we achieve a higher success rate than any other sharing platform.' Like the other coordination sites mentioned here, it suggests that to participate in this sharing culture is better for you, your pocket and the environment, and offers, through the site, fun, safety, convenience and connection. The site is eloquent about its backing: 'We get funding from the three areas we believe in. Namely the three pillars of the Peerby vision: social cohesion (i.e., funding from Stichting Doen), sustainability (we have won prizes in the United Postcode Lotteries and Clinton Global initiative), and social entrepreneurship (funding from media partner Sanoma Media). In addition, we are also testing several additional user-based revenue models, such as insurance for lenders and voluntary contributions from satisfied members.'

At time of study, the service was considering how to deal with disputes, asking people to trust rather than starting with regulation, resolution and feedback, in contrast to Ecomodo, where prevention of risk is a central premise. It offers its own understanding of responsibility: 'we are developing a feedback feature, which will allow Peerby members to rate each other after a (successful) transaction. The goal of this is to ensure that people who do not treat the belongings of others with the respect and care that they should will not be able to rent something in the future. … In the meantime we trust in the integrity of our Peerby members and we will do anything we can to help you, if you do happen to run into a black sheep that managed to sneak into our lovely community.'

Casserole club: http://www.casseroleclub.com

Casserole Club is included here as a more focused product than the generic sharing sites above, with a service that is closer to gifting than sharing: *'*Casserole is a project that connects people who like to cook and are happy to share an extra portion of a delicious home cooked meal, with older neighbours living close by who could really benefit from a hot cooked meal.*'*

Rather like the Mensheds example in Lewisham (p31), the service relies on cross-sectoral support and collaboration, taking referrals for "Diners" through local funders where it is active: *'*We've learnt that to rollout the Casserole Club in a new area we need to work closely with local organisations like councils, housing associations and charities to sign-up Diners and Cooks. It's also really important that we provide excellent customer service to our Cooks and Diners, and so we need access to funding to grow in new areas.*'* Meeting existing need, it provides a different model of feeding older people to councils' meals-on-wheels services. Indeed, behind the appeal for volunteers, is a *software as a service model* for helping local authorities to fulfil their commitments: it is built by FutureGov and licenced to councils for about £50K a year. "Cooks" are required to sign up on the site and undertake a short safeguarding process before they can search and contact local "Diners". Although it presents a friendly neighbourhood service, it offers a different kind of engagement from noticing that a neighbour might need support or sharing when you have left-overs. Although the initial commitment is to share one meal, thereby making the barrier to entry low; there is a specific hope on the site that the relationship will develop and endure.

Meanwhile, the service recognises that both donors and recipients are not necessarily the high-tech users of Ouishare, etc. 'As much as possible we encourage Cooks to sign up to Casserole using the website, as this makes our safeguarding process much simpler. The website also gives Cooks the ability to search for local Diners and to see who else is cooking in their area, and so is a really useful tool to help you get cooking quickly.' but 'We understand that not everyone has access to the internet, and this shouldn't stop you from being able to take part, if you can't access this internet but do want to Cook - please get in touch at hello@casseroleclub.com or 020 3475 3444.'

Taskrabbit https://www.taskrabbit.com

Task Rabbit is one of the more prominent services in the world of collaborative consumption, noted as being a success in terms of business model and scale (see Wiggin 2014, Yu 2012, Singer 2014). Again, the description on the site starts with a story of personal experience that led to a revelation: 'It was a cold night in Boston in 2008 when Leah Busque realized she was out of dog food and didn't have time to buy it herself before a dinner date with her husband, Kevin. Leah thought to herself, "Wouldn't it be nice if one of my neighbors - perhaps one who was already at the store at that very moment - could help me out?'. Again the rhetoric is of neighbourliness and care: 'From this everyday experience, TaskRabbit, an online and mobile marketplace that connects neighbors to get things done, was born. Fully-vetted, skilled professionals contribute their time and expertise, while busy people find a little extra time in their day to do the things they love. Neighbours helping neighbours — it's an old school concept reimagined for today.'

The site enables someone seeking domestic help to find a local person looking for the work and to employ them with confidence that they have been vetted, rather like a casual employment agency. This service also involves insurance. Trust is dealt with through regulation. Taskrabbit conducts identity checks, criminal record checks, an in-person interview and training session. This is clearly no longer coordinating tasks with your neighbours to help each other – despite the rhetoric of the founding narrative. This shift carries with it a similar message to that of Ecomodo: that mediation, online payment facilities and regulation make us safe.

Freecycle: https://www.freecycle.org

Last in this section, we point to the under-designed, organically-grown precursors that grew out of need, imagination, opportunity and effort in the digital realm, which includes services such as craigslist (http://london.craigslist.co.uk/) and, particularly, the international Freecycle. Botsman & Rogers (2010) discuss how services such as Freecycle paved the way for other more business-oriented activity. Freecycle is a series of email lists based on neighbourhood, co-ordinated centrally through a website and locally by list moderators, amounting to 5,000+ groups worldwide with 7 million members, grown over 10 years. Individuals sign up for the list and see everything offered or wanted in the membership – very busy in some areas, quieter in others. The website describes it as a: 'worldwide gifting movement that reduces waste, saves precious resources and eases the burden on our landfills.' Most notifications are from people with something to get rid of that might be of value to someone else. There is no mediation – you contact the most appropriate respondent to make an arrangement; then post again when the item has gone.

The Coordination Opportunity

It is apparent from the descriptions above that there are many contrasting kinds of activity being coordinated by these different sites, despite a relatively consistent positioning across each within the new social and environmental practices of the evangelists' world. We look at the differences and use them to pull out questions for use by designers and policy-makers thinking about positioning.

Ecomodo was picked as a starting point here because it straddles the social, environmental and economic concerns expressed across the whole group and seeks to facilitate a very wide range of sharing-related tasks and practices. It is possible to share skills, space and items and also to rent them out, for gain or for charity. It is designed to work at a neighbourhood level and build circles of trust through existing relations and developing new ones. Streetbank concerns itself additionally with the social elements of sharing space and helping out and Peerby with alerting potential sharers. Notably, in Ecomodo, there seems an assumption that people operate in a community where trust must be scaffolded on formal mechanisms and so control, fees, deposits, insurance, feedback and dispute mechanisms are all part of the design. By contrast, Streetbank and Peerby are more optimistic in their structures and expect to contribute to existing informal cohesion strategies and grow trust organically. At the far end of this spectrum is Freecycle, which leaves all the work to email list members.

Questions we might ask, then, are: How is the risk of loss, harm or embarrassment handled? Does the service encourage us to trust each other or trust the site and its procedures (and the wider legal and economic frameworks of society)? What model of social relations does it embrace and embed in trying to win a critical mass of users? What understandings of trust appeal in different contexts?

When we look at Casserole Club, we see another form of sharing proposed; one with a wider network of agencies implicated in the brokering function of finding and matching people. The service is integrated into existing provision for people's vulnerability and many referrals come from old-style infrastructures for assessing need for care. The language is of individual help and the relations may be transitory between donor and recipient, but the work of volunteer cooking can be used to replace and/or supplement support provided by municipal teams, where, instead of 'meals on wheels' produced at cost to the council, the local arrangement puts the costs and effort of food purchase, cooking and delivery in the hands of the donor. In its relations with other agencies, this service reminds us of the initiative of the time bank and the Mensheds project, though both of these are offering supplementary forms of interaction in the area and neither is based on an individual to individual mapping.

The site is, of course, concerned with promoting take-up of the service and offers Cooks a FAQ section to reassure them. There is little information about what Diners might expect, which indicates something about how the service understands its relation with Diners: they are not the site users. Nonetheless, we might ask if the new economy means greater choice and flexibility for recipients of traditional support? Do they have to eat broccoli? Is there a stigma or a social stress to having different people bringing meals? Being beholden to individuals is very different to being provided with supplies by the state, which is anyway a matter of concern to some. (As Philippe notes in talking about timebanking, reciprocation has a positive effect- p33.) Casserole Club is, perhaps most clearly of all these sites, the intersection of choice and need.

Questions, here, might include: How are relations with other agencies established and signalled? Is there dependency on existing infrastructure for success? Does the service augment local capacity or replace it? What opportunities and challenges come with a change in service provider and shape of provision? How is the transition to more personal services managed?

If we now look at Taskrabbit, we notice something of an inconsistency. Dubbed repeatedly part of the sharing economy (see Wiggin 2014, Yu 2012, Singer 2014), it is difficult to see in what way Taskrabbit is a sharing service. Instead, it creates a micro-market for tasks, based on availability relative to location. Unlike Amazon Mechanical Turk (https://www.mturk.com/mturk), it does not crowd-source execution of virtual tasks, but physical ones. It has come in for some scathing reviews for its 'gig economy' work ethic: 'all I found was hard work, low pay, and a system that puts workers at a disadvantage' (Kessler 2014). We note that a similar ambiguity appears in descriptions of sites like AirBnB, which rents out spare space in homes and empty flats, and has been criticised for undermining sharing (Tsur & Belsky2014) and doing too little to ensure people do not evade tax and health and safety responsibilities, as B&B marketing and trading opportunities are improved by the ease of use of an online trading space.

We can now ask: How is the sharing economy related to deregulated trading? Are these global-local economies putting financial value on former gifting activity?

If we take a moment to consider the marketing further, not only are we struck by the disparity between talking about sharing while actually selling, but also by tensions in the founding narratives that support the positioning of the services. In trying to sound open and accessible, both Taskrabbit and Streetbank imply that the projects were conceived after a revelation in adult life by, it would seem, people who had minimal previous experience or expertise in building local networks and strengthening communities through sharing action and responsibility. In contrast, Ecomodo stresses the research and design professionalism that has gone into developing the site, offers a commitment to ethical behaviour and uses links with commercial and media partners to show its credibility. Its use of links to reputable (insurance) companies and old-style press testimonials is a position far closer to that of Casserole Club, targeting its audience with stories based in different expectations from the other two. Meanwhile, Tonkinwise points out that: 'transparency about the intentions of a sharing system's administrators must be part of their brand, irrespective of whether it is associated with the fields of sustainability or sociality, etc.' (2012).

If we look back to the hOurworld site, we see the language of historically situated and embedded community organisation: 'As individuals we each participate in the community currency movement and we honor all the good people who have contributed to this system of service to others. We pay special tribute to the seven women that originally founded time banking in the United States.' The cooperative site has no need to claim originality and appeals instead to a long tradition.

Patchwork Present's Olivia also tells a personal story of arriving at her business idea, in the case study (p28) and on her site. We see similarity to the founding narratives that appear above. In her case, it is based on her understanding of how people come together to share buying gifts. We note that in the sharing economy it is in vogue to bring stories back to the personal in the narrative of innovation, and it is one of the markers that identify a service as entrepreneurial and not part of business-as-usual.

New questions, then, include: How does the rhetoric of the site position the service? Is it consistent with aims and target users' interests? How is it inventing the world it wants to see?

Last, we can see that take-up varies (with Freecycle showing what a system can do over time with little design or styling and no intervention) and we can relate this to design, marketing, target audience and other factors that will affect the success of any new product. We note that in our (limited) study of Facebook friends (p44), most people were managing their exchanges through existing media, such as social network sites, email, texting and phone. They were not turning to specialist sites to run their affairs, but were integrating their sharing needs into their everyday lives and, with that, their everyday uses of technology.

Challenges

In summary, we can ask what might prevent digital sites that facilitate local sharing from becoming popular. Without sufficient activity to reach critical mass *for any particular location*, take-up will continue to be low or patchy - there needs to be both some point in connecting and a local cohort to connect with. Apart from social awkwardness, lack of trust and absence of things to share, which are all on the user side, other more generic obstacles might include:

- new layers of effort are required to lend something for those who are already predisposed to do so, thus overcomplicating what could be simple;
- the site requires people to use digital tools in a context with no direct digital element;
- the ideas would work better if presented in groups' social network of choice rather than a custom service;
- the service appeals to people without much connection in the community so it does not plug into existing social structures;
- the service is too disconnected from what is going on in the local area to inspire interest.

Transport sharing tools: cars and bikes

Informal sharing of transport is a major existing cooperative endeavour, seen in casual ride-sharing, regular car-pooling, joint ownership of a car or shared car bills between neighbours. Brereton & Ghelawat (2010), in their study of the use of a ride-matching tool, describe how what may start out as a relationship based on getting to work efficiently often becomes friendship over time, leading to people walking as well as driving together.

To these many local sharing arrangements, we can add the more formal processes of transport systems such as Uber, Zipcar and the many city bike schemes that are now available to subscribers. These use technology, principally GPS location sensors and Internet of Things connectors, to provide a joined-up service across large areas, supported by mapping, coordinating and payment infrastructure on a large scale.

Uber (https://www.uber.com/) is like Airbnb in that it allows private individuals to use spare capacity: it brings together the drivers of private cars with time on their hands with those needing a ride, through a service that coordinates this pairing, handles payment and registers/vets. It is a paid taxi service, not a ride-sharing scheme. The service is highly mobile in that it has been designed for the phone and it has sufficiently dented London's regulated taxi services with its more 'open economy' model to have given rise to mass demonstrations by cab-drivers.

Zipcar (www.zipcar.co.uk) is, since 2013, owned by car rental giant Avis Budget Group. It is essentially a localised car rental subscription service that uses digital technology to organise short hires, releasing cars remotely to subscribers, unlocking doors, monitoring use and calculating and subtracting fees so that drivers can book a car based in a street nearby and return it without having to travel. It too makes big claims about changing attitudes and behaviour: 'we are redefining the way people think about transportation' and presents a mission with values beyond the immediate service delivery to enable 'simple and responsible urban living'. It is selling a lifestyle: 'We envision a future where car club members outnumber car owners in major cities around the globe. …Zipcar will be an integral part of these vibrant communities of well-informed, connected people who enjoy urban life and transportation options.' Zipcar is mentioned because its use of technology, distributed nature and support for shared use of resources speaks to the 'access not ownership' philosophy of service delivery in the collaborative economy, as does its rhetoric, but it is based on an updated model of a privately-owned fleet made available to people on an as-needed basis, by booking, which has long been with us.

By contrast, the provision of timeshare bikes as a city service in London is closer to the provision of buses, with a hop-on-and-off quality that cannot be booked, though also with digital technology as enabler. This is used similarly, to register, unlock, record duration and redock, but also to monitor clumping and trigger vans to transport bikes between locations, so that docking points are neither full nor empty as peak flows move bikes unequally round the city (see Cocozza 2014). As we have noted, when distribution of resources is in municipal control, it is largely invisible as sharing *per se*, though here it is the provision, not the distribution, that Transport for London oversees. Orsi comments in her analysis of sharing and coordination (2009), that this level of sharing requires formal infrastructure: 'Whether this is publicly or privately managed, … [t]aking sharing to the fourth degree can require getting government buy-in, mobilizing multiple players (legislators, investors, banks, developers, planners, etc.), or even restructuring our communities.' It is, of course, much harder to move car gluts around than bikes, and generally you return the car to where you found it. Again, the emphasis is on access as needed, replacing ownership of dormant resources.

Uber and Zipcar are global in ambition, and bike schemes are appearing in more cities each year. All respond to the need for access to be hyperlocal as a genuine challenge to ownership, but only the bikes are wholly reliant on the infrastructure of individual cities to support them. The different economic models imply different flexibility to local conditions, though all must be able to provide a service *as needed* or their value to subscribers will fast diminish.

What is notable about these activities is that only the resource is shared. No one has to meet anyone else, even virtually, to make this work. And as long as co-ordination of slots is managed successfully by the remote network, there need be no awareness of others in the system. By contrast, a car club where people allow other people to use their cars shows the contextual complexity that Benkler

(2004) says comes when you move away from the market. Decision-making about who can book when is done in an automated, anonymised way at Zipcar; whoever books first is able to use the car at the time of their choice. This process can be more subtle when people have to meet and interact to borrow or lease a neighbour's car. Decisions might be made on the owner's perception of merit, need or how well they think someone will treat the car. This has advantages and disadvantages.

Questions arise, such as: Is public transport sharing? How is sharing resources without encountering other users different from schemes which bring people together?

Discussion: Conspicuous Collaboration

In this report, we have described brokering, verifying, organising, scheduling, communicating, aggregating, scaling, sub-dividing, managing, paying and distributing, to mention a few activities that sharing can entail and digital tools can support. **We have looked at these in terms of innovation: first, what people will set up locally and why; and, second, the high-profile services of the sharing economy. The kinds of sharing seen in our local examples are quite different in tone, scale, ambition and practice to those that are most widely discussed in regard to the idea of a sharing economy.** In looking at the former, we notice the emphasis on organising together to create shared spaces for collaborative use of resources. In looking at the latter, we notice that they start from a position of sharer-to-sharer individualised activity and some have a tendency to monetise value, rather than encourage sharing as a social exchange mechanism. The sharing economy refers to a range of practices and business models which also feature significant elements of renting, leasing and hiring. While these digital initiatives may be a stepping stone to a more embedded use of technology, with connected devices and internets of things, they are also acting as powerful agents in our understanding of what it is to share and how we do it, so distinctions in interpretation matter.

Coming at a point when digital technology is increasingly able to offer a convergence of mobile and real-time cloud-based big data processing, sensors and social media, to sense, connect and control, we might speculate on a number of new types of engagement. We can make tools and things that know where they are and who is using them, when and for how long. We can imagine items that share themselves out and see prototypes in the car services that are centrally booked, but locally unlocked and returned. We can envisage time and space more evenly apportioned and sharing protocols managed autonomously by smart systems in our community resources. But inventiveness is best tempered by an understanding of contexts of use, group goals and personal interests. So our section here looks at what we have learnt about how, why, when and what people share in our studies.

Tool Use

Well-connected people, whether locally or more broadly, turn first to their networks and use existing tools to structure activities, communicate and mobilise. They rely on existing trust relations, sometimes extended over social media. This does not mean that they do not sign up to sharing economy sites (and may use rental services such as AirBnB or Zipcar for travel or crowdfunding sites such as Kickstarter to see their friend's movie made). However, to augment a neighbourhood with sharing facilities, it takes a critical mass *locally* and some added value in using a dedicated tool to make the overhead of take-up worth the trouble for people with a good network. Further, many people interested in the social and environmental wellbeing of their neighbourhood have little interest in using new technologies and are not aware of developments in digital sharing. Our research suggests that what exists does not match up closely to what they need, seen in responses from both the Brockley and Facebook sharer studies. This leaves open the question as to whether these sites increase the numbers of people who make connections in and across their communities by digital means. It also leaves plenty of opportunity for a new wave of sharing technologies that work with communities to support them, such as Heitlinger's watering can (2014).

We have not looked extensively at sharing support tools, but the structures of Time & Talents and its configurability are strengths when supporting enablers, such as the Rushey Green Time Bank team. People identified other management needs and it would be possible to develop tools that help support sharing in a collectivist or communitarian way, especially if accessible through existing interfaces and devices. Some of these might be one-off or limited-use custom pieces, which is likely to put their development in the same arena as their deployment: a product of the core economy and a labour of love, not money. Unsurprisingly, that is where Time & Talents sits, developed by timebankers, organised in a cooperative and distributed free.

So, in considering the digital, we need to look wider than tools explicitly for sharing to how sharing works with broader digital tools, and wider than the individual to the neighbourhood and the 'embedding' of the sharing. The study shows that individual enablers are a rich stimulus for local sharing and a move towards a culture of sharing (eg the acceptance and use of the microlibrary, p21), whereas becoming another enabler is not necessarily well-supported by sharer-to-sharer sites.

Place and Locatedness

Relation to place varies widely across the study. Some services attempt to scale across multiple areas; some can only be reproduced discretely in each, such as the microlibrary or garden. Whereas sites sharing virtual resources have no need of local moorings (Patchwork Present, which brings together money, can broker joint present-buying across the world), other services, which deal in physical resources, might be considered more accurately as thousands of local clusters within a global frame. This poses problems for all brokering sites promoting material exchanges, as they have to establish themselves with a sufficient number of people sharing sufficient material in each locality. Peerby's model of sending out notifications to all members local to a possible exchange could thus be a key factor in its success.

One way of embedding locally is to work with local partners, seen in Brockley (eg Mensheds) and more UK-wide (eg Casserole Club). This can provide a direct connection to potential service users, but, especially when the partner is a local authority or health service, may only target people already in the care system.

Granularity – and, with it, critical mass at the most local level - affects what will take off. For instance, the time bank has five clusters across Lewisham, the local council area. To further its work, it could develop more clusters and/or become denser in membership in each area and grow this way, but not usefully expand the width of its catchment areas, as Philippe notes (p33).

Even the positioning of a service as global can be problematic, as hyperlocal news sites manage to engender greater trust among and between users than services focused on a function (rather than a locale). More effective is to grow one or multiple local contexts at outset, a bit as Patchwork Present and Streetbank are seeking to do, since later these can be aggregated, whereas the absence of context, personal effort and a local element in global sites works against take-up. And, as we have observed, there is a tendency to produce sites with homogenising messages, interfaces and back-end mechanisms even when it is appreciated that local culture will impact take-up.

Further, if we look at the factors that have influenced the area on which we focused (the 30 minute walking radius from the researcher's home), we see the importance of mutual support (or *relational* assets) across organisations and from well-established sources, with 40-year-old Brockley Society functioning as incubator for other enterprises, such as the microlibrary and community garden.

Timeliness

In thinking about the initiating of sharing, we can see an intrinsic difference between what normally happens in trust networks and what happens on sites aimed at better utilisation of resources. Posts on Facebook ask to borrow something as needed; they are not random offers to lend. It is often easy to source something through a local network as people like to help out. But there is no reason to hawk about personal resources in case someone happens to want to borrow them. A courgette glut might be mentioned on social media with an invitation to collect; a power drill would not, unless you were, similarly, seeking to give it away. Even if posting news of resources is efficient in terms of promoting sharing (offset against travel time and fuel use if the person is not truly local), we still contend with lack of timeliness (Grice 1975). Without an angle that makes something worthy of mention, such as perishability, posting up resources can look like showing off your assets. So, it is hard to learn of possible community assets in the abstract (Philippe, p6). Ecomodo and its kin suggest that we show what we have tucked away so that others can make use of it. But, if there is no immediate issue to solve, we may not be disposed to list our items or skills; we may not even think of them. And at the point we receive micropayments for their use, we may now have an incentive, but it is no longer sharing as we know it. This is the tension in models that ask us to list assets out of context. The moment of need for mobilisation at the Ivy House pub revealed a range of skills people had not shared before. Use of Freecycle to dispose of unwanted goods is timely in this sense too; sharing sites are not (though Streetbank's use in Brockley showed similarities to Freecycle: disposing of unwanted goods, not reciprocal sharing).

Financing and economic models

The financial model for initiatives varies according to activity: building transaction mass and levying a percentage for brokering services (Patchwork Present, Ecomodo, etc), hourly rental (Zipcar, bike schemes), independent fund-raising (fair/advertising for Brockley Society), grants (Mensheds at startup), share issue (community pub), local authority fee (Casserole Club), awards from charity, etc. Much of the sharing economy is supported by venture capital - which dictates that profits must be made - and not co-operative models of finance. There was more potential for public subscription than we saw, ie the model that first supported public libraries and other membership-based or public facilities financed by the more affluent in a locale. This, in a more contemporary form, is what crowdfunding enables. The community pub's share issue (p23) comes close.

Overall, socio-economic structures embodied in the local resources in Brockley and those of the sharing economy digital services differed widely in terms of individualism, the trust motivators they assume and the degree to which they see sharing as a social or economic practice, although there is also a mismatch between the rhetoric and the functionality on some sharing sites, obscuring this to some extent. With digital platforms, back-end functionality, interaction at the interface and the message in the text of the site are all more or less deliberately designed to support a particular understanding of social engagement, but they do not all have to point in the same direction.

Accessibility and inclusivity

Accessibility issues affect both tools and services. The former includes using popular models of device, not assuming that everyone can or will access your services digitally and making the service easy to use for people regardless of circumstances. The latter includes ensuring that people who are not already in the network will find out about opportunities and feel comfortable taking them up. We might also call this inclusivity.

When we consider take-up across different groups, sharing supported by digital tools has the advantage, as well as social drawbacks, in circumstances when it does not require encounters between sharers. The digital system can be less personal, choosing dispassionately on the basis of first-come, first-served, rather than weighing up all the factors that personal lending might involve, which could include judgements on grounds such as need, but may privilege those known to the lender, then those who are similar in outlook, class, ethnicity, etc.

This is illustrated in the transport examples (p55). As noted, there need be little social involvement in using a car club: the technology can serve a rental service, car-pooling or a mesh of self-organising drivers. As further technologies come in, such as the remote on/off lock/unlock mechanisms of Verizon's Auto Share (see Hambleden 2014), access becomes manageable through a smart phone for a range of devices. We can look past whether money is involved or even if someone makes some profit somewhere, but consider this in terms of efficiency of service provision and maximising take-up across the whole community as budget allows. The advantage of these impersonal exchange systems is that no one has to work alongside anyone (or even take their first brave step into the community garden).

However, with remote monitoring comes the opportunity of remote control and penalising those who do not manage their resources. Rental cars are already being switched off when drivers run behind on payment, sometimes mid-drive (see Corkery & Silver-Greenberg 2014). For people without resources, the move to monetising everything would be the most exclusive condition of all.

Making Stuff; Choosing Our World

This report is written as part of contributing to the Sustainable Society Network+ and so we conclude by asking what kind of sustainability we are considering in thinking about our societies and our sharing practices. The next section looks at alternative visions for a collective approach.

In doing this, the politics of economic structures are critical: they underpin our understanding of resources and assets, their distribution and our day-to-day access to and management of them. Sharing is one filter for thinking about economics, who has what, who needs what and who gets it. It is also a series of existing practices, as evidenced by our study of a small area in South London (p20 to p39) and our widely distributed circles on Facebook (p44 to p47). It is also the engine driving the sharing economy in some form or another, which we have critiqued in looking at evangelists, coordination sites, transport sharing tools and sharing support tools (p48 to p57). Design for sharing creates the social and technical infrastructures for more participation and new relational assets, as the connections between people become as important as the things people possess. But with the question of what we design comes the question of what values it introduces into society and what we have embraced. And technology may be helpful, irrelevant or hindering when we answer that question.

The political challenges, as we face the future, are complex. They include how we negotiate the distribution of fewer overall resources, while keeping an emphasis on social justice. We have to do this in a careful balance between individualist and collectivist cultures holding very different interests and priorities, but competing in an increasingly global context with an increasingly unstable climate. This difficult work needs more than the rhetoric of sharing; it needs policies and clear sightlines about how to take decisions and what actions lead to which impacts.

Resources which were available in the UK through the 20th century mechanisms of a supportive state are being redistributed and distributed through new mechanisms. Thus, we ask how sharing can be made more sustainable in ways that replace or enhance services, not out of a reluctance to use a safety net, but because it is apparent that the safety net is no longer meshed finely enough to protect some of the most vulnerable amongst us.

Sharing at its more informal - of swapping goods and talents as needed - will continue because of and despite the network tools that can broker it. Trust is the currency here and trust may prove increasingly tractable to being extended to strangers, as the generation reared on social media engages with these issues. Tonkinwise takes a sceptical view: 'Just as the word 'friend,' now a verb, does not really correlate with what the word means in a flesh-based world, so the word trust appears to be stretched thin when digitized.' (2012). Nonetheless, informal patterns of exchange that cross existing networks to link communities and circles have an important social cohesion role.

But this is only working at the first degree of Orsi's four-part analysis (2009). At the next degree, the Mensheds and timebanking examples show how some of the vulnerabilities of ageing, infirmity, isolation and depression can be tackled in ways that are novel and come in addition to the standard forms of support available through the state. This layer of care can be supported by digital tools and a sensitivity to the challenges of building shared spaces where collaboration can happen. More types of space can be created that serve communal social needs as well as using resources efficiently.

Does sharing efficiently reduce the need for current scales of production? Innovation in social practices and technological support for them suggests new industries, while threatening older ones. Car manufacturing may shrink as it is hit by better utilisation of vehicles. Service management and related ICT development will grow. Replacing consumption with reuse models would ideally require a move away from upgrade cycles and built-in obsolescence, but this would be a significant change in conventional trading practices.

Sterling (2014) describes the rise of global superpowers that are commercial, not territorial, in their reach. He puts a powerful case that five US technology companies already have the means to manage our lives, determining what interaction is possible with things and media. Such strong commercialising forces could insist on monetising every transaction, furnishing the technology and providing the blend of surveillance and data gathering to send some social practices underground. Those with the wealth to shore themselves up against riot, rebellion and rising water levels do not have to be part of the societies created this way; the rest of us do. Sharing, in this context, is for the other 99%. Sharing, then, is beyond just a practice or rhetoric, it goes right to the heart of who is able to claim a place in the world.

So our question is not only how to design sharing, but in what direction. Stokes et al (2014) point to four complementary drivers of the collaborative economy:

- Creating value of out of idle assets
- Rebuilding social capital
- Creating new economic relationships
- Environmental sustainability

This diversity is entirely understandable when defining a sector, but to design for this sector requires either a way of integrating or prioritising these goals. Design is always to a particular end. We might have to separate our vision of sustainable futures before we reassemble them. When we design for sharing, do we:

- design for optimal resource management?
- design for improved societal relations?
- design for economic (global, national, local, communal) vitality?

Strategy One: Design for optimal resource management

Efficiency of resource use supports the environmental case and to some extent the economic one. As designers, we can use the power that computer networks give us to connect things up and 'design to manage and sustain limited sources of materials, time and enthusiasm' (Light 2011). The principle functions to support, then, are distributing, scheduling, unlocking and docking/returning, with guarantees at every stage. Payment systems would facilitate the growth of these services and this model makes the space for innovative and disruptive services of the kind we are seeing in the sharing economy, where use of spare capacity is creating secondary and new markets.

Just as local government services facilitate co-existence without requiring individuals to meet and work together (to dispose of waste or lay roads), so digital platforms can offer an infrastructure that takes the direct collaboration out of sharing. This corresponds to the fourth degree of sharing proposed by Orsi: 'You reserve a car using your cell phone, punch in a code on the car door, get in, and go! Whether this is publicly or privately managed, launching such a program involves significant investment of time and resources and a rather complex system of administration.' (2009). Both business and not-for-profit systems are appearing, and we note that this is made possible by digital platforms at each stage. In fact, Kiessling suggests, in writing about the Internet of Things (IoT) and sharing, that 'For free-floating car sharing the users must, for example, know where to find available cars. GPS receivers in the vehicles and an app that shows all available cars on a map assist them.' (2013). In each case, the introduction of the fourth degree of organisation goes hand-in-hand with providing a service that requires major resourcing and centralised management and, thus, little local coordination, allowing people to access resources without having to encounter each other. Just as the ATM removed the need to discuss money with a bank clerk, so mediated sharing manages the availability of resources while allowing people to ignore each other if they choose to. This weakens it as a social model, since it promotes exploitation of resources (including other people at times), may lead to disparities of access and requires no one to practice gainful co-existence as part of using resources wisely. And Sennett (2013) has warned us that cooperation is a craft to learn and practice.

Meanwhile, Kiessling's IoT-enabled sharing utopia (2013) is countered by Morozov, who links together the sharing economy, networked technology and micropayments in a way that fills him with dismay: 'This sharing imperative dictates that everything that we own, from tangible assets to intangible thoughts, be categorised and assigned some kind of a unique identifier like the QR code. When somebody somewhere – it could be our neighbour or an advertising company across the ocean – expresses an interest in "borrowing" an item that matches the description of what we own, our phone would notify us of their offer, pitting us against all the other "micro-entrepreneurs" with similar ownership profiles. Once we accept, the rest is logistics: a drone or a self-driving car would stop by to fetch the item … and the payment would safely arrive on our smartphones.' (Morozov 2014). As for a circulating stream of more fully utilised goods, he dismisses its usefulness in staving off environmental catastrophe as 'overstated and often risible', pointing instead to the need for big business and power generation companies to rethink their processes.

Strategy Two: Design for improved societal relations

To promote social wellbeing, systems need to help manage trust, access, connection, the breaking of social boundaries and social inclusion, while leaving intact a sense of culture and place. All of the initiatives seen in Brockley have this goal at their core, as well as interest in environmental and economic sustainability. Key digital functions might be recruiting, matching, minimising the work of building trust, voting, helping to structure and maintain collaborative spaces and make others aware of them and the opportunities they present. They need to support Ostrom's (1990) principles seen in managing the commons. (The questions peppered through p50 to p57 address these kinds of concern.) Further digital functions might especially target perceptions and realities of inequality, since inequality causes resentment and division. These functions might include mechanisms for distributing transparently and, even, collaboratively gearing contributions and distributions according to capacity and need.

Much of the design work that promotes social relations is at the second or third degree in Orsi's system (2009), where systems are informally or highly organised, but not out of the hands of the local communities that use them. So, for instance, we cannot conceive of Mensheds (p31) or the community garden (p26) as software, but we can anticipate digital tools to help run the spaces well. Would a door that opened upon being scanned with a phone, or a system that showed which plants were currently too dry and needed water be assets or complications in running a community garden? At what point, as the resource develops, could such interventions be usefully offered and how do we prevent the technical dictating the social system, rather than the other way round?

We might ask how to take forward Rushey Green's time bank with its brief to improve mental wellbeing? Bellotti et al (2014) raise opportunities that applying further new technology could bring for mobile timebanking applications: '[O]ur vision is to add 'smart' and context aware capabilities, to enable our system to direct service and resource recommendations to those best placed to perform or provide them most efficiently.' (Bellotti et al 2014: p2983). In working with mental health, how might such a system 1) keep members from ignoring jobs because they are awkward (p34), without telling timebankers what to do and violating the basic principle of 'dipping in, dipping out'; 2) keep in sight that the role of the facilitator is to employ subtle criteria in matching members and their (best) interests, and that context-aware systems are better able to judge a more market-style (ie formalised) set of options, like proximity or stated skills; and 3) keep configurability and self-determination for the group, so it can promote effectiveness for health rather than efficiency, ie not dictating the culture of timebanking with the application.

Creative measures could support a love of sharing and contributing to the community without a direct emphasis on management or efficiency. One example is the Community Hacking project on the edges of Glasgow (http://www.communityhacking.org/), which uses IoT tools to support a local sense of belonging. It shares stories, rather than tangible goods, but, in doing so, promotes conditions for other forms of support. This latest embedding of technical innovation joins a history of community broadband, meshes and other technical infrastructure that promotes access in alternative ways and acts as enabling community assets, and many devices for storytelling.

It is argued that community cohesion is a necessary prerequisite for action on environmental and economic resilience (eg JRF 2014) and that we need social infrastructure not as a platform, but as a core to handle crises from other directions. If so, bringing things down to the very local, so that large numbers of people become involved, sets this direction in contrast to Strategy One. The two visions are not incompatible, but we notice that scale is important - with smaller infrastructures and organisational platforms favouring social development.

Strategy Three: Design for economic (global, national, local, communal) vitality

While neither designing for efficiency or sociality is at all clear cut, designing for economic vitality is the most controversial strategy of the three, since there are so many ways that this vitality can be understood. Scale matters here too: do we want a global scene in which wealth is generated to the maximum, or thousands of localities where people are gainfully employed in activities that produce rich and sustaining outcomes? Can both or either be achieved by working at the level of the state?

On the one hand, we have Sterling's vision of a world run by what Cawson calls 'corporatocracies' (Cawson 2013), super-power companies that know more about our actions than we do and use this knowledge to generate more and greater data-led businesses. On the other, we have the argument for increasing local co-production and encouraging participation in the areas that the *Digital Social Innovation* report identifies: new ways of making; participatory mechanisms and open democracy; awareness networks; open access and information commons, mesh networks projects, etc. (Bria 2014 pii). Coote & Goodwin (2010) argue that our aim should be to 'devolve power and encourage local action wherever possible. … opening up opportunities for people to take control over what happens in their own localities and providing access to resources that will make local action feasible and effective' and that 'everyone, including those who are disadvantaged and disempowered, has assets and resources, not just problems to be fixed by others.' (p4-6). Of course, these views are only incompatible if the goal is any real self-determination at local level. Otherwise, initiatives produced at grassroots level may be exactly what global technology companies need to maintain their position as superpowers, outbidding each other for scalable innovations, and using this relation with the cutting edge of creativity to guarantee their competitive position and their chance to determine how people connect to the objects around them and the financial systems on which we all depend.

So, encouraging a culture of collective innovation can work both at a macro and micro level, despite different end-goals. And innovation can involve a new way of organising a shared garden as much as a software tool, even if the former cannot be taken up and scaled to make a profit. However, neither layer is necessarily supported at the level of nations, which are more concerned with balancing market provision and expenditure in answering to voters asking about the pension deficit.

To support economic vitality through sharing, the sharing tools of the digital commons are an obvious starting point, though we have not discussed them in this report. However, physical resources are increasingly relevant, as innovations include digitally-enhanced objects and smart homes and cities. A key point is that the tools relevant in generating vitality are those that support entrepreneurs developing new forms of infrastructure and exchange as well as supporting the practices enabled by them. Conceiving tools to inspire and support collaborative initiative takes us out of Orsi's categorisation system into a consideration of how any or all of the four degrees of sharing are supported. The functions for this include crowdfunding, networking, designing, finding ways of inspiring serendipitous encounters between people, and between people and things, easing red tape and mobilising communities. The story of the Ivy House pub and its move to community asset is a tale of luck and enterprise. But outstanding in it is the need they had to mobilise fast and effectively, adapting to changes in the original intentions for the pub, the law and the statutory regime at local level and communicating appropriately to drinkers, suppliers and well-wishers. We are reminded of Ostrom's point, in looking at the mechanisms of co-production (1996), that to do things together, one needs to ensure that a critical mass of people know what is going on and are available to decide on options together.

Thus, vitality can also be understood in units of time and enthusiasm. Although neither is financial, both are connected in ways that relate to sustainability. The initiatives in Brockley run on very little money, making use of in-kind support, currencies of time and so on, but they still need resourcing and they still have to operate within a value system that expects money for land, materials and work. What can support responsiveness is to pay attention to the particular ad-hoc qualities of this informal designing, such as the collaboration, context, interaction and learning that are often ignored (Dorst 2008).

Last, promoting vitality in its most beneficent form involves challenging the tendency for those that already have resources, confidence and skills to be able to take advantage of opportunities and for those that are already disadvantaged to find themselves even further outside the locus of benefit. This brings in issues of access to resources, including even access to (and confidence with) the digital tools we are discussing, such as smart phones and ubiquitous wifi.

We need people with initiative and imagination and the resourcefulness to use these qualities if we want to promote long-term societal sustainability and this has to be bred by creating opportunities to do so. Digital tools can be an important part of that, but targeting secondary markets is only a small part of that potential. The economy of sharing (saving, managing, growing) is wider than the sharing *economy* and more useful in environmental and social terms.

Having it all?

The strategies above highlight different alternatives but do not wholly preclude development on all fronts. They are artificially constructed to show emphasis and clearly social cohesion is improved by adequate resourcing and a buoyant economy, and vice versa. We can have most of it if we plan carefully, though Benkler warns that you have to protect sharing-based modalities of production if you want to keep them in a price-based, hierarchy-based world (2004). The only argument for choosing one strategy over others is to ensure that no facet of societal development is neglected at the expense of other, more dominant, ones.

Researchers are careful in approaching judgments that affect the validity of the data during collection and analysis, then form opinions based on the evidence they find. Here we advocate design that supports social cohesion, local community development and, with it, the modelling of initiative-taking that makes places to live both pleasant and resilient. Brockley's culture of initiative and support seems to be locally infectious, if still limited to certain groups, and it is no coincidence that one of the team chooses to live there. We advocate this strategy for several reasons. First, some of the narratives of the sharing economy are conflating resource management activities with social cohesion goals, sometimes deliberately as a sweetener to market what is essentially a new form of business that monetises spare capacity; sometimes out of enthusiasm; and, in the media, seemingly because it makes for a better story. This shows the attractiveness of the end state, but does little to create or promote it. Second, the market has been designed to take care of its own and social mechanisms need boosting in an economy where global enterprises are able to map and exploit every possible source of revenue and development opportunity by using the new monitoring potential that comes with service provision. And, third, because in the narratives we tell ourselves about life in the future, only those where we address issues of inclusion and access to resources in a fundamental way seem sufficient to mobilise the human elements of a world facing the extinction of a majority of life forms, the destruction of key natural materials, likely inundation, and so on, to find a way of life worth living.

References

Aigreen, S. (2011) Sharing: Culture and the Economy in the Internet Age, Amsterdam University Press

Asher-Schapiro, A. (2014) Against Sharing, 19 Sept 2014, Jacobin Magazine: https://www.jacobinmag.com/2014/09/against-sharing/

Agyeman, J., McLaren, D. and Schaefer-Borrego, A. Sharing Cities, Friends of the Earth: http://www.foe.co.uk/sites/default/files/downloads/agyeman_sharing_cities.pdf

Belk, R. (2007) Why not share rather than own? The Annals of the American Academy of Political & Social Science 2007, 611-126

Bellotti, V., Carroll, J.M., and Han, K. (2013) Random Acts of Kindness: The Intelligent and Context-Aware Future of Reciprocal Altruism and Community Collaboration. *Proc. CTS 2013,* IEEE (2013), 1-12

Bellotti, V., Cambridge, S., Hoy, K. et al (2014) Towards community-centered support for peer-to-peer service exchange: rethinking the timebanking metaphor, ACM CHI'14, 2975-2984

Benkler, Y. (2004) Sharing Nicely: On Shareable Goods and the Emergence of Sharing as a Modality of Economic Production, Yale Law Journal, 114 (November), 273-358.

Botsman, R. (2013) The Sharing Economy Lacks a Shared Definition: giving meaning to the terms, Collaborative Lab, online presentation: http://www.slideshare.net/CollabLab/shared-def-pptf

Botsman, R. and Rogers, R. (2010) What's Mine is Yours : The rise of collaborative consumption, Harper Business

Boyle, D., Coote, A., Sherwood, C. and Slay, J. (2010). Right Here Right Now: Taking co-production into the mainstream, NESTA

Bramall, R. (2013) The Cultural Politics of Austerity: Past and Present in Austere Times, Palgrave Macmillan

Brereton, M. and Ghelawat, S. (2010) Designing for participation in local social ridesharing networks: grass roots prototyping of IT systems. PDC'10, 199-202

Bria, F., Almirall, E., Baeck, P., Halpin, H., Kingsbury, J., Kresin, F., van Tongeren, S. and Tait, J. (2014) The Digital Social Innovation report, 2014, European Union, D4 Second Interim Study Report

Brynjolfsson, E. and McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, New York, W. W. Norton & Co.

Cahn, E. (2009) It's the Core Economy stupid: An Open Letter to the Non-Profit Community: http://timebanks.org/wp-content/uploads/2014/01/CoreEconomyOp-Ed_001.pdf

Cawson, A. (2013) The End of Neoliberalism: A public lecture by Professor Alan Cawson, given to the Political Science Department at the University of Crete on 22 October 2013: https://www.youtube.com/watch?v=VzgT6suUopQ

Cocozza, P. (2014) A day in the life of a London Boris bike, The Guardian, 26 July 2014: http://www.theguardian.com/lifeandstyle/2014/jul/26/day-in-life-boris-bike

Coote, A. and Goodwin, N. (2010). The Great Transition: Social justice and the Core Economy, London, New Economics Foundation

Corkery, M. and Silver-Greenberg, J. (2014) Miss a Payment? Good Luck Moving That Car, New York Times. 24 Sept 2014: http://dealbook.nytimes.com/2014/09/24/miss-a-payment-good-luck-moving-that-car/

Crompton, T. (2010) Common Cause: The Case for Working with our Cultural Values, WWF-UK

Dorst, K. (2008). Design research: a revolution-waiting-to-happen, Design Studies 29, 4-11

Edelman, B. and Luca, M. (2014) Digital Discrimination: The Case of Airbnb.com, Working Paper 14-054, Harvard Business School: http://www.hbs.edu/faculty/Publication%20Files/14-054_e3c04a43-c0cf-4ed8-91bf-cb0ea4ba59c6.pdf

Evans, P. (1996) Government Action, Social Capital and Development: Reviewing the Evidence on Synergy, World Development 24(6), 1119-1132

Gauntlett, D. (2011) Making is Connecting, Polity Press

Grice, H. P. (1975) Logic and conversation, in (eds Cole, P and Morgan, J.L.) Syntax and semantics: Speech acts, New York, Academic Press, 41–58

Hamblen, M. (2014) Verizon Auto Share service to launch later this year, Computerworld, 8 Sept 2014: http://www.computerworld.com/article/2602872/verizon-auto-share-service-to-launch-later-this-year.html

Heise, U. (2008) Sense of Place and Sense of Planet: The Environmental Imagination of the Global, Oxford University Press

Heitlinger, S., Bryan-Kinns, N. and Jefferies, J. (2014) The Talking Plants: An Interactive System for Grassroots Urban Food-Growing Communities, Proc. ACM CHI'14

Hester, R. (1984) Labours of love: a model for community design, in (eds Langdon & Cross, N.) Design and Society, Design Council, 138-142

Hipp, J. R. and Perrin, A. (2006) Nested Loyalties: Local Networks' Effects on Neighbourhood and Community Cohesion, Urban Studies, Vol. 43, No. 13, 2503–2523

Hollis L., (2014) The Heart of Urban Resilience is Trust, Not Technology, Shareable.net blog 2 April 2014: http://www.shareable.net/blog/the-heart-of-urban-resilience-is-trust-not-technology

JRF (2014) Practical action to build community resilience, Joseph Rowntree Foundation, March 2014: http://www.york.ac.uk/media/sei/documents/publications/community-sustainability-environment-summary.pdf

Kessler S., (2014) Pixel and Dimed on Not getting by in the gig economy. Fastcompany.com: http://www.fastcompany.com/3027355/pixel-and-dimed-on-not-getting-by-in-the-gig-economy

Kiessling, T. (2013) Culture of Sharing to Expand the Internet of Things, Wired Innovation Insights, October 23, 2013: http://insights.wired.com/profiles/blogs/connect-and-share-the-network-of-things-is-expanding

Light, A. and Miskelly, C. (2008) Brokering between Heads and Hearts: an analysis of designing for social change, in Proc. DRS 2008, Sheffield, July 2008

Light, A., Miskelly, C. and Thompson, S. (2008) An analysis of building habitat with networked tools, Proc. ACM OzCHI'08, 180-187

Light, A. (2011) Digital interdependence and how to design for it, interactions, 18 (2), March + April 2011, ACM, New York, NY

Light, A. and Akama, Y. (2104) Structuring Future Social Relations: The Politics of Care in Participatory Practice, Proc. ACM PDC 2014

Louridas, P. (1999) Design as Bricolage: Anthropology Meets Design Thinking, Design Studies 20 (6), 517-535

Makwana R. (2013) Values and the Sharing Economy, Shareable.net blog, 28 February 2013: http://www.shareable.net/blog/values-and-the-sharing-economy

Morozov, E. (2014) Don't believe the hype, the 'sharing economy' masks a failing economy, The Observer, 28 September 2014: http://www.theguardian.com/commentisfree/2014/sep/28/sharing-economy-internet-hype-benefits-overstated-evgeny-morozov

Mortberg, C. and Stuedahl, D. (2010) Why Do The Orders Go Wrong All The Time? Exploring sustainability in an e-commerce application in Swedish public school kitchens, Information, Communication & Society, 13 (1), 2010, 68–87

Orsi J. (2009) Four Degrees of Sharing. Shareable, 16 Sept 2009: http://www.shareable.net/blog/four-degrees-of-sharing

Orsi, J., Eskandari-Qajar, Y., Weissman, E. et al (2013) Policies for Shareable Cities: A Sharing Economy Policy Primer for Urban Leaders, Shareable and the Sustainable Economies Law Center

Ostrom, E. (1990) Governing the Commons: The Evolution of Institutions for Collective Action, Cambridge University Press

Ostrom, E. (1996) Crossing the Great Divide: Coproduction, Synergy, and Development, World Development 24(6), 1073-1087

Ostrom, E., Parks, R. B., Whitaker, G. P. and Percy, S. L. (1978) The Public Service Production Process: A Framework for Analyzing Police Services, Policy Studies Journal 7(s1), 381.

Phills, J.A ., Deiglmeier, K. and Miller, D.T. (2008). Social Innovation: Rediscovering Social Innovation, Stanford Social Innovation Review: http://www.ssireview.org/articles/entry/rediscovering_social_innovation

Reimerink, L. (2014) What Happens When an Entire Neighborhood Starts Using an App for Borrowing Housewares, Citylab.com, 4 Sept 2014: http://www.citylab.com/tech/2014/09/need-to-borrow-a-cup-of-sugar-theres-an-app-for-that/379655/

Rowe, J. (2013) Our Common Wealth: The Hidden Economy That Makes Everything Else Work, Berrett-Koehler

Rowland, C., Goodman, E., Charlier, M., Light, A. and Lui, A. (2015) Designing Connected Products: UX for the Consumer Internet of Things, O'Reilly Media

Sennett, R. (2013) Together: The Rituals, Pleasures and Politics of Cooperation, Allen Lane

Singer, N. (2014) In the Sharing Economy, Workers Find Both Freedom and Uncertainty. New York Times, 16 August 2014: http://www.nytimes.com/2014/08/17/technology/in-the-sharing-economy-workers-find-both-freedom-and-uncertainty.html?_r=0

Simms, A. (2012) The New Materialism: better not more, NEF Blog, 26 November 2012 http://www.neweconomics.org/blog/entry/the-new-materialism-better-not-more

Simms, A. and Potts, R. (2012) The New Materialism: How our relationship with the material world can change for the better, bread, print & roses, Schumacher College and The Real Press

Slay, J. and Penny, J. (2013) Surviving Austerity, New Economics Foundation: http://www.neweconomics.org/publications/entry/surviving-austerity

Steed, S. (2013) Money and Giving: Do financial incentives deter or encourage co-operative behaviour? , New Economics, Foundation: http://www.neweconomics.org/publications/entry/money-and-giving

Sterling, B. (2014) The Epic Struggle of the Internet of Things, Strelka Press

Stephens, L., Ryan-Collins, J. and Boyle, D. (2008) Co-production: a Manifesto for growing the core economy, London, New Economics Foundation

Stokes, K., Clarence, E., Anderson, L. and Rinne, A. (2014) Making Sense of the Collaborative Economy, NESTA and Collaborative Lab, Sept 2014

Tonkinwise, C. (2012) Sharing Trust: Tasteful Designs of Social Systems, Trust Design Publication 4: Public Trust, Premsela Foundation

Troncoso, S. (2014) Is Sharewashing the new Greenwashing?, 23 May 2014, P2P Foundation blog: http://blog.p2pfoundation.net/is-sharewashing-the-new-greenwashing/2014/05/23

Tsur, M. and Belsky, L. (2014) Why the sharing economy is bad for friendship, 16 March 2014, Forbes http://www.forbes.com/sites/deborahljacobs/2014/03/17/why-the-sharing-economy-is-bad-for-friendship/

Warneken, F. and Tomasello, M. (2008) Extrinsic Rewards Undermine Altruistic Tendencies in 20-Month-Olds, Developmental Psychology 44 (6), 1785–1788

White, R.J. (2009) Explaining why the non-commodified sphere of mutual aid is so pervasive in the advanced economies: Some case study evidence from an English city, Int. J. Sociology and Social Policy 29(9/10) 457-472

Wiggin, T. (2014) Taskrabbit, symbol of sharing economy, pivots focus on home services, inman.com, July 2014 http://www.inman.com/2014/07/10/taskrabbit-symbol-of-sharing-economy-pivots-to-focus-on-home-services/

Yglesias, M. (2013) There Is No "Sharing Economy", Slate.com, 26 December 2013: http://www.slate.com/blogs/moneybox/2013/12/26/myth_of_the_sharing_economy_there_s_no_such_thing.html

Yu, R. (2012) America's new business model: Sharing, USA Today, 16 July 2012 http://usatoday30.usatoday.com/tech/news/story/2012-07-15/social-sharing-economy/56243142/1

"If we want to understand design for sharing, we must balance a study of technological opportunity with analysis of the practice of people enabling sharing and building infrastructure on the ground. These are the people who are taking sharing from the level of serendipitous encounters to a more organised and mobilising series of engagements."

"The value of sharing means that you have access to new resources and they are people resources as well as material resources - so that, for me, sustainability is linked to the combination and you can start being creative. But you have to trust people and break down the barriers."

"You're making yourself a bit worse off because you're giving away some of your possessions, but the overall cumulative thing is that everyone becomes enriched by it: lots of people making a very small sacrifice."

"We've built a site that we hope to be global, but you can only ever build a site for an audience that you know and understand. … you make a connection and you all feel the better for it. There's the practical thing, but I think the emotional connection for community is really important."

Design for Sharing, by Ann Light and Clodagh Miskelly, is the summary report of the eponymous project, 2014

Interaction Design and Architecture(s) Journal - IxD&A, N.24, 2015, pp. 63-80

Setting up and Running a Sharing Service: an Organisational Perspective

Naemi Luckner, Geraldine Fitzpatrick, Katharina Werner, Özge Subasi
Vienna University of Technology (TU Wien)
Argentinierstr. 8/187, 1040 Vienna, Austria
{naemi, katharina, subasi}@igw.tuwien.ac.at, geraldine.fitzpatrick@tuwien.ac.at

Abstract. Enabled by web and mobile technologies, there has been an explosion of interest in the sharing economy and peer-to-peer exchange, with much high profile attention given to monetised exchanges such as in AirBnB and Uber. However there are also many other sharing initiatives, such as time banking, that focus on smaller, more local communities and do not involve monetisation of exchanges. While there is a growing body of literature elaborating participation and motivation in sharing services as well as analysing organisers' roles, little is discussed about the work involved in the day-to-day organisation and management of such services. In this paper we report on an interview study with ten participants discussing five different sharing systems from three different countries. A qualitative thematic analysis of the data points to significant on-going effort reported by all to establish, maintain and grow a service, not only focussing on its practical aspects but also on growing a community and building trust. How they engaged in this practical work though was not so much shaped by the service model (time banking, LETS and so on) but on a complex relationship between their funding model, the service goal and whether it was a top-down or bottom-up initiative. These findings have implications for the design of technical platforms to support services, not just in elaborating a range of possible tasks to be supported but also in where and how it needs to be tailorable to certain needs, how adaptive it is to different service models and how it facilitates monitoring and reporting duties for organisers.

Keywords: sharing service, exchange service, time banking

1 Introduction

Peer-to-peer exchange has existed throughout human history. Recently, with the rise of technological innovations and a growing public interest in alternative economies, technology-supported exchange platforms are thriving. Even though there are many well-known monetary platforms such as Uber (www.uber.com) and AirBnB (www.airbnb.com), many organisations focus on non-monetary services such as time banking. While there is an ever-growing body of work about sharing services, most of it looks at a specific type of sharing service and/or in a specific cultural context. Many authors have looked into sharing services from a members' and organisers' perspective and have given accounts of the roles involved and their responsibilities.

More recently, work has been done regarding the organisers' perspective and motivations that keep them working [1, 2], however, there is little detail on the more practical day to day work involved in setting up, running and maintaining a service. Furthermore, much of this work has taken place looking at a single service model and/or in similar contexts. We want to add to this understanding by looking more closely at the organisational work involved in facilitating a sharing service, as experienced across a range of different types of services and different organisational structures; and to also look at this in different contexts.

For this purpose, we first introduce the different exchange models we looked at in our field studies, describe important aspects of sharing services that are relevant for this paper and how these were addressed in literature in recent years. We then introduce our field studies and research approach. In presenting our findings, we explore the organisational tasks to administer both the back end and front end of a service, as well as mechanisms and strategies employed by the organisations in our field studies, giving an overview of key challenges of organising an exchange/sharing service. A key finding was that the differences between the services in terms of organisational concerns and effort was not so much based on their particular service model (e.g., time banking or not) but on other factors such as their available funding and resources and whether they were established as a top-down or a bottom-up organisation. Finally, we discuss the interrelation of different strategies, and highlight the importance of context and goals in the decision-making process. We conclude the paper with a list of design implications and suggestions on what to contemplate when designing a technical sharing service.

2 Related work

Sharing trends are a worldwide phenomenon and are currently subject to high attention, not only within the sharing community, but also within an international scientific community. There are different models of sharing. While some services, such as time banking or Local Exchange Trading Systems (LETS), are widely known, other models such as cohousing and neighbourhood peer-to-peer sharing networks are less often discussed in the peer-to-peer economy literature. Time banking is a sharing concept that enables participants to exchange services based on a time-based currency, often referred to as time credits [3–6]. Participants invest time helping others and can spend the same amount of time receiving services from other participants. LETS models of exchange are based on an alternative currency called LETS, which can be used in parallel to 'normal' currency. They are often used in local communities during times of recession [4, 7]. Cohousing initiatives [8] are intentional cooperative buildings or housing communities who co-own, plan and manage their buildings collaboratively, and share facilities as well as activities. Neighbourhood peer-to-peer sharing networks are local sharing communities where sharing takes place in the near vicinity of a sharer and favours are exchanged without the expectation of always getting something in return and without any 'valuation' through alternative currencies.

While there are different sharing systems, there are also different goals for creating a sharing service. Cahn [9] has reflected on the values and principles involved in this social economy [10] and summarised them as seeing humans as assets, redefining work, reciprocity and valuing social capital. These values have since been expanded, for example by Gregory [11] adding the goal to increase active citizenship and re-building a sense of community, and by Seyfang [12] to support sustainable consumption measures. Furthermore, sharing services have been researched as a means to facilitate social inclusion [7, 9, 13] and as a way to provide services in time of austerity [14].

A significant amount of work has looked at the sharing side of services, especially from a member's perspective [2, 10, 15–17], for example talking about motivations to join and continue using sharing services. Issues about sharing practices are also discussed, such as lack of trustworthiness [2], availability [3], not recognising skills to offer [18], the preference by members for giving rather than receiving [10, 18] or not wanting to ask others for help [19]. Researchers have also looked at the role of technology for sharing and exchanging, including standalone sharing systems e.g., for time banking [19] and more generic platforms such as Facebook appropriated for sharing [20]. However, while technology can act as an enabler if it is reliable [21, 22] it can also pose an issue for participants if they are not in the habit of using technology [23], or if they don't see the need for regularly recording their exchanges and hence disrupting the time tracking concept [24].

There is also some work on the organisational perspective. Voida et al. [25], Thoits & Hewitt [26] and Kane & Klasnja [27] for example, all talk about volunteering work and its beneficial impacts for those who volunteer; Seyfang [28], Bellotti et al. [21] and Gregory [3] and others discuss different roles of organisers in sharing organisations; Bellotti et al. [1] and Shih et al. [2] illustrate motivations that drive volunteers and organisers. While many of these have focussed on a single service, most commonly time banking, Light & Miskelly [29] evaluate the organisational needs across seven different types of sharing cases, but all within one local area.

Here we also focus on the organisational side of sharing services, building on and complementing this work. While there has been discussion about roles and motivations, there is less on the practical day-to-day work that is entailed in administering and managing sharing services, which we take up as focus in the interview studies to be reported here. Further, while there is some work looking across different types of services, there is an opportunity to do this beyond one local area, and so here we compare across different types of services in different countries and contexts to understand the similarities and differences and why these arise.

3 Interview Studies

To understand organisational and administrative issues on a more general level, we sought to talk with people who were involved in facilitating non-monetised sharing/exchange services. To this end, we recruited ten people from six organisations, covering five different types of exchange systems (time banking, alternative currencies, exchanging neighbourly help, neighbourhood centres, and cohousing), four different types of technical platforms (and two not using any particular technical sharing platform) and three different European countries (Austria, Greece and the United Kingdom). Table 1 provides a short overview of our field study sites, presenting a selection of information relevant for the content of this paper. While these organisations were recruited opportunistically based on our networks, they represent different sharing models and are comparable since they are all connected with facilitating sharing, but are interestingly different as can be seen in Table 1. These differences will be discussed in more detail in Section 4.1. This allows us to step above any particular service model and context and to look more generally at the issues involved in running a sharing or exchange service.

Table 1. Overview of the services in our field studies, using anonymised IDs for each service.

ID	Exchange system	Funding	Staff	Country	Sharing Platform Technology	Existing since	Members
TG	timebank	private donations, raising money	volunteers (no pay)	GR	OSCurrency	2011	1000+
TE	timebank	local council grant	paid	UK	Echo (based on Cyclos)	2014	~210
NH	neighbour-hood help	project grant, co-operation with other organisations	paid (time limit)	AT	in-house	2014	750+
EC	alternative currency, timebank	membership fees	volunteers (partly paid with hours)	AT	Cyclos	1995	~220
NC	neighbour-hood centre	government funded	paid	AT	none	1997	n/a
CB	cohousing	community funded	inhabitants	AT	none	2009	80

"Timebank England" (TE) and "Timebank Greece" (TG) use different technical platforms to support reciprocal or network-based exchanges in a local community using time as a unit of currency. "Neighbourhood help" (NH) is based on a technical platform to enable local neighbourhood networks to ask for and receive help, organise shared events and share information about their local environment. The "exchange circle" (EC) uses a technical platform for exchanges that has elements of both time banking in that they use hours as a way to track the amount of work that has been done and alternative currency in that they created a tangible representation of these hours, can give hours away as a present and facilitate exchanges. Neighbourhood centre (NC) is a government funded organisation attached to a large housing complex with the aim to provide support for the local community and people living in its surroundings by promoting neighbourhood activities or workshops and offering counsel. They also see connecting people as part of their tasks and organise sharing activities among the community members. Cohousing (CB) is a purpose-built shared house with about 80 inhabitants. The founders wanted to create a multicultural, intergenerational community that shares resources and promotes a more sustainable lifestyle.

In total we conducted seven in-depth, semi-structured interviews with ten participants, as introduced in Table 2. Participants included both active and former members of services who were key organisers, managers and volunteers. The data was collected in situ in the respective countries and, where possible, at the locations of the administrative centres. Each interview included one to four interviewees, lasted on average 1,5 hours, and was conducted by one or two researchers, in either English or the native language of the country, of which at least one co-author is fluent.

Table 2. Overview of the interviews and participants.

Interview ID	Interviewee/s (role)
TG.1	1 former volunteer (organiser/expert, mainly responsible for the time bank website)
TE.1	1 paid coordinator (service set up and general overview of the platform activities)
TE.2	1 paid coordinator, 3 volunteers (one of them a former LETS organiser)
NH.1	1 founder (media and online management, strategic overview)
EC.1	2 voluntary organisers (administration and organisation)
NC.1	1 paid social worker (organiser of sharing activities)
CB.1	1 founder (co-financer, co-planner, and co-organiser)

To gain insights into current practices and issues concerning the administration and management of exchange services, the questions probed on how they ran their service delivery, how they managed participant engagement, and what work was involved in doing this. While most services used some form of technical online platform to support their service and negotiate exchanges, we concentrate more on the broader service not just the platform.

The conversations were captured using written notes and audio recordings, which were subsequently transcribed and, where necessary, translated to English. Thereafter, we used an inductive approach of thematic analysis [30] to identify organisational issues and their implications. These issues were first extracted and then thematically clustered to gain an overview of their inter-relations. Henceforth, when we talk about specific platforms, we use the platform ID, e.g., TG, CB, NC and so on; when we quote from the interviews, we use the interview ID, e.g., TG.1, TE.2 and so on.

4 Findings

While our data is drawn from diverse types of services and in different countries, there was a surprising overlap of shared experience. At a general level, all services went through similar phases with similar challenges: to set up, grow, and maintain the service. All entailed a significant amount of back end organisational work, both strategic and practical decision making and practical tasks, across all the phases. They all experienced common challenges in how to access the necessary resources, including funding, infrastructure and people to do this work. There were also common challenges in how to promote and grow the service, both at a philosophical level in terms of how to define and grow the community and its culture and in particular to establish trust, and also at a more practical level in terms of how to encourage people (both as potential volunteers and as members) to translate their stated support for this sort of community engagement to practical action.

However, how each of the organisations specifically experienced these challenges and how they addressed them varied considerably and this depended not so much on the particular model of sharing/exchange they embraced but on a complex interplay of other factors such as whether they received external funding or not, what the particular goals were, and whether the service was a top-down or a bottom-up initiative. This interplay between these factors determined for example, whether they had resources they could rely on or not, whether there were external constraints they had to operate against or report to, and how long a view they could take in their more strategic planning around the service.

In the following, we discuss our findings in more detail to draw out the particular types of work needed across the phases of a sharing/exchange service, and organise these around key themes of getting started, administering the service and promoting service engagement. The case-based examples we present throughout these discussions illustrate some of the specifics of the interplay of the three factors mentioned above in shaping how they approached a task or issue and we will return to these at the end to summarise.

4.1 Getting started

A key difference across the services was whether they started out in a top-down or bottom-up way or as part of a broader community initiative. We saw that certain goals were an important factor of setting up new sharing services. The services in our study were all founded to fulfil a certain need of a local community or organisation.

TE and NH were both *top-down* but in different ways. TE came about as response to a local council who offered a grant for social projects and called for solutions to mobilise community resources and help people meet each other's needs. TE therefore started off with funding to employ a full-time coordinator who was located at the local volunteer centre, which had applied for the grant. NH on the other hand could be described more as a social start-up, started by entrepreneurs who passionately believed in local community building and sharing. They developed a technical 'neighbourhood network' platform and had a strategic plan to roll out and promote their service one district at a time. While they aim eventually to have a commercially viable business across many districts and cities, the service itself is meant to be free to community members and instead local businesses and building owners are targeted for revenue flow. Not being tied to a particular district, they do not have local offices in the districts and rely on local promotion for service uptake rather than the work of a particular local coordinator.

TG and EC on the other hand are examples of services that started in a much more *bottom-up* fashion. Both were initiated by grass-root communities, with the goal of benefitting from untapped resources to support and enrich their local community. TG was founded at the time of an economic crisis to connect locals in a dire situation, foster solidarity and facilitate the exchange of goods and services without the necessity of money (a pattern also noted by Gregory [14]). EC on the other hand was created to provide a sustainable alternative to the existing economy, add value to local communities and show people that their work is valued equally.

NC and CB differ from the above mentioned services in that they are not focussed solely on sharing, but have the goal to *integrate* sharing as part of their philosophy for their broader community-building initiatives. Both are tied much more explicitly to housing complexes, and try to add in, organise and promote sharing practices; to date neither have a specific technological back end for supporting sharing exchanges other than general mailing lists and community notice boards. NC is a local community centre that employs social workers and also recruits volunteers, and that is often approached with offers of goods to give away and share as well as requests from locals who are in need of certain support. CB is a cohousing project that was designed and built by its residents to cater for their wishes of leading a more sustainable life, closely connected with their immediate neighbourhood. The inhabitants of the cohousing complex were hand picked to create an international, intergenerational, diverse community. While NC has a top-down structure as a government funded institution, they hold regular meetings together with volunteers in a more bottom-up fashion. CB is a bottom-up organisation where all residents contributed in a participative design structure to the establishment and on-going management of their building. All inhabitants are organised in working groups to discuss and further develop different parts of their co-existence.

In summary, and interestingly, the actual underlying service exchange model was not such an influencing factor on the patterns of organisational work we saw across the sharing services.

Funding. It is not surprising then that *funding* was also one of the key differentiators among the services and had numerous implications. Most services used a mixture of funding opportunities, but over all, bottom-up services were largely self-funded [TG, EC, CB], whereas top-down services were often funded through (sometimes time-limited) government or local council grants [NH, TE, NC, CB].

In the self-funding model, members of the staff would often contribute personal money when required to [TG, CB] or money was raised through parties and events [TG, EC]. Some exchange services required their participants to pay a membership fee [EC], which was used for organisational purposes. This model depended on the participation of the service's own community, rather than on external funding sources. The contributions were monetary on a monthly or yearly basis, alternative currencies such as hours, or a combination of the two, as explained by one of our informants [EC] *"from the second year on, members also pay 'hours' for administrative work in addition to their membership fee"*.

Another way to acquire funds was to charge transaction fees for business participants, as was proposed by NH. These fees could be settled by the external organisations that cooperated with the exchange service and strived to gain something by contributing, be it access to the participants or promoting their businesses to potential customers. Some services also received start-up funding from government [CB], local council [TE] or project grants [NH]. Some grants were often attached to terms and conditions on which they had to report on regularly, such as additional safeguarding strategies, security checks or insurance for exchange activities. Grant funding was especially attractive for start-up companies with limited options, but could tempt them to adapt their original goal to better fit the project's requirements and was sometimes seen critically [NH.1] *"grants help us now to start-up, but there are many projects that jump from grant to grant and that always adapt their projects a little. I think that is very dangerous since you lose sight of your focus because you only concentrate on getting the next grant"*. Needing to find such follow up funding was certainly an issue for [TE] who were only funded for a start-up phase. Light & Miskelly [29] similarly show that there is a *"trade-off between raising funds and having time for anything else, especially in a voluntary capacity"*.

4.2 The organisational team - tasks and roles

Regardless of how the services were initiated, once started, they needed people who could take on a wide range of tasks and activities to promote, run and grow the services. Literature points to the different types of roles organisers take on in sharing services, e.g., street ambassadors [11] or time brokers [28]. In our interviews we also saw a diverse representation of named roles, necessary to start-up, run and maintain a sharing service, for example community managers [TE, EC], moderators [TG, NH], brokers [TE] and ambassadors [EC, TE]. However, even though everyone seems to have a clear picture of the tasks within their respective service, they tended to use different terms and wording for similar roles, depending on the scope for people in paid or formalised positions or whether people were volunteers and/or took on more than one role within the organisation. Hence, we refrain from talking about the work in the context of specifically defined roles or individuals but rather discuss the key organisational tasks that needed to be done across all services from the organisers' perspective and how our informants dealt with accomplishing these tasks.

An often-observed phenomenon [TG, TE, NH, NC], especially in bottom-up services, was that of a *driving force* among the organising team, as also mentioned by Bellotti et al. [21]. This driving force was usually one person who kept a general overview of what needed to be done at any given time. They got things done, or, if they did not have the necessary expertise, knew whom to ask to do it. An interviewee described the importance of this person as follows [TG.1]: *"there was one person that did everything and she is a very dynamic, very giving person [...] she would make the phone calls, speak to the other people, get people involved, [...] she was really the engine of everything that happened"*.

Apart from someone keeping the general overview, there was usually a group of other *volunteers or employees* who kept the system up and running. This included a lot of work at the 'back end' administration of the system, such as setting up policies and procedures, administering the technology, looking for funding or maintaining the infrastructure. There were also many tasks more visible to the service's participants, such as overseeing activities, communication and exchanges on the platforms; acting as proxies for more vulnerable people such as children, older adults, or people with disabilities who were interested in using a service but needed support doing so [TE]; or promoting the services and informing potential future participants about possible advantages of joining the service, which according to Bellotti et al. [21] is necessary to attract and keep participants. We discuss these in more detail in Section 4.4.

Who actually undertook these tasks in the organisations varied. In TE for example, a single person, the paid coordinator, took overall responsibility for doing most of these, co-opting volunteers as needed, while in EC the tasks were distributed across a number of volunteers. In NC, sharing was only one of many concerns the social worker had in her role and so she played more of an oversight and facilitating role, identifying opportunities and co-opting community members to help as needed.

Across the services, we saw *motivations* of the organisers similar to those reported in Bellotti et al. [1], i.e., “Society/Community/Utopia” reasons, to explain why the people were prepared to put such effort into providing sharing services. We also saw some other more specific and personal motivations, e.g., creating a thriving community of neighbours rather than convenient connections [NH.1]; overcoming personal loneliness [NC.1]; making changes following a change in life circumstances such as divorce or retirement [EC.1, TE.2]; or being personally challenged e.g., by dealing with technological challenges such as data analysis to refine the platform’s strategies of informing participants [TE.1].

In recognising the importance of keeping people, especially volunteers, engaged, one organisation made a point of trying to find out details about their volunteers’ personal goals in order to better cater to their needs and keep up their motivation. An interviewee [NC.1] explained this in an example, stating that they sought to understand their volunteers’ motivations to find appropriate tasks for them to work at. Consequently, volunteers who joined to battle their own loneliness, for example, would be asked to do group activities rather than run errands.

The challenge of not keeping volunteers active and engaged was highlighted by one of the participants in TE.2 who had been a member of a LETS previously but it had to fold because they could not get enough volunteers to take on the administrative work to enable it to run.

4.3 Administering the service

Across all services it was evident that there was significant work involved in ‘making a sharing service work’ (to borrow from Bowers [31]), from the overall processes and policies to the day-to-day tasks in actually administering the service.

‘Back office’ work. Even though we recognised some major differences between bottom-up and top-down services, there were similarities in the administrative structures and rhythms of the organisations. The similarities could mostly be seen in regard to regular meetings and establishing some forms of community etiquette, while the differences were often grounded in reasons of accountability, such as the need for insurance for some externally funded services.

In most cases, no matter if top-down or bottom-up, the organising teams followed a schedule of *regular meetings* to discuss upcoming ‘to-dos’, strategies or just reconnect. These meetings played an important role in letting the services run smoothly and addressing pressing matters. While regular meetings seemed to be of bigger importance to bottom-up services such as TG, where a team of volunteers had to collectively decide about strategies and goals, NC as a top-down organisation also employed this practice to connect with their volunteers regularly and discuss upcoming happenings.

EC was interesting in how they organised their meetings. While they evolved from a bottom-up initiative, their organisation turned out to be very structured in that volunteers worked fixed hours per month, part of it in exchange for service currency. They also manage their volunteers and members through two types of meetings. As described by the interviewee [EC.1], before each monthly meeting that was open for everyone, they would hold an administrative meeting for the core team *"we are meeting for about two hours before the general meeting to discuss what could be renewed [...] it's partly voluntary work, but half of that time is paid in 'hours'"*.

For many organisations there was also work involved in establishing and agreeing to a *working culture* within the administration team, pointing to the importance of trust and transparency of internal processes. We observed very different approaches of team culture and collaboration, often tied up with the structure of the services. Since most top-down service structures were integrated in bigger organisations, they often inherited clear organisational rules of how to work together [TE, NC], sometimes adapting them to more approachable guidelines [NC]. Bottom-up services on the other hand mostly built up their team cultures more organically, and decided on internal strategies to set guidelines. Some interviewees talked about how helpful it was to formulate rules [TG, CB] or recommend guidelines [NC] to fall back on when needed in order to successfully work together. The importance of this is illustrated by an interviewee [TG.1] who expressed his frustration with working in a bottom-up group without such agreed upon guidelines *"I didn't care if someone could vote at their first meeting or not, I just wanted it to be formalised [...] and I ended up feeling that this resulted in hurting the functioning of the group, because there was no transparency on who we were and what we were, what our purpose was"*. This became "*the straw that broke the camel's back*" for him and it resulted in him withdrawing from the platform. CB, on the other hand, undertook decision-making by small working groups, even though the decisions directly influenced the lives of all parties in the building. Hence, they deeply depended on a culture of trust between its organisers who in this case were also inhabitants of their cohousing complex.

While being part of a bigger organisational structure might mean lower overhead in negotiating own processes, it also came at a cost of greater administrative workload in terms of accountability and reporting tasks according to the conditions of accepting external funding. At a minimum, for example, TE's organisers often needed to report on budget matters, service use, e.g., number of exchanges, and so on, in great detail.

A particular 'hidden' administrative overhead for TE was related to *insurance*. It was required as a matter of its agreement with the local council, and as a condition for its funding, that insurance be provided to all participants undertaking activities within the context of the service. An implication of this is that *all* activities needed to be documented in advance on the digital platform in order to be covered by the insurance [TE]: *"So, if you are going to an exchange and […] you are not going to be able to do the exchange exactly [as agreed upon …] you need to update your request or your* *offer before you do that, so that it is covered on the insurance"*. The hidden work was in always having to monitor exchanges taking place and to convince people to put them into the system. This was difficult as participants sometimes found this onerous, as it did not reflect their daily practices and generated additional effort on their part, and they did not understand the implications for TE from the service side of supporting an exchange that wasn't insured. Other services too [TG, NH] struggled with getting people to document all their activities digitally, especially once the participants had built personal relationships with other participants. While this also raises the question about where the boundary lies between participating in an exchange platform and helping out a friend, it also points to the additional back end work required once a platform is part of the picture.

Additional administrative work was also created around organisations needing to be *accountable for the volunteers* in their service and many had some sort of quality assurance/vetting process in place. NC, for example, interviewed their potential volunteers before they could join. These interviews were a way for the organisation to get to know applicants, and also to find out about their goals and wishes and understand what kind of work they wanted to do. They organised supervisions, trainings and excursions for their volunteers, not just as a way of vetting but to keep them involved and motivated. These could be informative events, but were also intended to create a social team spirit and sense of community. For TE, vetting of their volunteers was of great importance because they were often asked to act as proxies for vulnerable members of the community and special safeguarding policies were in place. As TE.1 explained *"... obviously with vulnerable people you want to make sure it is safe"*.

As previously noted, there is also another tranche of back-office work around funding, not just the day-to-day management of budgets but also towards more strategic longer term survival, especially for more top-down services relying on external sponsorship and funding, e.g., needing to apply for new funding grants [e.g., TE] and so on. All of these types of activities entail additional back-office work as part of running and monitoring the service, being accountable, and being sustainable.

Managing the IT. While a key enabler for sharing services, IT does not run itself and also requires significant back end effort to run and maintain. What that effort entails is also tied in with what expertise is available and how IT systems are sourced.

In most of our field studies, technology platforms played a vital part in the organisation. Even EC, which was funded in 1995, switched from providing only printed hard copies of offers and requests, to incorporating an online presence as part of their service (the hard printed copies still exist though). The studied services had different strategies to deal with technology within their projects. While some developed their own platform [NH], others used customisable open source solutions [TG], and others outsourced the development or used existing solutions [TE, EC, NC]. Each approach led to different issues and advantages. NH, for example, had in-house computer expertise, giving them full control of their system so that they could adapt it based on their participants' feedback: *"We caught [a usability issue] at the beginning [… because] at a test shortly before going live, we found out that even though some people liked the feature, all were overburdened with it"*. TE used an existing generic sharing platform solution but had additional work to document and negotiate feature requests with the external developers (e.g., he asked for a bespoke front end and additional functionality to support bespoke processes such as their vetting process and being able to set up different member groups with different permissions). This could be mutually beneficial work though because, as stated by TE.1: "*this is in their [software provider's] interest as they are developing the software for availability to all timebanks and thus are keen to put in the thinking/work now to make the software as bespoke as possible to any timebank.*" His next wish was for a mobile version of the platform to make it more accessible to people in a timely way compared to desktop-web-based solutions, advantages of which are also noted in Bellotti et al. [19].

Other services [TG, EC] did not have access to in-house or external expertise, both suffering from limited technical knowledge among the team. Relying on volunteers instead, they could not respond to bugs or feature requests in a timely or reliable fashion. TG.1, using an open-source sharing platform, illustrated this problem: *"[The platform] had a lot of bugs, a lot of missing functionality. I didn't have that much time [… and] just made a few small bug fixes. And people got kind of frustrated with the platform"*. EC.1 were concerned too and connected the issue to generational age: *"Because we, the core, are […] too out-dated for things like that. And we would wish for more young people [...] I'll not interfere with that, with the homepage, and putting things online and all of that"*.

In principle we know that good design of technical infrastructure can lead to a more systematic management of people and resources [29]; and that both easy communication and smoothness of coordination are crucial elements in the entire exchange process [16]. However, our cases suggest it is a particular challenge for non-monetary, bottom-up services to find reliable volunteers with the right expertise, or to be able to finance external experts, also mentioned by Lampinen et al. [16], and running good computer systems to support the exchange platform can greatly impact the members' as well as organisers' experience of using a service.

4.4 Promoting service engagement in the community

Apart from the practical back end management, a large part of organisers' efforts was directed towards growing community engagement with the service itself; here the issue of top-down versus bottom-up was less influencing and all cases had very similar needs for promoting the service, getting people to participate and catering to their members, especially with a concern for how to build trust within the community. Across our cases we saw a number of different strategies for how to promote community engagement, all with different resourcing implications.

A community is its people and for sharing services to work, there needs to be some critical mass of members (how many people comprise a 'critical mass' is relative for each group). Thus, a key focus of organisational effort across all services was set on how to *attract new participants* as both members and possible volunteers. Organisations had different active promotion strategies. Most commonly used were flyers [NC, TE], business cards [TE], media coverage [TG, TE, NH, EC], social media [CB, TE], events and word-of-mouth [TG, TE, NH, EC, NC], advertisements on public screens in subway stations [NH], as well as workshops [TE, EC]; all interviewees pointed to the considerable effort it took to decide on, design and deploy/organise any strategy. TE for example, organised workshops and events, often in the context of other volunteer-based organisations, such as churches. Their strategy was to lower the barrier to join by letting people register in-situ at the workshop [TE.1]: *"you always intend to do things and the minute you walk out the door it's […] harder, so it's sensible trying get people to sign up at the initial meeting"*. However while the interviewee [TE.1] recognised the value of such events, he also noted: "*I wouldn't say that event organising is one of my biggest strengths. I can do it but it is not a biggest strength. So I would try to either have an employee or even identify a member, who could do that for me*".

The next critical challenge for most of the services was how to translate everyone's enthusiastic belief in this type of service as a 'good idea' [TE.1] into on-going *active participation* as a regular exchanger/sharer. Again a variety of strategies were used to both remind people about the service and to inform them e.g., about activities posted on the website, as well as about upcoming public meetings or events with other service participants [e.g., TE, EC]. Strategies included: sending out regular printed newsletters [EC] and emails listing offers and requests [TE]; sending out more general informative emails [TE, NH]; explicitly phoning potential exchange partners to inform them [EC]; and using physical notice boards in highly populated areas [NH, CB]. NH also had future plans to increase its reach into the neighbourhood by putting screens into houses and building blocks or using existing in-house TV channels as a way to reach all neighbours, especially those who might otherwise be overlooked [NH.1] *"... a neighbourhood consists of all inhabitants of a house […] we think that especially older adults could contribute a lot in a neighbourhood community"*.

We did not hear any one strategy being highlighted as most effective and it was often a case of trial and error, as well as access to the physical resources e.g., to print materials. One interviewee [TE.1] for example, talked about how he is continually evolving strategies to encourage awareness and active engagement. His emailing strategy is one example: he used to send out a lengthy, monthly email to members, but now sends out shorter emails more regularly, making sure to include only a single "call to action" and a related link. He also monitored the effectiveness of the emails using a marketing service, which enabled him to track how many emails were opened and links were clicked, and what sorts of content generated most interest. This again entails work, not just to promote, but also to monitor, review and adjust strategies.

Building trust. Another key challenge mentioned by all participants was how to build trust in the community and the system, also mentioned in [11]. Repeatedly, interviewees talked about the challenge of creating a trustworthy service in which participants feel secure. Being able to trust a community enough to communicate via an online platform was one thing, but inviting another person into your house for a service or exchange was a different matter [NC.1]: *"Only by following the rules will we succeed in creating a secure environment for everyone"*.

One way to build trust was through having visibly *rigorous processes*. We saw this already for vetting volunteers and some also had vetting processes for new members before they are able to engage in exchanges. As part of the registration process at TE for example, new participants had to give the names and contact details of two other people who had known them for more than six months. These people were then contacted by a member of the organisation and asked five standard questions about the participant. Additionally, new members had to undergo a 75-minute training before their accounts were unlocked. Others, such as NH also verified their participants' identity, but took a more low key approach by sending a postcard with a code to the given address to unlock their account. That way they could at least be sure that the participant's contact data was correct. In these ways, both organisers and fellow members could be more confident in who they were dealing with.

Another way to build trust was through more *social strategies*, e.g., inviting new and potential members to meetings or social events [EC]. Face-to-face meetings were regarded as particularly important for creating a trusting environment [TE.2]: *"I think the opportunities to meet each other are useful. [...] People might prefer to exchange when they've seen someone face-to-face"*. An interviewee [EC.1] also addressed differences between city and countryside *"[Exchange culture] works better in the countryside than in [a city]. In the city it is harder, because people feel more insecure and are more withdrawn. [...] Some people don't even know their neighbours"*.

Other services depended on *community self-regulation* to create a trustworthy environment e.g., encouraging 'thank you' messages [TG, EC] or hand-picking participants [CB], while others formulated explicit member guidelines for community behaviour on the platforms and in service exchanges [NH, TE]. An intriguing insight was that even though some interviewees [NH, TE] had reservations about member verification during the sign-up process, and hence a delay for participants in being able to use the system, this seemed to be of low impact on the registration numbers.

Not only trust among members and towards organisers is important, but participants also need to *trust the technology* they are using. Something our interviewees mentioned concerning the technology platforms was the issue of how to ensure a person's privacy while still displaying expertise and possible exchanges online. In NH for example, participants have different "circles", and when posting something they have full control over who sees the post, e.g., only people who live in the same house, or everyone in their close vicinity. Also, when it comes to a person's profile, they could often manage exactly who sees which information. NC.1 mentioned that they sometimes received requests to directly manage personal data, leading to a different type of brokering role: *"sometimes, when someone asks us not to circulate their number, we collect the number of the other interested party and they [the provider] will get in touch later. That usually works well"*.

In summary, all of these approaches to building trust take organisational time, effort and resources. Thus, how to handle trust among participants was not only entangled with the ideological orientation of a service but also with funding, e.g., the formal processes as required by the funding body [TE] and having resources to do things [TG], as well as the people power to be employed towards this goal.

5 Discussion and Conclusion

The ideal of sharing services, and associated notions of community and sustainability and so on, is clearly tapping into something that people consider important, evidenced by the recent growth in such services. However, the reality of actually setting up and running a sharing service is not a trivial matter. While there is a growing body of literature reporting on sharing services, these are often focussed on one type of platform or exchange model e.g., time banking. In this paper we have explored sharing services in six organisations, covering five different types of exchange systems, four different types of technical platforms (and two not using any technical platform) and three different European countries. Regardless of the type of service or general set up, the data here has shown it to be a complex and strategic undertaking, not just to set up but to run and grow the service.

What we see in particular is that the actual underlying service exchange model was not such an influencing factor on the patterns of organisational work across the sharing services. Rather it was whether services emerged top-down or bottom-up or as integrated within a broader service, the service goals, what funding they had and any reporting/accountability requirements that were entailed. For example, both TG and TE operate time-banking models of service exchange, supported by their different technical platforms, but the administrative context for their timebanking services is very different: TE was sponsored by a local council with a start up grant and had access to offices as part of the volunteer centre and a stable externally supported technical platform, whereas TG started as a self-funded grass-root movement in an economically challenged country, using public space or coffee places for their regular, open meetings, and relied on internal expertise to support an open source platform.

Despite this diversity, all services went through similar phases and there was a considerable shared experience of the common challenges around accessing the necessary resources and infrastructure, how to manage the myriad of daily administrative 'back-office' tasks, how to attract and keep members and volunteers and how to promote ongoing active engagement. We can characterise these activities across three time horizons: short-term (e.g., responding to requests, carrying out vetting activities), medium-term (e.g., developing organisational culture policies, monitoring of service engagement) and longer-term concerns (e.g., strategic development, future funding).

The findings also point to other complex inter-relations among these aspects and their implications for day-to-day decision making, particularly in relation to the implications of funding sources and entailed reporting and accountability requirements, e.g.: 'How will registration and verification processes impact the participant's motivation to sign up to a platform?', 'How much technological support is needed to track participants' activities and what are the resource implications?', and 'What are the implications of seeking or accepting funding from outside sources versus being self-reliant?' and so forth.

To illustrate this in more detail, one decision a service needs to make is whether it is free of charge for participants (as in most of our cases) or not (as in EC who charged membership fees). On the one hand, a decision to be free of charge could be perceived as an expression of social values and ideologies behind a service, valuing all participants equally and lowering barriers for people with lower income to join. This can also impact the motivation of organisational team members to engage, especially if they joined out of altruistic reasons. On the other hand, it can have practical implications, such as the need to search elsewhere for funding; a higher number of passive participants because they can just sign up to have a look without being sure they want to participate; and no or limited possibility to employ security checks because of the resources needed to handle this number, which may impact trust. This example shows the complexities entailed in every choice.

Summing up, there is significant ongoing organisational work to set up, run and maintain a service, but how this work happens arises out of situated negotiations in relation to specific contextual factors, such as: goals, funding, service model, start-up model, technology choice, and other available resources; and people are a critical key resource for this work. Here the rhetoric and idealised motivations and philosphies around sharing services meet the practical realities of the day-to-day work in constrained envionments required to run a service.

5.1 Implications for Design

Light & Miskelly [29] as well as Bellotti et al. [19] propose design implications based on their work on peer-to-peer sharing services. While Light & Miskelly analysed different service models, their proposition for technology to support sharing practices is more focussed on suggesting additional platforms to facilitate the organisational work and shifting to supporting shared spaces more than individualised exchanges. Bellotti et al. on the other hand propose very practical design implications, such as creating a more social environment and facilitating the recording process, but their recommendations are specific to time-banking.

During our work on five different sharing services and when seeing how they played out in different contexts we identified a list of opportunities for the design of technical platforms to support sharing more generally. Addressing those needs could result in technology that better caters to the needs of the organiser and volunteer team and facilitates their work on the services.

We saw that each organisation, even those that used the same type of service, e.g., time banking, had different needs depending on factors such as financing, organisational structure and so on. Also, most organisations we talked to grappled with the issue of not having enough technical expertise to easily customise their digital platforms. Hence, these organisations could be supported by creating sharing technology that is *modular and individually tailorable*. For example, platforms could offer ways to define different roles in the system, account for different skill sets and ideally have a diverse variety of exchange models to choose from, all of which a particular service could tailor to their context. Using a modular approach would minimise the techncial skills required compared to bespoke development.

As discussed previously, some services, especially when they were organised top-down and financed by an external funding source, had a need to regularly report on the project's status. To facilitate this task, sharing platforms could provide *ways for monitoring as well as reporting on activities* on the platform. Organisers could for example be supported by the platform through providing them with report templates, standard query sets, or dashboard views of service activities on the platform.

We live in a time, where a range of different devices is equipped with internet capability and people are no longer bound to a certain platform such as a desktop computer. The interviews showed that the organisations from our field studies also either planned to make use of different platforms in the future (e.g., creating a mobile app and putting up public screens in the neighbourhood to promote their service), or had already put this strategy into action (e.g., providing easily accessible material to put up on notice boards, link to and so on). Also literature talks about the use of mobile devices for sharing services [21, 22]. Currently many of these cross-platform and cross-media initiatives take place independent to the sharing platform. There is an opportunity to re-think the technical support of services as running across a modular suite of media and platforms. Additionally, sharing platforms could offer *different interfaces to easily connect with other technological devices* such as public displays and screens. Such a *responsive design approach, catering for different devices* that link across different media and platforms, could help sharing platforms to be better embedded in physical as well as digital spaces, reach a more diverse set of participants and equip organisers and volunteers with better access to their service, also addressing issues of inclusion for those who are not so technically able.

Furthermore, since we saw similar challenges across different platforms, a meta-service, such as a *'sharing about sharing' platform* could support organisers to exchange their experiences and strategies to deal with issues, and would be well in line with the altrusitic philosophy of many organisers and services.

Acknowledgements

We would like to thank our participants from the sharing projects for sharing their knowledge and experiences with us. Thanks as well to Lone Malmborg and our collaborators in the Give&Take project team for interesting discussions on these topics. This work is partially supported by the AAL-JP and FFG project Give&Take (no:841963).

References

1. Bellotti V., Ambard A., Turner D., Gossmann C., Demková K., Carroll J.M.: A Muddle of Models of Motivation For Using Peer-to-Peer Economy Systems CHI '15. pp. 1085–1094 (2015)
2. Shih P.C., Bellotti V., Han K., Carroll J.M.: Unequal Time for Unequal Value: Implications of Differing Motivations for Participation in Timebanking CHI '15. p. 1075-1084 (2015)
3. Gregory L.: Spending Time Locally: The Benefit of Time Banks for Local Economies Local Econ., 24, pp. 323–333 (2009)
4. Seyfang G., Longhurst N.: Growing Green Money? Mapping Community Currencies for Sustainable Development Ecol. Econ., 86, pp. 65–77 (2013)
5. Carroll J.M., Bellotti V.: Creating Value Together : The Emerging Design Space of Peer-to-Peer Currency and Exchange CSCW '15. pp. 1500–1510 (2015)
6. Sotiropoulou I.-C.: Exchange Networks and Parallel Currencies: Theoretical Approaches and the Case of Greece (2012)
7. Seyfang G.: Tackling Social Exclusion with Community Currencies: Learning from LETS to Time Banks Int. J. Community Curr. Res., 6, pp. 1–11 (2002)
8. Sherwood H.: How to Create Happy Communities Through Co-housing, http://www.theguardian.com/cities/2014/nov/21/how-to-create-happy-communities-through-co-housing?CMP=share_btn_fb
9. Cahn E.S.: No More Throw-away People: The Co-production Imperative, Essential Books, Washington, D.C., (2000)

10. Seyfang G.: “With a little help from my friends.” Evaluating Time Banks as a Tool for Community Self-help. Local Econ., 18, pp. 257–264 (2003)
11. Gregory L.: Change Takes Time: Exploring Structural and Development Issues of Time Banking Int. J. Community Curr. Res., 13, pp. 19–32 (2009)
12. Seyfang G.: The New Economics of Sustainable Consumption: Seeds of Change, Palgrave Macmillan (2009)
13. Seyfang G.: Growing Cohesive Communities One Favour at a Time : Social Exclusion, Active Citizenship and Time Banks Int. J. Urban Reg. Res., 27, pp. 699–706 (2003)
14. Gregory L.: Time in Service Design: Exploring the Use of Time Credits to Deliver Social Policies Social Policy Association Conference (2010)
15. Collom E.: The Motivations, Engagement, Satisfaction, Outcomes, and Demographics of Time Bank Participants: Survey Findings from a U.S. System Int. J. Community Curr. Res., 11, pp. 36–83 (2007)
16. Lampinen A., Lehtinen V., Cheshire C., Suhonen E.: Indebtedness and Reciprocity in Local Online Exchange CSCW '13. pp. 661–672 (2013)
17. Collom E.: Motivations and Differential Participation in a Community Currency System: The Dynamics Within a Local Social Movement Organization Sociol. Forum, 26, pp. 144–168 (2011)
18. Ozanne L.K.: Learning to Exchange Time: Benefits and Obstacles to Time Banking Int. J. Community Curr. Res., 14, pp. A1–16 (2010)
19. Bellotti V., Cambridge S., Hoy K., Shih P.C., Handalian L., Han K., Carroll J.M.: Towards Community-Centered Support for Peer-to-Peer Service Exchange: Rethinking the Timebanking Metaphor CHI ’14. pp. 2975–2984 (2014)
20. Ganglbauer E., Fitzpatrick G., Subasi Ö., Güldenpfennig F.: Think Globally, Act Locally: A Case Study of a Free Food Sharing Community and Social Networking CSCW ’14. pp. 911–921 (2014)
21. Bellotti V., Carroll J.M., Han K.: Random acts of Kindness: The Intelligent and Context-Aware Future of Reciprocal Altruism and Community Collaboration International Conference on Collaboration Technologies and Systems (CTS) 2013. vol. 34. pp. 1–12 (2013)
22. Carroll J.M.: Co-production Scenarios for Mobile Time Banking in Dittrich, Y., Burnett, M., Morch, A., and Redmiles, D. (eds.) IS-EUD 2013. pp. 137–152. Springer, Copenhagen, Denmark (2013)
23. Luckner N., Werner K., Subasi Ö., Fitzpatrick G., Rauhala M.: ”But it’s only online!” - Inclusion in Exchange Platforms Designing for Sharing in Local Communities Workshop. CHI ’15. pp. 1–4 (2015)
24. Collom E.: Key Indicators of Time Bank Participation: Using Transaction Data for Evaluation Int. J. Community Curr. Res., 16, pp. 18–29 (2012)
25. Voida A., Yao Z., Korn M.: (Infra)structures of Volunteering CSCW ’15. pp. 1704–1716 (2015)
26. Thoits P.A., Hewitt L.N.: Volunteer Work and Well-being. J. Health Soc. Behav., 42, pp. 115–131 (2001)
27. Kane S.K., Klasnja P. V.: Supporting Volunteer Activities with Mobile Social Software CHI EA ’09, pp. 4567-4572 (2009)
28. Seyfang G.: Harnessing the Potential of the Social Economy? Time Banks and UK Public Policy Int. J. Sociol. Soc. Policy, 26, pp. 430–443 (2006)
29. Light A., Miskelly C.: Design for Sharing, (2014)
30. Braun V., Clarke V.: Using Thematic Analysis in Psychology. Qual. Res. Psychol., 3, pp. 77–101 (2006)
31. Bowers J.: The Work to Make a Network Work: Studying CSCW in Action CSCW ’94. pp. 287–298 (1994)

Think Globally, Act Locally: A Case Study of a Free Food Sharing Community and Social Networking

Eva Ganglbauer, Geraldine Fitzpatrick, Özge Subasi, Florian Güldenpfennig
HCI group, Vienna University of Technology
Argentinierstrasse 8, 1040 Vienna, Austria
{eva.ganglbauer, geraldine.fitzpatrick, oezge.subasi, florian.gueldenpfennig} @tuwien.ac.at

ABSTRACT
Social networking has a long history of supporting communities online. In this paper we are concerned with a specific community that has formed around free food sharing to save food from being wasted. Specifically, Foodsharing.de is a platform that enables consumers, farmers, organizations and retailers to offer and collect food. Associated with this is the Foodsharing Facebook group where broader community discussions take place. We report on a qualitative analysis of the Foodsharing Facebook group to understand its role in emerging and sustaining the community. The Facebook group is a place where the individual values and motives, socio-political discussions and mass media interrelate and create new social patterns through narratives and local community building. We present our findings as interplay between a number of factors: individual, community, and organisational levels; public relations and media, the operational platform Foodsharing.de that enables local communities and the Facebook group where global ideological framing of the community takes place.

Author Keywords
Social networking sites; communities; Facebook; food waste; empowerment; free share economy; activism;

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Sharing food has been a cultural and social practice in communities since ancient times and is re-gaining popularity now. This is because societal developments of overproduction of food in industrialized countries results in half of the food being produced thrown away along the food chain [10]. The advance of web and social technologies offer new possibilities to connect people who want to offer or receive food for free to save it from being wasted. This gives rise to interesting online and offline sharing interactions, as people who meet online have to meet offline to actually hand over the food articles. However little has been researched about the values and ideals of such free share economy [27] communities and how technologies facilitate these online-offline practices.

The focus of the study here is a German community that shares food supported by the platform Foodsharing.de. This platform enables activities such as offering food, accepting the offer, and negotiating a time and place to meet and hand over the food. Complementing this, active discussions around Foodsharing and its community happen at the Foodsharing Facebook page, where the Foodsharing Association posts messages to invite community interactions and discussions. Community members are also allowed to post to the Facebook group. 3242 posts and comments from December 2012 until March 2013 on this Facebook group served as the material for our qualitative thematic analysis. The focus of our analysis was on individual values and the overall picture that makes such a community work. We identified individual values and needs of social, ecological and economic nature as motivations to engage in sharing food. Here public relations and mass media not only played a significant role in promoting the community but also in creating through narratives a new social pattern of sharing food. Public relations also played a role in building local communities, as sharing food requires a critical mass of active participants in a geographically bounded area. We also analyzed interactions between community members and found pro-active appeals and critical awareness of community members. The key contribution of this paper for the CSCW community is how participants use the social media and web platforms to facilitate fluid transitions between online-offline and local-global interactions. We draw out the interplay between individual, community, organisational levels, public relations and media, the operational platform Foodsharing.de that enables local

CSCW'14, February 15 - 19 2014, Baltimore, MD, USA

ACM 978-1-4503-2540-0/14/02…$15.00.
http://dx.doi.org/10.1145/2531602.2531664

communities and the Facebook group where global ideological framing of the community takes place.

RELATED LITERATURE

This paper builds on prior research on food as a social and cultural phenomenon, the phenomenon of food waste, communication and social networking sites, and food communities.

Food and its passage to waste

The main motivation behind the Foodsharing community is to save food from being wasted by giving it to other individuals and institutions for free. In this section we first want to engage with food itself, as food is an inherent part of our lives. Everybody has to eat to survive. Beyond this primary need there is a social and cultural dimension, where food is inherent in the enactment of cultural and social practices. Food is graspable, tangible, visible, and related to living organisms. Evans points to this materiality of food as "susceptible to spoilage, decay, and rapid transformation" that "suggests a vitalism that animates foodstuffs" [7].

If food is wasted somewhere along the food supply chain, it is not just the food itself that is wasted but also the energy that has been invested in growing, nurturing, harvesting, producing, packaging and transporting. According to Gustavvson et al. [9] medium and high income countries are mainly responsible for food waste, with 95-115 kg a year per person, whereas in developing countries it is 6-11 kg a year per person. The majority of food waste happens at the household consumer and retailer level for industrialized countries like Germany, which cause up to 40 percent of all food waste. The remaining 60 percent of waste happens during food production, agriculture, post-harvesting and processing [9]. Besides the ecological effects it is also the ethical implications that accompany food. It is important to emphasize that for individuals, food waste usually occurs unintentionally. For example, the participants in studies of Evans [7] and Ganglbauer et al. [8], in most instances, did not want to waste food. Food was wasted inadvertently, accompanied by ethical as well as economic concerns and because somewhere else in the world 'people suffer from hunger' or 'it just feels wrong to throw food away'.

Food communities

Given the importance and social nature of food in people's everyday lives, it is not surprising that communities around food are formed. Communities serve as a foundation where experiences and knowledge are discussed and reflected upon. There are various communities with food as a common theme and these include alternative food cultures that might be supported by technology [2]. Odom [19] looked at the agricultural aspect of urban food production and how this might be supported by technology, as technology can potentially play a role in tracking food to 'grower management software' or 'garden sensors'. Communities are often built on the idea of sharing, such as the one studied by Parker et al. who looked at the reflective aspects when people share healthy eating ideas through audio-recordings on health [21]. Participants in this study discussed and shared more than healthy eating ideas and also had debates about wider systemic implications of healthy food and 'become advocates of change apart from the tool' [22]. Gross et al [11] talk about Foodmunity, a social networking site that facilitates people meeting together around shared experiences with food. Our Foodsharing community is extended by the notion of actually sharing food articles, which in turn results in discussions and engagement around broader issues we want to present in this paper.

Communities and social networking sites

Individuals are often part of a broader community, where there are different dynamics at work. An integral part of communities is communication, as it is only through communication that the interior values and motives can be exteriorized [9]. It is also communication that enables cooperation and collaboration [14] between members, which is central to a community. Social networking is an umbrella term for the infrastructures where people with similar interests can meet, interact, create, share and exchange information online. Platforms like Facebook provide the possibilities to connect, investigate and network socially to share identities, content and statuses [12]. For people with common interests there are Facebook pages that enable an online space for people who like and follow a certain theme. Facebook users 'liking' the page and participating in it can benefit from being able to socialize, entertain, seek self-status and information [20], contribute and discover [13]. The members of such online communities are often geographically distributed and independent as a common interest acts as social lubricant for the community. Such social networking sites can also be used in local geographical areas to build social relationships, negotiate ways to take collective action and social norms [17], negotiate local events and services, and share information and advice [15]. In this case interactions happen online as well as offline [18].

ABOUT FOODSHARING

Foodsharing.de[1] is a community platform in Germany that enables consumers, farmers, organizations and retailers to offer and collect food articles to save them from being wasted. Sharing food in this community involves no transactions of money and attracts all sorts of participants. Foodsharing therefore is theoretically open to all levels of the food supply chain.

The Foodsharing initiative originated in Cologne when several committed people came together to form an Association. It mainly started around Valentin Thurn, a documentary film-maker. Thurn had created a documentary called "Taste the Waste"[2] that presented the problem of

[1] http://foodsharing.de/

[2] http://www.tastethewaste.com/

food being wasted from different perspectives including farmers, wholesalers, food retailers and consumers. Besides many other vey active members he was key in actually starting the Association and is still part of it. Following the documentary the newly formed Foodsharing Association started a crowdfunding campaign and raised enough money in Germany to get funding for an online platform, giving evidence that there were already many people who believed enough in the idea to invest money for its development. Foodsharing.de was subsequently released on Dec 12th 2012. While it was started in Cologne, the site now can be accessed by anyone anywhere in Germany and actual food sharing could take place wherever people could physically access each other.

Figure 1: Screenshot from Foodsharing.de webpage

As at March 2013[3] the Foodsharing.de community had 17000 active members distributed over Germany and 1788 food baskets had been handed over. On the date we downloaded our data for analysis, there were 271 food baskets on offer to be collected. A food basket is created by someone who has food to offer and can contain one or more food articles. The food basket page of Foodsharing.de displays all currently available food baskets across a map of Germany and also provides the same information (available baskets) in a list view (see Figure 1). Community members can filter on different parameters such as location and adjacency on a map, timeliness of food baskets, content, or ending time for collecting them. If somebody wants to take up the offer of a food basket, s/he can then send a request to the person, organization or institution offering the basket. The offering side in turn can accept or decline a request. If both sides agree, they then negotiate where and when to hand over the food basket in the offline world.

The Foodsharing Facebook community

The Foodsharing webpage also links to a Facebook group called Foodsharing[4], which will be referred to as the Foodsharing Facebook page or just Facebook page in this paper. The Facebook page was started on September 13th 2012, before the platform itself started, as a forum for interested people. Whereas the Foodsharing.de platform is mainly functional, enabling the practical sharing of food, the Facebook page is the place where we can see the emergence of the community itself: where broader community discussions take place, and where members are invited to post, comment or 'like'. This paper will therefore focus on the community interactions on the Foodsharing Facebook page.

METHODS

To understand the mechanisms of this community and the role of the platform the posts of the Facebook community were qualitatively analyzed using inductive thematic analysis [3]. As at March 2013, there were 22405 'likes' (Facebook's mechanism for showing support for a page), and 1012 contributing members. The data set comprised 3242 contributions, made up by 243 posts contributed from the Foodsharing Association, 401 posts by members and 2598 comments to posts. Everybody who 'likes' and is subscribed to the page has access to all the posts and comments.

To create the dataset for analysis, we expanded all posts, starting from the beginning of September 2012 until beginning of March 2013, to make all comments visible, and printed out 208 pages of material. As a first pass in the thematic analysis, two coders made notes on two separate prints-outs, from which they identified and agreed 14 broad types of contributions, such as users requesting help or the Foodsharing Association celebrating a milestone. We then started to look for the deeper themes underlying the posts and repeatedly reviewed the material together to draw out important issues. The themes and codes were further reviewed in a collaborative analysis session with other group members. In this paper we focus on two themes, 'individual values and needs' and the 'emergence of the community'.

From the first analysis we could see that public relations and mass media played important roles in promoting the community and engaging new members. Hence, we analyzed 11 videos that were still available from the 17

[3] As at October 2013 the Foodsharing.de community had 27600 members and 4800 food baskets have been handed over.

[4] https://www.facebook.com/foodsharing.de

video links that were posted to the site. We observed these 11 videos and made notes of the main content topics. We were also looking for themes in these videos to analyse how they contribute to the emergence of the community. One illustrative video that was highly influential (surpassing all others by number of comments, likes and shares) was transcribed and thematically coded to uncover the possible new relations of sub-themes to the emergence of the community.

At our request, the Foodsharing Association also made us 'insight analysts' for this group, which granted us access to aggregated Facebook data and descriptive statistics behind the group. This provided us with information about basic demographic data of the members and to identify posts that were most 'viral', meaning we could see how many people commented, liked or shared a post.

FINDINGS

As background to understanding the Foodsharing Facebook community, we first provide a picture of the age group and gender of people engaging in this group, drawn from an analysis of the aggregated data of the Foodsharing Facebook page.

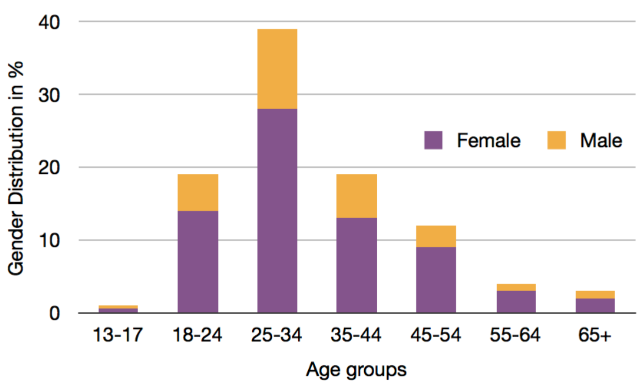


Figure 2: Gender and age distribution within the Foodsharing Facebook community.

More users 'like' the page who identify themselves to be female (69.6 percent) than male (27.4 percent). This resonates with other reports that women often play "central roles in shaping and furthering alternative agrifood movements and institutions" [1, p.12]. Allen and Sachs discuss the ways in which food practices such as cooking are often still a predominantly female domain, with women being mainly responsible for food-related work at the home as well as at the labour market [1]. This could be one aspect why more women (at least on the Facebook page) than men are present.

The most common age group for Foodsharing Facebook 'likers' is between 25 and 34 (39 percent of total). Almost a third (28 percent) of the users in the Foodsharing Facebook group are female and between 25 and 34 years old. See Figure 2 for more details.

We go on here to characterise the posts and comments in terms of topics. There were a plethora of issues discussed by the community, ranging from dumpster diving (freeganism)[5], agriculture, gardening, and everyday practices of food and waste, to sharing experiences offline and online, food waste cooking events and other initiatives.

To determine the nature of the information provided and how users engaged with Foodsharing, we categorised different types of contributions from Foodsharing users:

- *Foodsharing experiences*: Experiences, both positive and negative, from online and offline interactions with other members with whom food was shared.
- *Finding a local community*: Requests to connect to/ find others in a specific local area.
- *Calls for internationalization*: Remarks about wanting to have such a community in their country. (Foodsharing.de is provided for Germany only at the time of analysis).
- *Offering help:* Offers not only to share food but also to engage in voluntary work.
- *Plaudit to the Foodsharing initiative:* Appreciation for the initiative and how useful it is.
- *Wider systemic implications:* Discussions of the systemic effects of Foodsharing and if individual actions might have consequences if Foodsharing gains a critical mass.
- *Links to other initiatives:* The dynamics of Foodsharing seemed to attract users to link to other initiatives with a similar mindset.
- *Everyday practices and food waste:* Discussions about how food waste in private households, food retailers, restaurants or agri-industry emerges or not.
- *Feedback on the design of the Foodsharing.de platform:* Remarks about what could be improved and which features would be desirable.

The Foodsharing Association also made particular types of contributions:

- *Requesting help*: Asking for help on a range of issues, from asking for legal expertise (e.g., lawyers to consult about food legislation and Foodsharing activities) to asking for volunteers to be interviewed and report about Foodsharing experiences on TV.
- *Providing feedback to members:* Giving answers to questions, contributing to discussions.
- *Showcasing Foodsharing:* Sharing links to media where Foodsharing was presented, ranging from reports on TV to newspaper magazines.
- *Promoting networking:* Providing links to other initiatives with a similar mindset or political intention, such as for petitions, events, etc.

[5] The practice of foraging dumpsters for edible food (freeganism) or other goods.

- *Celebrating collective community milestones* such as the thousandth food basket that has been handed over.

Within these contributions, we identified emerging themes following the thematic analysis procedure described in the methods section. Here we start with unpacking the underlying motivations and values of individuals to participate in such a community. Posts may also entail a number of other themes than the one we primarily identify, as topics were often discussed in a non-linear way, where personal experiences are mixed with arguments for political perspectives or general expressions of praise or dispraise for the community. Statements of the members have been translated from colloquial German into English to make it understandable for an international audience. Pseudonyms are used to refer to members.

INDIVIDUAL VALUES AND NEEDS

This community included a highly diverse set of active Foodsharing Facebook members. To enact practices of sharing food requires additional effort in people's everyday lives – to create and/or respond to posts, to negotiate meeting places and times, and to physically meet to exchange the food. This implies that there are motivations and added values beyond the food that is provided to members. We were looking for the motivations that are inherent in the practice of sharing food in such a community. Two underlying and interdependent aspects as incentives to take action in this food sharing community were identified, namely *social and ecological* values and *economical* needs.

Social and ecological values

When people described their experiences on sharing food on Facebook, we often came across statements such as "*gives me a good feeling*" or "*doing a good thing*" in sharing food to save it from being wasted. So there seems to be something in not throwing food away that feels inherently right for these members and gives people an intrinsic reward from being 'socially responsible'.

There were also social side effects and some specific instances where people reported that they built new social relationships through Foodsharing activities. This is because the online interaction on Foodsharing.de leads to people actually meeting up, i.e., while the initial contact is made online via the platform, actually meeting and handing over the food has to happen offline at a place users can decide themselves.

> Isabel: "Even if it [food] was only a small amount I gave away, it gave me a good feeling. I get rid of my food baskets so quickly and you get to know so many brilliant and interesting people. Next week I have an arrangement with one for dinner … it is fun doing good and at the same time making new friends."

For others, the social motivation was more about wishing to "*do good*", to help and support people "*who don't have it so easy*" (Sophie). This social motivation in some cases went even further. Michael solely wanted to offer to people who are in need and not those who are economically well off:

> Michael: "Foodsharing is a great idea, I can finally give left over food to people in need of help. […] I just don't see a point in helping people who are NOT in need of it."

This statement was then discussed with other Foodsharing members, questioning who is in need and how one might be sure that only people in need are receiving food. The discussion, along the dimension of who has or has not economic and social need, points to wider systemic discussions we often encountered in the data when members discussed Foodsharing, an aspect which we will discuss later on.

Given the effort involved in sharing food, it is not surprising to see that people also expressed frustration when members of the community made an appointment to hand over the food but those collecting were not reliable. The Facebook group then acted as a forum for those who had been disappointed by a member who did not show up. It was suggested by members that these happenings should be translated into technological changes insofar that not only members offering food baskets should be rated but also those who collect them.

Besides intrinsic social values that mattered for sharing food, active members also demonstrated ecological motivations. These played out not just in terms of local practices, but connected to broader concerns for societal change.

> Karoline: "I hope even more people become enthusiastic about Foodsharing, at least this would be great for humanity and the environment."

This statement implies the humanistic nature of sharing food and how the social and environmental impacts go together in an idealized account of striving for a better world. This statement also reveals that users think about Foodsharing having systemic effects in a social (humanity) and ecological (environment) way. It is not only about the practice of sharing food on a local micro-level, but, given that more people participate, it will have macro-effects and systemic changes along the food chain. Doris similarly expresses this:

> Doris: "Our resources are limited and we should ALL catch on to this finally".

Such discussions pointed to the importance of getting a critical mass in order to achieve a notable effect on scarce limited resources of the environment. However, these environmentally optimistic posts were often counter-argued by people who pointed out that more than just a critical mass is needed, that it needs interventions from the state to reduce food waste on the agri-industrial side. There were a vast number of posts, particularly during discussions about more systemic and political aspects, that condemned the food industry for resource and food depletion, the state who does nothing against it, and retailers who deliberately prefer to throw away instead of giving to people in need. Those in need refer to our next motivation, which is an economical as well as social one.

Economic need

It is clear from other posts the economic need played a big part in food sharing, given how often people in the 'giving' position commented about people in need. However there was only a minority of posts to the Facebook group from users who were in need of food support to make or improve their living. Of all the 3242 posts there were only five instances where people explicitly articulated their own economic need. This points potentially to feelings of shame that might go with social and economical needs of sharing food and discussing this publicly.

> Tory: "If I go to the food bank twice a week already und cannot give a lot [of food] myself…can I still get any?"
>
> Anna: "Just register online and have a look if somebody has to give something away. I don't think that it is about "who collects also has to give" but rather those who have, give, and those who need, just take."

Though we could not find many posts from users who collect food baskets, the free-rider phenomenon that is criticized in many other communities is actually welcome for the Foodsharing community. Tory is seeking help and asking for the conditions under which she can get food, and is encouraged and supported through Anna, telling her it is endorsed if "*those who need, just take*". Michael's social motivation to share food (noted previously) was even to give only to people who are socially disadvantaged. We could not find at any point a member complaining (at least on Facebook) about people who only seek and not give food.

In summary, in this section we were looking for the motivations that are inherent in the practice of sharing food in such a community. Social, ecological as well as economic values and needs are incentives to take action. The motivations between and within participants are manifold, some emphasizing a general '*doing good*', some writing more about their social, ecological and/or economic motives. Help-seekers, help-givers, social, ethical ecological and economic values and engagements are all able to co-exist and in some cases mutually re-enforce each other. These values are reflected and made visible by the various discussions at the Facebook community. What is interesting to note too is that while the different roles of giving and receiving make this sharing community work, since both roles are needed for any food exchange, the discussions on the Facebook community are largely presented from the activists and giver perspective.

EMERGENCE OF THE COMMUNITY

We can see the emergence of Foodsharing at both local and global levels, with the public media also playing a key role. We define the term 'global' in this paper to have a non-local, geography-independent and issue-based connotation. This section is started with the initial role of the media.

Creating visibility and narrative through media

Public relations and mass media played an important role in the emergence and sustainability of the Foodsharing community. Specifically we focus on the emergence of the Facebook community through the inter-relation between Facebook and public relations and mass media, and how members support local community building through advertising and pro-active appeals. We also focus on the development of critical awareness through community interactions and the emergent narratives that are used to communicate the values and practices of the Foodsharing Association.

As noted previously, Foodsharing started off with a crowdsource campaign to attract funding for the development of Foodsharing.de, at a point where the Facebook group already existed to promote and discuss Foddsharing.de's development. The platform was released on December 12th 2012 with a press conference, accompanied by local strategies such as posters, flyers and billboards close to food retailers. Foodsharing has since had significant media interest, with a very frequent presence on prominent TV news, newspapers and online news. The Foodsharing Facebook page links to: 17 reports about the community on TV channels, 3 of those channels being the biggest in Germany who broadcast about Foodsharing in their main evening news; 44 newspaper articles, with 6 of those being amongst the biggest national newspapers or magazines in Germany; and 3 links to radio entries and 2 mentions on blogs. This mass media coverage served as a starting point for motivating people to get active themselves in Foodsharing and posting this to the Facebook page. The activating potential was visible on the Facebook page with 31 posts where members got to know Foodsharing through a TV report in the main news on a prominent German TV channel and were encouraging about the initiative.

> Pam: "We just watched it on TV, tried it out and classify it as PERFECT! Great idea :D".

Media coverage not only prompted people becoming actively engaged in food sharing, but also to actively talk to other people who had more power and control over the distribution of food and so try to change their instant environment.

> Cora: "Just watched it on TV … it's a great thing … I am working in a big supermarket chain and will talk about it with my boss".

The various responses showed how media coverage could have important effects on awareness with follow-up actions, moving from watching TV reports to actively engaging in the community or their specific local environment. In fact so many people tried to visit the Foodsharing website after one broadcast report about it on one of the major German TV news shows, that the Foodsharing.de page was accessed unexpectedly often and was not reachable for days. This also resulted in numerous posts on the Facebook page remarking on the unattainability of Foodsharing.de.

To understand the nature of these reports with their activating potential, we observed reccuring issues across these reports: Stories started with presenting the problem space of food waste to raise awareness, presented Foodsharing.de as an alternative, and showed role-playing and exemplifying how the platform can be used. To understand the way these TV reports encouraged interested members, we analysed the narrative content of an illustrative example of the TV report that Facebook users 'shared' the most (most viral) on Facebook; 29.3 percent of all users who saw the post with the TV report also reacted (liked, shared or commented) to it.

The TV report starts off with the overall story of wasted food and introduces Foodsharing.de as a platform that enables individuals to waste less through the new evolving practice of meeting online to share offline. The main narrative behind the report told an individual story to "*show how it works*" (speaker announcing the report). The report then depicted a woman to show online interactions with the platform Foodsharing.de as well as offline interactions when another woman comes by with her children to collect the offered food. She says:

> "One has to bring the right attitude to this, others would throw it away, and that you can accept Foodsharing confidently and don't have to feel weird doing this."

The content of this TV report serves to attract people who want to engage in food sharing and directly addresses stigmas that could potentially be attached to it. This is illustrated in the remarks "*don't have to feel weird about doing this*" and "*accept Foodsharing confidently*" about collecting food baskets. That such social stigmas might otherwise exist is suggested by the fact that only five people posting to the Facebook page identify themselves as being in economic need (see section economic needs). The TV report also serves to practically demonstrate sharing food and how this constructs and narrates a new social pattern, potentially aiming to achieve cultural change ("*bringing the right attitude to this*") for TV audiences. The moderator's words to announce the report, "*show how it works*", point to the demonstration and play-acting of the new social pattern.

In sum, the values of Foodsharing and its social patterns are narrated and exteriorized through broadcasting and at the same time promote acceptance of it. The Facebook page played a critical role in a) keeping the report alive by linking to it and b) keeping the issue alive and being able to mobilise the energy, concerns and debates arising from the report by providing a focal point for people to gather, discuss and learn.

Building local communities

While the media played a role in making Foodsharing an accepted social pattern globally in a cultural, societal, and political way, local communities are still needed to make Foodsharing productive and sustainable. Active agents enabled through the global Facebook page promoted local community building and Foodsharing members engaged in broadcasting, advertising strategies and local interactions.

> Dora: "Is there a possibility to advertise for Foodsharing in your own town? Ideas anybody? It just works if enough people participate…"
>
> Carla: "It depends how you imagine advertising, you can have advertising material sent to you, I have done that too and was sent posters, stickers and flyers in different sizes. There is an email address I forgot that you can write to"
>
> Foodsharing: "info@foodsharing.de"

This conversation shows that to make Foodsharing reach enough people it is inherently dependent on local communities and pro-actively engaged members such as Dora. Therefore public relations and the organisational means of a community play a crucial role on a local level to advertise offline with flyers and posters, both of which can be ordered from and sent out by the Association. This conversation also points to online-offline interactions in advertising online and offline, similar to the practice of sharing food itself. Building and instantiating local communities are necessary to make the community as a whole work, to reach a critical mass, as Dora remarked. Such interactions on the global Foodsharing Facebook page resulted in 17 new local Foodsharing Facebook groups being founded to enable local interactions and food sharing.

Online support in the community did not necessarily come from the Foodsharing Association itself and often happened between members. Support often takes the form of encouraging statements or pro-active appeals.

> Sandra: "I really love the idea, unfortunately there are no food articles provided in my city."
>
> Kathy: "Sandra, offer food articles yourself, mine have been requested and collected within minutes! If everybody just waits until others are offering food articles, it will not work."

There were many other instances where community members supported each other. Examples include: advice on how to use the platform; how to initiate a local community; and discussing relevant food topics such as how to start dumpster diving.

Pro-active appeals were not only exchanged between members of the community to form bottom-up local Foodsharing communities, but also from the active agents behind the Foodsharing Association in a top-down manner. Hence it is also the pro-active involvement of the Foodsharing Association that matters, such as in Dora's and Carla's case where the exchange of information is accompanied by someone from the Association actively providing information.

The Association also repeatedly posted pro-active top-down appeals for engagement to the Facebook community members to offer foodstuffs which members not need any more.

> Foodsharing: "Dear Friends, holidays are coming soon and everybody still has foodstuffs at home that could be offered on Foodsharing.de before they spoil. Take part and comb through your pantry, fridge and kitchen. Now is the right time! Over 15.000 Foodsharing members are waiting for your food basket."

The role of the remark "*over 15.000 Foodsharing members are waiting for you*" is to communicate to interested members that they are part of a bigger movement and story, that another 15000 active people have already adopted the practice of sharing food. This pro-active call also provides direct instructions ("*comb through your pantry*") to engage potential or existing members in sharing food.

Creating critical 'global' awareness

Accompanying the emergence of the community was also the development of critical awareness at a more general level, through which people developed a critical understanding of the socio-political sphere their community moves in. Members regularly engaged with topics on the Facebook page that were actively discussed, questioned and negotiated, such as hunger in the world and the context of wasted food, genetically modified organisms, the role of marketing at food retailers, product packaging or practices of the agri-industry, etc. The vast numbers of topics, though not directly associated with Foodsharing, were actively discussed and provide evidence that food practices are inherently cultural and political.

Discussions also allowed members to develop critical awareness towards potential systemic impacts of Foodsharing and often took place within the context of discussed topics. An illustrative instance for the ongoing development of critical awareness was Tom and Hannah discussing the wider systemic impacts of Foodsharing:

> Tom: "Foodsharing cannot change the throw-away practices of agriculture and industry. Foodsharing can also not contribute to reduce hunger in the world. Foodsharing should then only communicate what it can do: Saving food at the consumer level. Not more, not less."

> Hannah: "But Foodsharing connects people with each other – and this is the basis for all other changes, because enterprises will not change their strategies voluntarily, together we are strong. Foodsharing raises awareness, and awareness is the key."

Here Tom and Hannah discuss their individual belief of what Foodsharing can achieve, where Tom questions and negates wider systemic implications, and Hannah argues that people together can achieve change though raising awareness and collective action ("*together we are strong*"). It is this interaction between members - where they provide different critical perspectives, the debate between them - which potentially contributes to raising critical awareness of individual members. The Facebook page offered the medium through which people could engage critically with the community, its purpose, its aims, its attitude, its technologies and systemic consequences. Through this critical process people acquire a greater understanding of the cultural and social circumstances that shape their lives.

DISCUSSION

In this paper we have been concerned with understanding an activist community around the issue of food sharing and its use of social networking, drawing particularly on contributions from the first 19 months of its Facebook page. Overall, what is impressive across the data is how quickly this community grew over a short period of time and, by definition, how engaged so many people needed to be. The very emergence of Foodsharing as a grassroots initiative and the growing levels of activity both on Foodsharing.de in the food baskets exchanged, and on the Facebook page in the number of 'likes' and the active contributions and discussions, give evidence of people feeling and being empowered to act. Empowerment links to levels of individual, community and organisational empowerment [23]. The Facebook group has been a key focal point and enabler at the levels of the individual, community and organisation (here, the Association) as well, along with key roles of Foodsharing.de and public media, in enabling the emergence of a community that engaged in intertwined 'global' thinking and local acting.

Think globally, act locally

As for many communities the principle of '*think globally, act locally*' is a valid description of the interactions between individuals, the community, the Association and across the data there were strong patterns of *global-local* as well as *online-offline* interactions.

The Facebook page, representing the online world, provided the basis to form global identities and ideas that guide and frame this community. Various discussions, links to similar interventions, the links to mass media where Foodsharing is portrayed, all act as *ideological framing* processes for the community. The Facebook page also has a global-local dimension as people there connect with each other to build new local communities, which resulted in 17 new local Facebook groups. The platform Foodsharing.de acted on a national level and is provided for all Germany as a functional and operational tool. But there is also a strong local element that is the lifeblood of Foodsharing. First users search, request, accept and meet online to negotiate where and when to meet. It is then at the offline local place where the act of handing over of food takes place.

Apart from the overarching patterns of global–local, and offline–online, we could observe interactions to be *top-down* as well *as bottom-up*. This was most visible for e.g. local community building which was a bottom-up approach by engaged community members, at instances supported top-down by Foodsharing through providing advertising material. Pro-active appeals were used as encouragement between members as well as top-down by the Foodsharing Association as posts to all members of the Facebook community. This was, for example, the case at the instances where they pro-actively promoted Foodsharing with mass media. Mass media presented narratives of *new social patterns* how the community works and direct instructions and encouragements to engage in sharing food. Thus actions happen at multiple levels, by local agents as well as the Association.

Individuals of the community enact through social, ecological and economic motives, to save food from being wasted. Through the social networking platforms, they are enabled to translate needs, values, and 'good intentions' (around eco-beliefs, concern for environment, social good, etc.) into practical offers or collection of food. They are also able to connect with others to exchange food in local geographical areas. It was also the belief in a wider systemic change and being part in a bigger intervention that encouraged people to participate. The relationship for individuals to the narratives told via various (mass) media was, according to our data, very influential. Not only the guiding values of the Foodsharing community were shown, but also how the new social pattern of meeting online to share food offline is played out in very explicit and explanatory ways. This had impact on individual encouragement to pro-actively engage in the community. It is the Facebook page that makes this engagement visible and notions of *empowerment*, the process of being motivated to act [23] were visible - such as for Michael who reported to "*finally give left over food to people in need of help*" or Cora who watched a report about Foodsharing on TV and felt encouraged to talk to her boss in the supermarket to actively change her environment. Individuals saw stories where they could identify themselves as part of a bigger movement able to change their circumstances that empower them to act. This was also discussed by Dimond et al. [4] who described the positive impact of collaborative storytelling online. Overall it is the individuals that can realize an emerging and sustained community only, individuals that need to feel agency to change their respective environments according to their values, needs and beliefs.

The **community** itself lives and is enlivened by the various interactions between individuals that fulfil different roles. Mutual understanding, helping behaviours between and within community members, engaged voluntary action, and receiving help add up to collective problem solving. The Facebook group of the community acts as a forum for direct encouragement (pro-active appeals) to act, to post questions, find answers, get support, connect to others and being pointed to most relevant resources in a just-in-time way by other people in the community responding to questions and comments. This is accompanied by tensions and hot debates about political and cultural implications of food and waste practices that characterise this community. Interactions between community members can shape the nature of debates and support the development of critical awareness. Members discuss wider possible or non-possible systemic change through the community or question the systemic impacts of Foodsharing, as illustrated in the conversation between Tom and Hannah. Moreover the Facebook page provided a platform for people to form a community of interest, passion and activism around the issue of food waste and sharing food. It enabled people to mobilise and to act as a 'global-issue-based' community, to seed new local communities, while Foodsharing.de enabled people to form a local community of practical action to hand over food between the members.

The Foodsharing Association provided **organisational** means and technological resources to enable the emergence of this community. They provided the development and maintenance of the operational platform Foodsharing.de that made this free food sharing community possible. Through the Facebook group they were able to provide the information, materials, resources and respond directly to people, to point them to these resources, contribute to conversations, discussions through their posts, and make more powerful use of public media by linking stories through to Facebook. They could effect change, both by empowering individuals to act locally and form local food exchange groups, and empowering people more generally, even if they didn't have a local group, to change thinking, to be more aware and to act politically through giving information, stimulating discussions, pro-active appeals and establishing public discourses.

From a design perspective, there were two key characteristics that made this particular type of local-global, online-offline community work. Firstly, the platform Foodsharing.de enabled practical and operational local community exchanges. Secondly, the Facebook page facilitated broader discussions and framing processes for the community. These sites of inter-dependent global interactions then also facilitated the development of local interactions and communities, which is a crucial aspect for distributing power to local agents of change, such as the instances where local communities were built and advertised by local agents to initiate their own local Foodsharing Facebook page.

It is crucial to acknowledge for CSCW and HCI researchers who aim to design for communities that the foodsharing Association itself was the main actor who intervened and was visible as a change actor beyond the provided platforms Foodsharing.de and Facebook. The community would not have worked without their guiding initiative, help, support and intervention. However the community is still growing because different roles are offered to individuals that they can fulfil and act out. The effect of media also adds to a holistic perspective on how this community could emerge. Community policy and framing also includes rather than excludes political issues and values that are [22] or should be addressed [5, 6, 16].

Moreover the Foodsharing community is a very specific one as the nature of food and food waste is a tangible, graspable and visible one, as opposed to the invisibility of electricity or water [26], resources that are implicitly, naturally, and invisibly delivered to the household. For these reasons, some of the dimensions of local-global, online-offline [18] might be inherently part of this type of free share economy [27] community.

Limitations of our study

While this study had access to a large number of posts starting from the beginning of the Facebook group of the Foodsharing community, the findings might not reflect all members. This is because we only had material about those members who post to Facebook. We could not include the voices of members who are using Foodsharing.de but do not engage with the Facebook group. Conversely, we might also have heard voices that engage with the Facebook group but do not actually engage in food sharing.

CONCLUSION

Foodsharing is a vibrant active community of members engaging in very practical ways at local levels to exchange food, mediated by Foodsharing.de, and in more political and mutually supportive ways at a global level, using the Foodsharing Facebook group. Here we have focused on the role of the Facebook group and explored how the discussions and links provide a means for the values, motivations and growth of the Foodsharing community to play out and evolve. The key contributions are showing how participants use the social media and web platforms to facilitate fluid transitions between online-offline and local-global interactions and the empowering impact of these interactions. We draw out the interplay between individual, community, organisational levels; public relations and media, the operational platform Foodsharing.de that enables local communities and the Facebook group where global ideological framing of the community takes place. The study also points to the relationship between mass media coverage and the follow up public communication on the Facebook page, which proved to be of central importance for establishing cultural change and new institutional practices and social patterns that are oriented towards sustainability and social values.

For future work we plan to interview members of the Foodsharing community who are actively using the operational platform Foodsharing.de to gain more understanding how and why such new social patterns evolve within a community.

ACKNOWLEDGMENTS

Thanks to the Foodsharing Association for providing the aggregated Facebook page data and all the members of Foodsharing who make this community happen.

REFERENCES

1. Allen, P. and Sachs, C. 2007. Women and food chains: The gendered politics of food, *International Journal of Sociology of Food and Agriculture*, 15(1), pp. 1–23.
2. Blevis, E. and Morse, S. 2009. SUSTAINABLY OURS: Food, dude. *interactions* 16, 2 (March 2009), 58-62.
3. Braun, V. and Clarke, V. 2006. Using thematic Analysis in Psychology, *Qualitative Research in Psychology*, 3(2). p. 77-101.
4. Dimond, J. P., Dye, M., Larose, D., and Bruckman, A. S. 2013. Hollaback!: the role of storytelling online in a social movement organization. In Proc. of CSCW '13. ACM, New York, NY, USA, 477-490.
5. DiSalvo, C., Light, A., Hirsch, T., Le Dantec, C. A., Goodman, E., and Hill, K. 2010. HCI, communities and politics. In CHI EA '10. ACM, NY, USA, 3151-3154.
6. Dourish, P. 2010. HCI and environmental sustainability: the politics of design and the design of politics. In Proc. of DIS '10. ACM, NY, USA, 1-10.
7. Evans, D. 2012. Binning, gifting and recovery: the conduits of disposal in household food consumption. *Environment and Planning D: Society and Space*. 30(6) pp. 1123 – 1137.
8. Ganglbauer, E., Fitzpatrick, G., Comber, R. 2013. Negotiating food waste: Using a practice lens to inform design. ACM Trans. Comput.-Hum. Interact. 20, 2, Article 11 (May 2013), 25 pages.
9. Godemann, J., & Michelsen, G. (2011). Sustainability communication: interdisciplinary perspectives and theoretical foundations. Springer Science+ Business Media.
10. Gustavsson, J., Cederberg, C. and Sonesson, U. 2011. Cutting food waste to feed the world global food losses and food waste, Interpack2011.
11. Gross, S., Toombs, A., Wain, J., and Walorski, K. 2011. Foodmunity: designing community interactions over food. In Proc. of CHI EA ’11, New York, NY, USA, 2011, ACM, pp. 1019–1024.
12. Joinson, A. N. 2008. Looking at, looking up or keeping up with people?: motives and use of facebook. In Proc. of CHI '08. ACM, New York, NY, USA, 1027-1036.
13. Karnik, M., Ian Oakley, Jayant Venkatanathan, Tasos Spiliotopoulos, and Valentina Nisi. 2013. Uses & gratifications of a facebook media sharing group. In Proc. of CSCW '13. ACM, New York, NY, USA, 821-826.
14. Lozano, R. (2007). Collaboration as a pathway for sustainability. *Sustainable Development*, 15(6), 370-381.
15. López, C. A. and Butler, B. S. 2013. Consequences of content diversity for online public spaces for local communities. In Proc. of CSCW '13. ACM, New York, NY, USA, 673-682.
16. Le Dantec, C. 2012. Participation and publics: supporting community engagement. In Proc. of CHI '12. ACM, New York, NY, USA, 1351-1360.

17. Lewis, S. and Lewis, D. A. 2012. Examining technology that supports community policing. In Proc. of CHI '12. ACM, New York, NY, USA, 1371-1380.

18. McCully, W., Lampe, C., Sarkar, C., Velasquez, A., and Sreevinasan, A. 2011. Online and offline interactions in online communities. In Proc. of WikiSym '11. ACM, NY, USA, 39-48.

19. Odom, W. 2010. "Mate, we don't need a chip to tell us the soil's dry": opportunities for designing interactive systems to support urban food production. In Proc. of DIS '10. ACM, New York, NY, USA, 232-235.

20. Park, N., Kee, K. F., and Valenzuela, S. (2009). Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *CyberPsychology & Behavior*, 12(6), 729-733.

21. Parker, A., Landry, B. M. and Grinter, R. E. 2010. Characteristics of shared health reflections in a local community. In Proc. of CSCW '10. ACM, New York, NY, USA, 435-444.

22. Parker, A. Kantroo, V., Lee, H. R., Osornio, M., Sharma, M. and Grinter, R. 2012. Health promotion as activism: building community capacity to effect social change, in Proc. of CHI '12, New York, NY, USA, 2012, ACM, pp. 99– 108.

23. Rappaport, J. (1987). Terms of empowerment/exemplars of prevention: Toward a theory for community psychology. *American journal of community* psychology, 15(2), 121-148.

24. Robson, C., Hearst, M., Kau, C and Pierce, J. 2013. Comparing the use of social networking and traditional media channels for promoting citizen science. In Proc. of CSCW '13. ACM, New York, NY, USA, 1463-1468.

25. Shirky, C. Here comes everybody: the power of organizing without organizations, Penguin Group USA, 2008.

26. Strengers, Y. A. A. 2011. Designing eco-feedback systems for everyday life. In Proc. of CHI '11. ACM, New York, NY, USA, 2135-2144.

27. Weitzman, M. L. (1984). *The share economy: Conquering stagflation*. Harvard University Press.

A Muddle of Models of Motivation For Using Peer-to-Peer Economy Systems

Victoria Bellotti[1,3], Alexander Ambard[1,3], Daniel Turner[1], Christina Gossmann[1], Kamila Demková[1] and John M. Carroll[2]

[1]Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA, 94304
bellotti@parc.com

[2]College of Information Sciences and Technology
The Pennsylvania State University
University Park, PA 16802 USA
jcarroll@ist.psu.edu

[3]Jack Baskin School of Computer Engineering
UC Santa Cruz, CA, USA

ABSTRACT

This paper reports on a study of motivations for the use of peer-to-peer or sharing economy services. We interviewed both users and providers of these systems to obtain different perspectives and to determine if providers are matching their system designs to the most important drivers of use. We found that the motivational models implicit in providers' explanations of their systems' designs do not match well with what really seems to motivate users. Providers place great emphasis on idealistic motivations such as creating a better community and increasing sustainability. Users, on the other hand are looking for services that provide what they need whilst increasing value and convenience. We discuss the divergent models of providers and users and offer design implications for peer system providers.

Author Keywords

Peer-to-peer economy, field study, user experience design.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

This paper examines a phenomenon that, some argue, offers a promising new way of getting things done in the world. It is variously called the collaborative economy, collaborative consumption, the sharing economy, and the peer-to-peer economy. These overlapping terms, which we will collectively refer to as the *peer economy*, refer to a diverse range of *peer services*—both for-profit and non-profit—that have two common characteristics. Firstly, in most cases they are enabled by web and mobile technologies, in which power and the production of value are distributed between users. Secondly, efficient utilization of assets including labor, material goods, and knowledge is characteristic of these services [10]. It has been argued that use of such services will help avert a looming global resource depletion calamity by promoting sharing, reuse and sustainability [10,24,27,44].

CHI 2015, April 18 - 23 2015, Seoul, Republic of Korea

ACM 978-1-4503-3145-6/15/04...$15.00
http://dx.doi.org/10.1145/2702123.2702272

While advocates of the peer economy might claim these laudable alternative collectivist motivations as differentiating this new movement from the conventional economy, it is not clear how important they are when weighed against monetary and other self-interested motivations. Indeed there is already a backlash from commentators pointing out that much of the so-called sharing economy appears to have little to do with sharing [11,15,38]. So, what, then, are the real motivations for participation in the sharing economy?

Un-muddling these possible motivations is the goal of this paper. We present a qualitative investigation of peer service users' motives for participating in peer-to-peer transactions, both as providers and as receivers of value. And, unlike prior authors [17,18,34,46,50,51], we have also examined this question from the perspective of the peer service provider. We seek to understand what motivations providers believe their users have and what motivations they are trying to elicit in the design of their systems. Our inquiry took the form of 68 semi-structured interviews of individuals who used or who were founders of or employed by peer services. These services ranged from supporting entirely instrumental transactions—a means to achieve a personal goal, like renting out one's car to earn money—to apparently more purely altruistic transactions, such as giving items away for free. We frame our findings in the light of prior work on motivation, which provides theoretical foundations for our analysis.

BACKGROUND AND RELATED WORK

Theories of Motivation

There is a plethora of theories of motivation [4]. This is not surprising as understanding motivation is critical in many domains including education, the workplace, fitness, health, and marketing. In this section we discuss a selection of more influential theories that are relevant to our study.

Perhaps the most famous theory of motivations is Maslow's hierarchy of needs, represented as a pyramid with 5 or 8 levels, [35,36]. Maslow placed basic physiological needs at the bottom of the pyramid, followed by safety, then love and belonging, then esteem, and finally self-actualization at the top. Additional theories with ordered levels of needs, from basic animal up to more human aspirational, have also been proposed, with three being the most agreed upon number [32]. According to Maslow, higher level needs can only be articulated after basic needs are taken care of. But there is evidence that, in many situations, higher needs can supersede basic needs. [52]. Still, the hierarchy of needs has received support in a recent global survey of subjective well-being [47] and provides a useful framework for management and education guidance [e.g., 19,32,39]. Because of its accounting for the many possible needs that drive human activity, it is quite relevant to peer-to-peer economy participation, which offers a means to satisfy many of those needs ranging from the bottom to the top of the pyramid.

Many theories account for motivations for behavior change (e.g., see [53]). Most of these are concerned with health and wellness behaviors and assume the behavior in question, such as exercise or eating healthily, is challenging for people and depends on certain conditions (for example, coming to believe one's ability). Ajzen's Theory of Planned Behavior (TPB) [1,2] seems to be especially well developed, building on Bandura's theory of self-efficacy [5] and Ajzen and Fishbein's Theory of Reasoned Action [3]. TPB models beliefs and attitudes towards behaviors, acceptance and enforcement of those behaviors as norms, and beliefs about one's own capacity to adopt them as well as intention to do so. In a meta-analysis comparing TPB with other behavior change theories Webb et al [53] found TPB to be the most effective as a foundation for the design of internet-based behavior change interventions. Taking a behavior change perspective on motivation is valid for our work due to the fact that switching to new peer systems instead of conventional economic systems is a type of behavior change.

Self-determination Theory (SDT), another prominent theory of motivation [43], was developed in the context of education. It is founded on the surprising observation that extrinsic rewards such as money or a prize for an activity can 'crowd out' willingness to engage in the activity when not getting rewarded in subsequent trials. [23]. SDT expands on Herzberg's two-factor theory [30], which is generally referred to today as intrinsic and extrinsic motivations. SDT places motivations on a spectrum. On the leftmost extreme is complete lack of motivation. The center of the spectrum represents increasing degrees of internalization of and identification with external values. At the left—extrinsic—end of the center, one is simply responding to rewards and punishments aimed at sustaining values. At the right—intrinsic—end of the center, one has integrated those values into one's own beliefs and value system. On the rightmost extreme is intrinsic motivation, which is fully self-determined, meaning the person finds an activity rewarding in itself and is highly inclined to engage in it. The theory was highly influential in Daniel Pink's book, Drive [41], which applies the lessons from SDT to a variety of settings, especially the workplace. SDT has also been applied with predictive success in education [55], health care [56] and sports [49]. It seems applicable to peer economy participation in that services seem to range from relying on extrinsic motivations (mainly the reward of money) to depending on intrinsic drives such as amusement, curiosity or flow [21] or the simple enjoyment of being able to help others [8].

Reciprocal altruism [48] is a highly specific theory of the drive to repay favors to unrelated individuals. It is not typically counted amongst theories of motivation, and yet it is an undeniably powerful motivator [42], highly relevant to the peer economy. We might expect that participation in peer systems should be at least partially driven by a desire to reciprocate after having received a good or service.

Another relevant theory is Social Exchange Theory (SET) [31] Cook et al [20] provide a review of the importance of SET in the context of modern systems of exchange. The theory makes testable predictions (in the laboratory and in the real world) about the formation of transaction relationships and motivations to engage in exchanges in a network of actors, given variations in power of actors, value of resources, costs, and unpredictability of outcomes from exchanges. It predicts, for example, that an individual, fearing uncertainty, will tolerate significant opportunity costs to stay within an exchange relationship and that high status actors are preferred by high and low status actors as transaction partners. Reciprocity as described above plays a significant role in SET [25,26]. SET predicts that reciprocity, as a prevailing means of social exchange, is less likely than *generalized exchange* to promote solidarity or bonds of trust and affective regard. Generalized exchange is characterized by an indirect, non-reciprocal or pay-it-forward system, just as we see in many peer systems [20].

These theories informed analysis of the data collected in our study and the emergent framework that was used to analyze motivations reported by our participants.

Motivation for Participation in the Peer-to-Peer Economy

Whilst there has been considerable research on motivations for participation in online communities with some peer-to-peer characteristics, there are few publications focusing on motivations for participating in the peer economy itself, where transactions involve physical resources and services.

Suhonen et al. [46] surveyed users of a college campus lending, giving and helping service called Kassi. They found that motivations for use were primarily *altruistic* or inspired by the desire to *reciprocate* and, to a lesser extent, *instrumental*. Socializing was not listed as a motivator. Demotivators were: lack of promising opportunities for exchange, uncertainty about what to offer, and unfamiliarity with the service. Lampinen et al. [34] followed up on this work by interviewing users to reveal how they offset their sense of indebtedness for services received by "(1) offering small tokens of appreciation to exchange partners, (2) understanding and accepting the indirect nature of generalized exchange, (3) managing expectations by framing offers and requests carefully, (4) minimizing efforts needed in exchange processes, and (5) bartering and exchanging for a third party" (this last item meaning finding opportunities for transactions on behalf of friends). A limitation of this work is that it focuses on just one system, Kassi, with a user population largely made up of students.

Collom, [17,18] administered surveys to members of a time bank, finding that members were motivated by (1) personal *needs* and *wants*, i.e., self-serving needs, (2) the *values* of the timebank, and (3) by a feeling of *altruism*. Social motivations were the least important in both of these studies. His results also confirmed Knocke's [33] hypothesis that people participate most in the aspects of a system that address their motives. However, once again, this work is limited to one system, a timebank, and little open-ended, qualitative information was collected.

Van de Glind [50] conducted a qualitative interview study of participation in three peer services; a *Product Service System*, Peerby, in which users made their possessions available for usage or rent by others; a *Redistribution Market*, Shareyourmeal, in which users redistributed food by cooking for others; and a *Collaborative Lifestyle System*, Konnektid, where users could offer their skills for hire by others. He also conducted a quantitative survey of informants' estimates of how likely it was that they would use any of nine services including the three above and motivations for doing so. In his qualitative work, he found the following motivations; practicality (the most dominant), *socializing* (the second most dominant), *curiosity*, *environmentalism* and *finance*. In his quantitative study, the following incentives were statistically supported: *financial*, other people's *recommendations*, social *attitude towards the neighborhood*, *socializing* in general, and *environmentalism*. Overall, peer system users tended to be highly educated millennials, accustomed to using social media and were often early adopters of new systems. The limitation of this work, however, is that it relies on speculation about use rather than actual use of peer systems.

Finally, a recent survey of 2,500 individuals provides a ranking of 'considerations' for participation in the collaborative economy, ([51] p19). *Convenience*, followed by *price* and then *better service quality*, came out top. Again, though, this was a simple survey with fixed questions and predetermined answer options, which means interesting insights could have been missed.

Another limitation of the prior work in general is that, apart from focusing on only a few services, it does not clearly differentiate between the motives of users as receivers of goods and services and users as providers. Nor does it consider the provider's point of view on users' motivations and what motives the provider is trying to appeal to. We set out to address this gap. We also covered a larger and more diverse set of services than most studies, including for- and non-profits, ranging from primarily instrumental to primarily altruistic exchanges.

INTERVIEW STUDY OF MOTIVATIONS FOR PARTICIPATION IN THE PEER ECONOMY

We have been studying the peer economy with the goal of designing technology that will increase participation in productive exchanges, particularly in the domain of timebanking [7,8,12]. Our current work was conceived as a means to learn from the best practices of the peer economy as a whole, which is typically better resourced than timebanking in terms of its ability to create positive user experiences and to deliberately appeal to a diversity of attracting motivations.

Participants

Participants were recruited by various means, including postings on social media, snowball recruiting, and announcements on email distribution lists. We interviewed 45 users (31 female and 14 male) and 23 providers (employees or founders of organizations familiar with the user experience strategy of their company). A total of 43 different peer services were represented. The average age of users was 33.1 (SD 10.2, range 22-78) and amongst them were 10 students, 8 IT workers, 8 designers, 6 scientists, 4 managers, 2 teachers, and a few other occupations (so they resembled the well-educated participants of [50]). We did not collect demographic information from providers.

Method

We devised a semi-structured interview template that we used to collect information about motivations, roughly in line with the various stages of a *'customer journey,'* i.e., through initially being attracted to the peer service to becoming a long-term, fully engaged user and recruiting others. We also asked about problems and obstacles. The user and provider interviews contained slightly differently worded questions covering the same issues. Almost all interviews were conducted remotely using telephone or VoIP (Internet phone) and recorded for transcription. Audio problems made 5 transcripts difficult to transcribe in full, so we took notes from these and transcribed parts for quotation.

Analysis was bottom-up and focused narrowly on collecting statements from transcripts that articulated motivations for participating in peer-to-peer systems. These statements were entered by a team of researchers, one quote per row, into a shared Google Docs spreadsheet. Researchers worked synchronously and asynchronously going through transcripts and initially coded 596 quotes from the 68 interviews as being concerned with motivations. These quotes were then coded in more detail, according to an emergent motivational coding scheme that was informed by the data and prior work on motivation. This scheme will be discussed in more detail in the Analysis of Motivations section below. The coding of each quote was checked and revised as deemed necessary by at least two people. Meetings and 1-on-1 calls among coders were used to increase agreement on coding. Each quote was assigned as many motivations as were expressed (up to six at once), since people often gave multiple reasons for participation in a single quote. Motivations were also coded according to the source's role in the transaction being described by the interviewee. For users there were *user-receiver* (UR) and *user-provider* of resources (UP) roles. For providers there were service *provider* claims about *user* motivations (PU) and service *provider* describing *design* stratagems (PD) to inspire motivations. Some users talked only about the UR role, others talked about both, UR and UP and all providers talked about both PD and PU roles.

RESULTS AND DISCUSSION

In this section, we present findings with discussion and peer system design implications, since our analysis is in large part qualitative. We begin with a review of the coverage and thus representativeness of our data.

Representation

Due to providers' anonymity concerns, we obscure some company names in our analysis. In Figure 1, services are grouped into categories to show the represented distribution across 21 different types of services (43 organizations, including 6 timebanks). Eleven of these categories (shown in black in Figure 1) would fit the description of a *collaborative lifestyle* service like Konnektid described in [50], although this category is extremely diverse. Five (in diagonal hashing) are *product service systems*, like peerby in [50]. Two (in checker board pattern) are *redistribution market systems*, like Shareyourmeal in [50]. The categories shown in grey remain unclassified, being general purpose or advocacy entities.

Unsurprisingly, the two top categories are transportation and accommodation; currently the most commercially successful types of peer-to-peer business, including Uber and AirBnB. Next is 'marketplaces for selling', which may be slightly under-represented, since the popular service eBay fits this category. The general-purpose category includes services where multiple kinds of transactions are either explicitly supported or occur spontaneously in online communications. Craigslist and Nextdoor.com fit this category.

We classified 52 interviews as representing for-profit services and 15 as representing non-profits; one service did not yet have a clear business model. We also classified services as "instrumental" if they provided a financial incentive for being a resource provider (40 interviews) and "altruistic" if they supported transactions for no financial reward (35 interviews). This means that 7 interviews represented services that support both instrumental and altruistic transactions. The two platform providers could support either altruistic or instrumental applications and were classified as both. The other 5 were services that allowed people to offer resources for free or for a fee. Craigslist was classified as self-interest, rather than both, as interviewees who used it only did so for financial transactions. Timebanks were classified as primarily altruistic, even though members receive time dollars for services, since their philosophy is strongly oriented to an altruistic ideology.

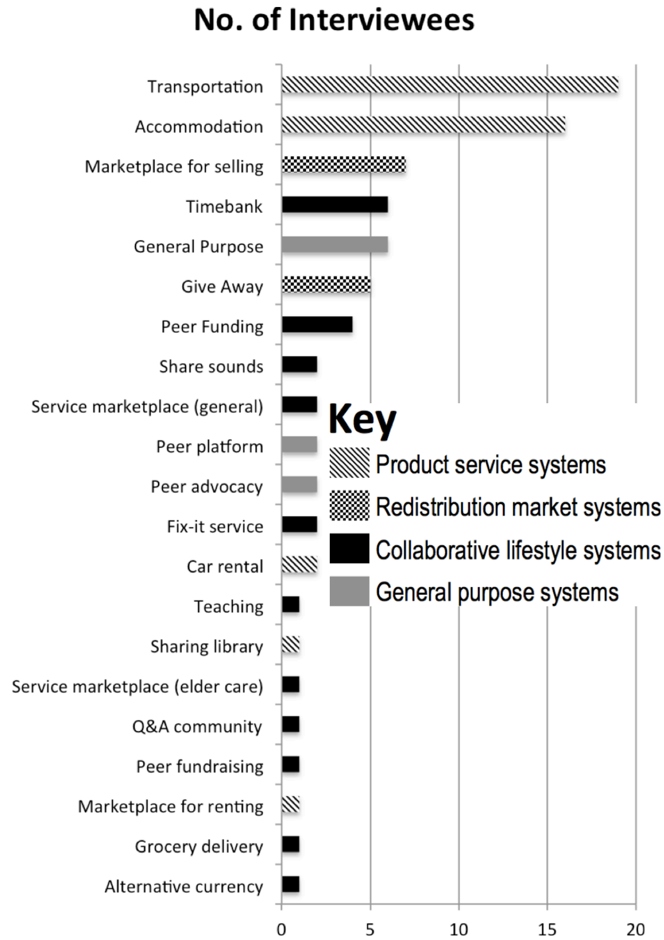


Figure 1. Number of interviewees associated with different categories of service

One challenge we experienced was obtaining interviews with individuals representing the most successful providers. Seven providers worked for somewhat-to-very-famous services. The other 16 represented less well-known or very obscure start-ups, several in beta status.

Analysis of Motivations

After a review of related work and a preliminary review of our interview transcripts, we created a rich motivational framework for data coding (see Table 1, leftmost columns). The framework lists seven theoretically informed psychological roots on the left, driving 21 variations labeled with more common terms for motivation. They are listed (following Maslow's example) roughly from more complex and uniquely human down to basic and simple motivations.

- *Value/Morality* is the drive to improve oneself or the world in terms of some value or principle including justice, equality, and wisdom [57]. Sustainability was broken out of Society/Community/Utopia, because it is often called out as a benefit of the peer economy.
- *Social influence* motives were inferred (following factors in [1,2,3,16]) when someone spoke of a normative motivation (e.g., to act like everyone else), reciprocation

Psychological Root	Common Term	User Receiver (UR)	User Provider (UP)	Provider Claims User Motivations (PU)	Service Provider Designs (PD)
Value/Morality	Self-improvement/Integrity (reduce guilt & hypocrisy)	9%	43%	24%	7%
	Society/Community/Utopia	19%	46%	85%	100%
	Sustainability	4%	23%	9%	13%
Social influence	Norm (internalized influence)	2%	3%	3%	9%
	Reciprocity (internalized morality, fear of disapproval)	1%	9%	3%	0%
	Persuasion (extrinsic influence, pressure from other)	6%	0%	21%	21%
Status/Power	Self-improvement/Increased power (future instrumentality)	7%	23%	12%	7%
	Status/Reputation/Social capital (future instrumentality)	9%	31%	26%	20%
Empathic/Altruistic	Help/Provide service to other	6%	51%	85%	36%
	Give something to other	4%	54%	24%	14%
Social connection	Social connection/Relationship	74%	100%	76%	95%
Intrinsic/Autotelic	Amusement	34%	14%	6%	5%
	Compulsion (e.g. addiction, competition)	0%	3%	3%	0%
	Interest/Engagement/Curiosity/Flow	56%	66%	44%	27%
Safety	Safety	4%	3%	6%	4%
Instrumental	Payment	2%	91%	47%	20%
	Get service/thing	100%	14%	100%	25%
	Increase value/Increase convenience	99%	31%	50%	34%
	Decrease burdens/Decrease inconvenience	18%	51%	21%	4%

Table 1. Relative frequency of code occurrence for motivations across 596 quotes from peer system users and providers. Cells show the occurrence of a code expressed as a percentage of the maximum number of instances of a code recorded for each of the 4 roles

of a resource provided by another or others, or persuasion by a contact (but not by an advertisement).

- *Status/Power* covers drives to succeed in society by gaining power and skill and by advertising such attributes through self-presentation, building reputation and acquisition of social capital over time [14,29,54].
- *Empathic/Altruistic* is based on the human instinct to empathize with others [13], which inspires the desire to assist, by providing a service or by giving something.
- *Social connection* oriented motivation comes from the human drive to start and build social relationships. It is termed 'Belonging and Love Needs' in Maslow's hierarchy [35,36].
- *Intrinsic/Autotelic* motivation, according to self-determination theory [43], is when something is done for its own sake, which could be amusement, compulsion, or interest.
- *Safety* is a basic need in our framework of motivations. It is also listed next to the bottom, above physiological needs in Maslow's hierarchy [35,36].
- *Instrumental* motivation means seeking a practical end, such as receiving a payment (of money or other valuable tokens); getting a service or thing; increasing value over prior solutions; or decreasing burdens related to prior solutions. It would include food, shelter, warmth and so on and roughly equates to Maslow's satisfaction drive at the bottom of his hierarchy of needs [35,36].

The four columns of numbered cells represent the four interviewee roles coded as user-recipient (UR), user-provider (UP), provider-claims-about-user-motives (PU), and provider-design-to-inspire-a-motivation (PD). Rather than presenting absolute numbers of instances of occurrence of each code, we present counts as percentages of the highest count of instances of a code for each column. This allows us to compare to what extent interviewees—acting in each of the roles URs, UPs, PUs, and PDs—talked about motivators, compensating for the fact that we obtained approximately twice as many UR-coded quotes, as any of the other types.

Note that Table 1 summarizes a large *qualitative* analysis. We take the obtained percentages in our data as suggesting *relative* importance to our participants, not as robust statistics. Our use of these tallies is heuristic.

We now discuss the more outstanding features of the data (summarized in Table 1) and their design implications.

Mismatched User and Provider Motivation Models

The biggest finding that Table 1 reveals is a misalignment of the motivations expressed in the different roles, suggesting that roles strongly modulate expressed motivations. In particular the greatest mismatch seems to be between system providers' design intentions (PD) and the self-reported reality for users of peer system services (UR and UP). The most important motivation category system providers talk about designing for seems to be Society/Community/Utopia—if we had not broken Sustainability out as a separate category it would have been by far the most important. Even when we dug deeper and broke out data from for-profits, non-profits, altruistic and non-altruistic services separately, this misalignment *held for all categories.*

Why might this be? We tried to compensate for selection bias in our interview study but must consider that it might be a

Figure 2. Uber advertisement for drivers on the side of a bus

contributing factor. Providers who were willing to talk to us tended to be in smaller, newer start-ups, whereas users were more often engaging the services of the giants in the peer economy. Interestingly, Figure 2 suggests that one giant, Uber (company employees declined interviews), is clear on how to motivate would-be drivers (user-providers). It is possible that only the more socially and altruistically motivated providers agreed to make time for an interview or that social desirability bias [21] influenced their responses.

Nonetheless, this striking diversity of motivations resonates with the current contention between the sharing economy cheerleaders who argue it is about creating a better world [10,24,27,44] and the skeptics who say it is self-interested business as usual [11,15,38]. We will examine the issue of self-centeredness in a later subsection. One interviewee, an AirBnB host, summed this misalignment up very crisply:

Interviewer: *And what interested you most about AirBnB?*
Interviewee: *Money. [laughter] That sounds terrible; AirBnB's whole brand is like, "This is an experience for sharing and togetherness and la-la-la-la-la-la." But I was like, I have bills to pay, and this seems like a pretty good platform to get money. [laughter]*

Providers' utopian motives would be at the top of Maslow's hierarchy [35,36] and on the more internalized end of the SDT spectrum [43]. On the other hand, user-providers are highly motivated by payments and user-receivers are far more interested in getting a service or good and increasing value (getting a better deal) than providers seem to anticipate. These motives are mostly consistent with [17,18,50 and 51] and seem to be related to more basic needs in Maslow's hierarchy; food, drink, shelter, resources, although the hierarchy does not easily map onto things like getting a better deal, for example; *"as someone who enjoys thrift shopping in general, just the concept of finding a good deal on something can be fun at times."* User-receivers' motives are very much on the extrinsic end of the SDT spectrum and very much in line with what social exchange theory would predict, due to users' evaluation of costs and value of peer services versus their competition [20,31]. User-providers' motives are apparently more mixed (see Table 1).

Design Implication: Basics Needs First

If our sample is biased towards less experienced peer system providers, at the very least, the results reported in this paper should help in refocusing their design efforts. Providers need to recognize that, while users may resonate with the ideals of a more egalitarian, helpful, and sustainable world, a user's motivation to use a peer service will often hinge upon its ability to satisfy basic needs—what Maslow called physiological and safety needs [35,36]—more effectively than the competition. Any peer service that is attempting to displace a conventional service must strive to do so more conveniently and at a lower cost to users.

Motivations to Make Social Connections

What users and providers do agree on (as suggested in Table 1) is the importance of the social element of peer systems. This is consistent with one of the previous studies [50] but not the others [17,18,34,46,51], which either do not mention it, or rate it as relatively low in importance. We find that users are attracted to systems where they can connect with other people and forge relationships or simply enjoy the company of others. Indeed, social motivations are the ones most frequently mentioned by user-providers (UPs) overall. But in breakouts we found that the UP social motive dropped off sharply in the for-profit and instrumental services.

System Providers recognize that users are motivated by this and explicitly design their systems to support it, even in the for-profit and instrumental services. As one provider put it; *"I think that there's a lot to be said for having a community of people that have shared values and being able to interact with them."* Another said; *"The whole idea is that neighbors can use it to help each other out, borrow stuff, sell stuff, prevent crime, that kind of stuff, and then ultimately build stronger relationships that help reduce crime and build better community."* There is clear understanding that people prefer to interact with others who share their values and who may be a source of help when needed.

As an example of the user perspective on the same issue, one user of Couchsurfing stated; *"I was also attracted to kind of having a social connection to a place where I had never been before and already know that there were people there who were kind of trying to meet the same kind of people, you know, people traveling. You know, you're going somewhere where you don't know anyone and then realize that there are already people there who want to get together and meet up."*

While building social relationships seems important to peer systems' success, a successful system may also foster social connections that are transient, as a ride-sharing system user described; *"we were driving and then we're talking and discovered that we're hungry and so we just tried […] one restaurant on the way and like have a meal. I had a good fun meal. It's like you don't try to be that, you know, let's become more than friends for something longer term or something. No. It was just good and it was very fun."*

In some services, particularly those that offer accommodation for travelers, social connections may also be a means to enhance the overall value of the service; *"The interesting thing is that you stay with people or that you stay with locals or in locals'... […] You learn about their culture. During* *most of the time they're really interested in you. They're interested in showing you their city and telling you about their culture. And that is for me the most thrilling part because you really get to know people's ethnicity."*

Design Implication: Go Social

Of the theories we reviewed earlier in this paper, only Maslow's hierarchy and social exchange theory (SET) explicitly deal with socialization and only SET makes predictions, related to the instrumental value of exchanges versus the costs of engaging in them [20,31]. This suggests that peer systems should support easy and positive social experiences with clear social benefits for their system users, especially experiences consistent with the needs the service addresses, since those needs are what drew the user to the system in the first place. For example, a service for gardeners who need to exchange plant seeds and cuttings should also support users in making plant-related social connections, communicating about plants, sharing pictures of plants, recommending each other for having certain horticultural skills and so on. Rather than just motivating users with the instrumental value of getting cuttings or seeds, users can be motivated by various kinds of social interactions and these enable access to instrumental and other kinds of benefits (e.g. entertainment and self-improvement through learning).

Self-interested Motivations

As mentioned above, users are motivated to participate in peer systems as receivers by the prospect of the services and things they need, especially if there is added value. These motivators count perhaps even more than social motivators. For some, in some systems, the social element doesn't matter at all compared to getting a good deal:

Interviewee: *I don't see it building more social capital as compared to Craigslist or eBay because basically there's no face-to-face interaction that's why?*

Interviewer: *Right.*

Interviewee: *It's just people coming in to find a better deal than there is on Craigslist. It's not that I'm here to make friends or be proud that I'm part of this community who recycle things.* [Later] *There's not much community feel to this and you see a very small picture of the other person. Yeah. I don't think so. It's really just a market place.*

It could be that the designers of this service have not considered the potential social aspect of their system and this may be a lost opportunity. However, the 'better deals' seem to be sufficient for the user being interviewed. For some, this may even become a compulsion as one user of a system that allowed bidding said, *"I mean if you are much of a gambler, I guess it will give you some sort of thrill to bid, out bid other people. I know some of my friends do enjoy like you can out bid people at the very last minute and then you got the deal, so…"* Notwithstanding this claim, we found little evidence of compulsive deal seeking in general in peer systems.

Increased value may not be based on price alone, but with the extra customer care that peer service providers may offer, as one user put it; *"A lot of my friends […] who are looking for a place temporarily […] they like AirBnB now, even though some might be on the more expensive side. But they'd be like, 'I wouldn't mind doing it for a couple of days, at least, because it's any day better than staying in a hotel or being secluded. Because now a lot of them give you feedback or some people even make you breakfast or take care of basic things,' which is quite nice."*

The user experience is a key part of the value proposition and multiple users mentioned the superior experience of ridesharing services compared with the competing taxi services. Superior value included higher predictability and reliability, better information for both drivers and customers (e.g. GPS visualizations) and superior customer service.

On the user-provider side, for some systems that are trying to be more altruistic in character, where the idea is that people donate items or provide services, credits or time dollars can become powerful self-centered extrinsic [23,30,43] motivators for providers. One user talked about getting a hair dryer from someone working hard for points:

Interviewer: *Why do you think she was so eager to give away her hair dryer?*

Interviewee: *Oh, she wants the credit.*

Another talked about earning *"completely pointless internet points. But nonetheless I wanted to get them."*

In both of these cases, superficially altruistic acts of giving were described—in the former case goods, and in the latter case information—but motivations were also coded as instrumental; in the former case, for the purchasing power of credits and in the latter for status, which entails the promise of future instrumental advantages [29,54]. We coded status as distinct from self-improvement—defined as increasing one's skills or abilities (power). Status was defined specifically as increasing the perception of such assets, by enhancing one's reputation and acquiring social capital. The same user who talked about the pointless points explained what they enabled; *"you gain a lot more privileges as you gain reputation for answers. You get to moderate questions, you can ban users, you can close questions, you can basically participate more in the administration of the site."* The points, in other words, are not pointless at all. Table 1 shows that self-improvement and status are moderate incentives for user providers and again, not emphasized to an equal degree when service providers explained their designs.

Other basic, self-interested motivations include safety (occasionally mentioned in relation to ride sharing and local community services) and intrinsic motivations (*autotelic*: for its own sake). User receivers and providers are drawn to interesting, engaging, and amusing experiences; peer service providers at least seem to understand that users appreciate the engagement part, but design for interest, engagement, and amusement was not apparently as high a priority for providers as it probably should be (see Table 1).

Design Implication: Make it Worth My While

Self-interest is central in social exchange theory, which predicts that humans are mainly driven to transact by this kind of motivation. Prior work also backs up this idea [17,18,50,51]. Given its primacy, service providers should consider self-interested motivations from our framework in Table 1, even if the primary objective of their service is to promote sharing and sustainable resource usage. Self-interest takes many forms other than financial rewards. For example, providers can emphasize a user's ability to learn and then advertise valuable skills through providing a service, or advertise the delights of having more space at home by giving away unwanted clutter or sharing infrequently used tools. On the other hand, providers should always be aware of the possibility of crowding out effects [43] on those who might otherwise be intrinsically motivated to participate.

Altruism and Morality

Service providers' top design priority appears to be powerfully driven by societal or utopian motives and they seemed to assume their users would be too (see Table 1). For example, when asked about what was driving his service, one said, *"... just sort of reaction to the consumer culture that has become so prevalent in our society and a desire to fight the power, if you will."* Another said, *"We are about building the community right!"* Yet another said, *"We believe in using [anon] as a way of promoting the advancement of social justice for all people in the [anon] county."* These motives clearly belong at the top of Maslow's hierarchy, being related to morality and problem solving.

Although providers talked a lot about their users helping each other (see Table 1), they said little about how their designs supported users in helping and giving. This is another mismatch, because these, although not top motivations, were important for user-providers; *"that would be a way that I could share my experience, my education with other people,"* and, *"I really liked the idea of offering accommodation for other travelers,"* and, *"I love the feeling of coming home and knowing that you've made someone's stay like really fun. So that was great. That would be one motivation."*

Another disconnect is that some user-providers (UPs) spoke of possibilities for self-improvement from a moral integrity perspective; *"it's kind of balancing my guilt for having all this other stuff that I can afford,"* and, *"I think I was opened to it because it matched my values,"* and, *"the possibility of supporting people in a car-free lifestyle."* Providers rarely mentioned designing to encourage this motivation.

Finally, in breakouts, we did find that UPs in non-profits and altruistic services talked a lot about the motive of giving to others whereas it was underemphasized by service providers. UPs mentioned this far less in the for-profits and instrumental services with payment being their top motive. Meanwhile, service providers in for-profit and instrumental services did not focus nearly so heavily on payment in talking about their designs, although they did recognize it as a motive.

Design Implication: Help Me Become My Ideal Me

The notion of moral self-improvement may be another opportunity largely missed by providers to draw people into using peer services. Appealing to identity-congruent values can influence consumer behavior [40] with large positive impacts as in the successful anti-littering campaign in Texas in the 1980's organized by Dan Syrek which appealed to the pride of the Texan male [28]. So peer system providers would do well to find out who their target users believe themselves to be and forcefully proclaim identity congruent personal ideals that can be achieved by using their service.

CONCLUSION

We conducted a unique, qualitative, interview-based study that allows us to compare the motivations that users have for participation in peer services as both receivers and providers of resources with motivations anticipated by service providers. What we have found is a number of mismatches highlighted in Table 1 that suggest areas for providers to improve their designs.

At the top level it seems fair to say that providers we talked to want to create a better world. Their model for success is to give people an opportunity to behave in a more sustainable fashion, by sharing resources and helping one another. Users, on the other hand, are operating on a model of how best to satisfy instrumental needs, particularly basic ones: just getting what they need to survive at a competitive price, with maximum convenience. It therefore seems that the sharing economy backlash [11,15,38] is grounded in an unavoidable fact that, to be successful, peer services must first serve self-interest and that a *muddle of models* will remain in play unless providers embrace what really motivates users.

Nonetheless we do see that user-providers can be motivated by altruism and moralistic impulses and that philanthropic service providers might do better to switch their focus on promoting utopian ideals and concentrate more on encouraging and rewarding these more personal motivations.

LIMITATIONS AND FUTURE WORK

Our study and report here have some limitations. First, for the sake of space, we presented different types of systems' motivations in one table, rather than breaking them apart to show some of the variations we mention in our discussion. In addition, we were unable to talk to users and providers of the exact same systems. It may be that providers of some of the services we did not cover have a deeper understanding of what users want. Certainly some have been immensely successful. But of course there may always be room for improvement and our findings may also be of value to them. Further work also remains to be done to implement some of our design ideas in some real peer systems and to collect data on pre- and post-intervention user behaviors and subjective impressions. In addition, there is more to be done to develop our analytical framework and to use it to create different ideal motivational blueprints for different kinds of service that can guide developers of peer services.

ACKNOWLEDGMENTS

Thanks to the many users and providers of peer services who gave us interviews and to Shawna Hein and Matthew Valente for conducting some of the interviews. To Josephine Giaimo for input to an early draft of the interview guideline document. This paper is based upon work funded by the National Science Foundation under Grant No. IIS-1218544.

REFERENCES

1. Ajzen, I. From intentions to actions: A theory of planned behavior. In J. Kuhl, J. Beckmann (Eds). *Action Control: From Cognition to Behavior*. Springer Berlin Heidelberg (1985), 11-39.
2. Ajzen, I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes 50,2* (1991) 179–211.
3. Ajzen, I. & Fishbein, M. *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall, (1980).
4. Aunger, R. & Curtis, V. The Anatomy of Motivation: An Evolutionary-Ecological Approach. *Biological Theory, 7,3* (2013) 49-63.
5. Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review,* 84,2 (1977) 191-215.
6. Bauwens, M. P2P Foundation. Category: Relational. http://p2pfoundation.net/Category:Relational
7. Bellotti, V., Carroll, J.M., & Han, K. Random Acts of Kindness: The Intelligent and Context-Aware Future of Reciprocal Altruism and Community Collaboration. In *Proc. CTS 2013,* IEEE (2013), 1-12.
8. Bellotti, V., Cambridge, S., Hoy, K., Shih, P.C., Renery Handalian, L., Han, K., & Carroll, J.M. Towards community-centered support for peer-to-peer service exchange: rethinking the timebanking metaphor. In *Proc. CHI 2014.* ACM Press, (2014) 2975-2984.
9. Botsman, R. The Sharing Economy Lacks A Shared Definition. *Fast Company Co.Exist Article* November 21, 2013. http://www.fastcoexist.com/3022028/the-sharing-economy-lacks-a-shared-definition
10. Botsman, R. and Rogers, R. (2011). *What's mine is yours. How Collaborative Consumption is changing the way we live.* HarperCollinsPublishers: London.
11. Cagle, S. The Case Against Sharing: On Access, Scarcity and Trust. Article in *The Nib*, May 27th, 2014. https://medium.com/the-nib/the-case-against-sharing-9ea5ba3d216d
12. Carroll, J.M. & Bellotti, V. Creating Value Together: The Emerging Design Space of Peer-to-Peer Currency and Exchange. To appear in *Proc. CSCW 2015*. ACM Press (2015).
13. Chakrabarti, B. & Baron-Cohen, S. Empathizing: neurocognitive developmental mechanisms and individual differences. Anders, Ende, Junghöfer, Kissler & Wildgruber (Eds.) *Progress in brain research, 156* (2006), 403-417.
14. Charlton, B.G. The inequity of inequality: egalitarian instincts and evolutionary psychology. *Journal of Health Psychology, 2* (1997), 413-425.
15. Chase, B. Sharing is caring. But not in the sharing economy. *Huffington Post, 7/25/2014.* Retrieved from http://www.huffingtonpost.com/brad-chase/sharing-is-caring-but-not-millenials_b_5618963.html
16. Cialdini, R. *Influence: The psychology of persuasion.* New York; Quill (1984).
17. Collom, E. The motivations, engagement, satisfaction, outcomes, and demographics of timebank participants: survey findings from a U.S. system. *International Journal of Community Currency Research 11*, (2007), 36-83.
18. Collom, E. Motivations and differential participation in a community currency system: The dynamics within a local social movement organization. *Sociological Forum, 26, 1* (2011), 144-168.
19. Conley, C. *Peak: How great companies get their mojo from Maslow*. Jossey-Bass, (2007).
20. Cook, K.S., Cheshire, C., Rice, E.R.W. & Nakagawa, S. Social exchange theory. In J. DeLamater & A. Ward (Eds.), *Handbook of social psychology.* Dordrecht: Springer (2013), 61-88.
21. Crowne, D., & Marlowe, D. *The approval motive: Studies in evaluative dependence*. New York: Wiley (1964).
22. Csikszentmihalyi, M. *Flow: The psychology of optimal experience.* Harper & Row (1990).
23. Deci, E. L., Koestner, R., & Ryan, R. M. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125,6* (1999). 627–668.
24. Eisenstein, C. *Sacred economics: Money, gift, and society in the age of transition*. Evolver Editions, 2011.
25. Emerson,R.M. Exchange Theory, Part I: A Psychological Basis for Social Exchange and Exchange Relations in Network Structures. In *Sociological Theories in Progress*. Houghton Mifflin, Boston, MA (1972a), 38-87.
26. Emerson, R.M. Exchange Theory, Part II: A Psychological Basis for Social Exchange and Exchange Relations in Network Structures. In *Sociological Theories in Progress*. Houghton Mifflin, Boston, MA (1972b), 38-87.
27. Gansky, L. *The Mesh: Why the future of business is sharing.* Portfolio Trade, 2012.
28. Heath, C. & Heath, D. *Made to Stick: Why some ideas survive and others die*. Random House (2007).
29. Henrich J, Gil-White, F. The evolution of prestige: freely conferred deference as a mechanism for enhancing the

benefits of cultural transmission. *Evol Hum Behav, 22* (2001) 165–196.

30. Herzberg, F. (1966) *Work and the nature of man*. Cleveland, OH: World.
31. Homans, G. C. Social behavior as exchange. *American Journal of Sociology, 63*, (1958), 597–606.
32. Huitt, W. (2007). Maslow's hierarchy of needs. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from, http://www.edpsycinteractive.org/topics/regsys/maslow.html
33. Knoke, David. Incentives in Collective Action Organizations. *American Sociological Review 53,3* (1988), 311–329.
34. Lampinen, A., Lehtinen, V., Cheshire, C. & Suhonen, E. Indebtedness and reciprocity in local online exchange. In *Proc. CSCW'13*. ACM Press, (2013), 661-672.
35. Maslow, A.H. A theory of human motivation. *Psychological Review, 50,4* (1943) 370–96.
36. Maslow, A. H. *Motivation and personality*. New York: Harper & Row. (1970a).
37. Molm, L. Collett, J. & Schaefer, D.R. Building Solidarity through Generalized Exchange: A Theory of Reciprocity. *American Journal of Sociology, 113,1* (2007), 205–42.
38. Morozov, E. The 'sharing economy' undermines workers' rights. *Notes EM.* Retrieved 7/13/2014 from: http://www.ft.com/intl/cms/s/0/92c3021c-34c2-11e3-8148-00144feab7de.html
39. Msroka, M. *Motivating workers in educational institutions: Adams' equity and Maslow's need hierarchy theoretical implications.* GRIN Verlag (2013).
40. Oyserman, D. Identity-based motivation: Implications for action-readiness, procedural-readiness, and consumer behavior *Journal of Consumer Psychology 19* (2009) 250–260.
41. Pink, D. H. *Drive: The surprising truth about what motivates us.* Riverhead Books (2011).
42. Regan, R. T. Effects of a favor and liking on compliance. *Journal of Experimental Social Psychology* 7 (1971), 627–639.
43. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68–78.
44. Schor, J. B. *True Wealth: How and why millions of Americans are creating a time-rich, ecologically light, small-scale, high-satisfaction economy.* Penguin Books, (2010).
45. ServiceSpace Blog Post. *From sharing economy to gift ecology.* Posted by Nipun Mehta on May 31, 2014. http://www.servicespace.org/blog/view.php?id=14918
46. Suhonen, E., Lampinen, A., Cheshire, C. & Antin, J. Everyday favors: a case study of a local online gift exchange system. In *Proc. GROUP '10.* ACM Press, (2010), 11-20.
47. Tay, L. & Diener E. Needs and subjective well-being around the world. *J Pers Soc Psychol, 101,2* (2011) 354-65.
48. Trivers, R. L. The evolution of reciprocal altruism. *The Quarterly Review of Biology* 46, 1 (1971), 35-57.
49. Vallerand, R. J., & Fortier, M. S. Measures of intrinsic and extrinsic motivation in sport and physical activity: a review and critique. In J. Duda (Ed.), *Advancements in sport and exercise psychology measurement* (pp. 83–100). Morgantown, WV: Fitness Information Technology (1998).
50. van de Glind, P. *The consumer potential of Collaborative Consumption.* Research MSc in Sustainable Development – Environmental Governance Faculty of Geosciences, Utrecht University, the Netherlands August, (2013).
51. Vision Critical. *Sharing is the new buying. How to win in the collaborative economy* (2014). Retrieved from: http://www.visioncritical.com/sites/default/files/pdf/sharing-new-buying-collaborative-economy-report.pdf
52. Wahba, M. A., & Bridwell, L. G. (1976). Maslow reconsidered: A review of research on the need hierarchy theory. *Organizational Behavior and Human Performance, 15,2* (1976) 212–240.
53. Webb, T.L., Joseph, J., Yardley, L. & Michie, S. Using the Internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res. 12,1* (2010).
54. Willer R. Groups reward individual sacrifice: the status solution to the collective action problem. *Am Sociol Rev 74* (2009) 23–43.
55. Williams, G. C., & Deci, E. L. Internalization of biopsychosocial values by medical students: a test of self-determination theory. *J Pers Soc Psychol, 70* (1996). 767–779.
56. Williams, G. C., Grow, V. M., Freedman, Z. R., Ryan, R. M., & Deci, E. L. Motivational predictors of weight loss and weight-loss maintenance. *J Pers Soc Psychol, 70* (1996). 115–126.
57. Wright, R. *The moral animal: Why we are, the way we are: The new science of evolutionary psychology.* Vintage; Reprint edition, (1995).

THE CONVERSATION

Why we should design smart cities for getting lost

April 7, 2016 9.08pm BST

Marcus Foth
Marcus Foth is a Friend of The Conversation.
Professor, Urban Informatics, Queensland University of Technology

The Conversation's partners

The Conversation UK receives funding from Hefce, Hefcw, SAGE, SFC, RCUK, The Nuffield Foundation, The Ogden Trust, The Royal Society, The Wellcome Trust, Esmée Fairbairn Foundation and The Alliance for Useful Evidence, as well as sixty five university members.

View the full list

The 'Lose Yourself in Melbourne' ad was onto something: instead of being directed to the fastest or shortest route, some people might want to take a diverting detour. 'It's Easy to Lose Yourself in Melbourne', Tourism Victoria

The internet has reached our cities. A smart city is optimised for efficiency, productivity and comfort.

The smart city uses intelligent transport systems. It is administered by integrated urban command centres, which analyse the omnipresent raw material of the digital era: big data. As citizens go about their everyday lives, they leave data traces everywhere, even in the sewers.

Many technology companies and city governments celebrate the new *enfant terrible* of smart city research: the urban scientist who finally imposes a rigorous scientific (that is, positivistic) mindset on city governance. However, Jeremy Kun confirms that:

> *… being quantitative doesn't protect against bias.*

Commentators such as Cat Matson, Charles Landry and Paul Mason advocate a people-centred approach to city design. In our own work, we warn that ignoring decades of research by architects, geographers, urban planners, designers and sociologists could lead to a dystopian future where humans lose agency if we mindlessly pursue convenience and efficiency.

Wall-E presents a vision of a dystopian future created by techno-scientific determinism.

Algorithmic culture of like-mindedness

Big data requires analysis by algorithms, and they in turn create filter bubbles. Corporations such as Facebook and Google deploy sophisticated algorithms to help us navigate the otherwise bloated social mediascape. The content displayed on Facebook's news feed is selected based on a user's profile, location, interests, online habits – what they post, share, recommend and "like".

The popularity of social media stems from its power to create personalised spaces, walled gardens, which are tailored to individual preferences and favour content *relevant* to each user. Proprietary algorithms determine what is deemed *relevant*.

Without ethics, it is these algorithms that determine the make-up of the Facebook news feed, Google's top search results and the recommendations on whom to follow on Twitter and what to buy on Amazon. They are optimised to prioritise content that generates more business.

As Gilad Lotan has observed:

> *We're not seeing different viewpoints, but rather more of the same. A healthy democracy is contingent on having a healthy media ecosystem. As builders of these online networked spaces, how do we make sure we are optimising not only for traffic and engagement, but also an informed public? … The underlying algorithmics powering this recommendation engine help reinforce our values and bake more of the same voices into our information streams.*

The diversity advantage of cities

As more and more social media platforms embrace urban environments as their playground, this algorithmic culture has important implications for cities.

People come together in cities not just for the infrastructure and convenience they provide, but for offering choice. Cities are fundamentally about possibilities, opportunities and diversity.

Jane Jacobs notes:

> *Cities have the capability of providing something for everybody only because, and only when, they are created by everybody.*

Ethan Zuckerman thinks of cities as serendipity engines:

> *By putting a diverse set of people and things together in a confined place, we increase the chances that we're going to stumble onto the unexpected.*

However, Zuckerman also asks:

> *… do cities actually work this way?*

It's a timely question for smart cities governed by big data and algorithmic analysis. How can a smart city become a serendipity engine? Can we design smart cities for getting lost?

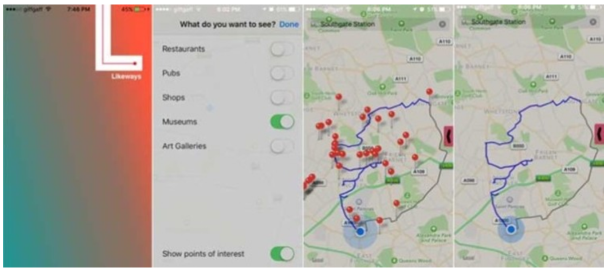
Here are some examples of why that may not be such a bad idea.

Getting lost and getting to know strangers

Public transport journey planners are usually optimised for two factors: the fastest speed and shortest distance to get you from A to B. Yet there are opportunities beyond telematics.

Why don't we offer the choice to go slow, to take the least polluted route to work, or the scenic way home?

Experimental prototypes such as Martin Traunmueller's Likeways and Mark Shepard's Serendipitor allow you to lose yourself and rediscover your city.



The Likeways app lets users choose routes that wander past restaurants, pubs, shops, museums or art galleries.
Screenshot, courtesy Martin Traunmueller

In addition to a diversity of places, cities also offer a diversity of people. However, all too often we stay within our existing social networks of friendship and convenience. Eric Paulos's Familiar Stranger Project investigated anxiety, comfort and play in public places.

Yet our ability to unlock the advantage of a city's social diversity is still in its infancy. Early examples include co-working spaces and meet-up groups that bring diverse people together, airlines offering social seating, and design interventions such as Jokebox that foster playfulness and curiosity.

Co-working spaces aim to spark creativity and innovation by bringing diverse people together. flickr/Impact Hub, CC BY-SA

Saskia Sassen warns that the privatisation of public spaces in the city:

> *… has deep and significant implications for equity, democracy and rights.*

This also stifles innovation, as people are lacking the kinds of “skunkworks” that foster creativity and diversity.

Deliberative democracy and the city

Besides the nascent opportunities vested in people and places, what may well be the final frontier of a truly smart (as in intelligent) city is content and discourse. Seeking to burst the filter bubbles, Eli Pariser created Upworthy:

> *… on a mission to change what the world pays attention to.*

It curates news you should read rather than just “the news you want” to read.

Another illustrative example is Rebecca Ross’s project London is Changing. Large digital displays visualise local community voices and juxtapose diverse opinions about the impact of gentrification in London.

A digital billboard displays local community voices and diverse opinions about gentrification in London. London is Changing

To uphold a citizen’s right to the digital city and strengthen the role of cities in a deliberative democracy, cities should empower citizens to be smart. Smart cities should allow us to get lost and find new places, to meet strangers who may become new friends, and to engage in discussions with diverse others so we may form new opinions.

Alexandros Washburn observes:

> *In talking about the anticipational Smart wonders of 'our city', we really mean 'my city'. We confuse the collective with the personal.*

If collective, that is, civic intelligence is what makes cities smart, we need more serendipity engines. Smart cities should be open and agile and employ what Bob Dick calls "dialectical processes" and what Anna Cox calls "design friction".

That way we may change algorithmic filters for the sagacious discovery of diversity in the city. Diversity fuels innovation, and innovation is what we need to be sustainable.

Cities
Social media
Melbourne
Innovation
Digital media
Diversity
Urban design
Big data
London
Cultural diversity
Smart cities
algorithms
digital rights
Cities & Policy

From Users to Citizens: Some Thoughts on Designing for Polity and Civics

Marcus Foth[1] **Martin Tomitsch**[2] **Christine Satchell**[1] **M. Hank Haeusler**[3]

[1] Urban Informatics Research Lab, Queensland University of Technology, Brisbane, Australia

[2] Design Lab, Faculty of Architecture, Design and Planning, The University of Sydney, Australia

[3] Smart Cities Research Cluster, UNSW Built Environment, Sydney, Australia

m.foth@qut.edu.au martin.tomitsch@sydney.edu.au christine.satchell@qut.edu.au m.haeusler@unsw.edu.au

ABSTRACT

This paper presents an essay aimed at prompting broad discussion crucial in keeping the interaction design discourse fresh, critical, and in motion. We trace the changing role of people who have advanced from consumers to producers, from stationary office workers to mobile urban nomads, from passive members of the plebs to active instigators of change. Yet, interaction designers often still refer to them only as 'users.' We follow some of the historic developments from the information superhighway to the smart city in order to provide the backdrop in front of which we critically analyse three core areas. First, the issue of echo chambers and filter bubbles in social media results in a political polarisation that jeopardises the formation of a functioning public sphere. Second, pretty lights and colourful façades in media architecture are increasingly making way for situated installations and interventions fostering community engagement. And third, civic activism is often reduced to forms of slacktivism. We synthesise our discussion to propose 'citizen-ability' as an alternative goal for interaction designers to aspire to in order to create new polities and civics for a better quality of life.

Author Keywords

Civic engagement; community engagement; polity; civics; media architecture; activism; smart cities; smart citizen; civic intelligence; participation; urban informatics

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

As both technology and technological practices evolve, human-computer interaction (HCI) has expanded its focus from the design and assessment of particular interaction styles, to encompass the role that interactive systems play in connecting people with their world. The focus of HCI is no longer grounded by the notion of the stationary user moored to a fixed desktop PC in an organisational context. Rather, users are interacting with technologies across many contexts and in all areas of their lives. The design and development of interactive technology has similarly broadened. These days, users are not only exposed to a standard TV set at home and a desktop computer at work, but to a plethora of different interfaces and interactive devices that blur many of the previously drawn boundaries between home and work, mobile and static, public and private. Despite an arguably richer and more nuanced perspective of *use* emerging from HCI research, there is a need to similarly expand the scope to encompass notions that provide alternative or complementary perspectives than just use and usability. In this paper, we provide an analysis of what they may be, by focusing on how political activism and civic engagement can give rise to the idea of 'citizen-ability,' that is, design not just in pursuit of a better user experience, but a better citizen experience and in fact a strengthening of the efficacy of our citizenry and its polity.

Figure 1: Occupy Central, Hong Kong, 2014

We argue that design research into fostering civic and urban engagement is a significant and timely topic for three key reasons: First, *place* – the feared 'death of distance' that had been heralded when the internet first became commercially successful, never happened. In the contrary, the bold rhetoric that predicted face-to-face to make way for a proliferation of e-commerce, distant education, telework, and other remote online transaction capabilities, never became entirely true as predicted. Local place thrives, and does so in ways ameliorated by new forms of situated engagement, locative media, and location-based services (Gordon & de Souza e Silva, 2011). This is further corroborated by the rapid pace of urbanisation that has tipped over 50% across the world, and reaches close to 90% in places such as Australia where the majority of the land is sparsely populated.

OzCHI '15 , December 07 - 10 2015, Melbourne, VIC, Australia
Copyright © 2015 ACM 978-1-4503-3673-4/15/12... $15.00
http://dx.doi.org/xx.xxxx/xxxxxxx.xxxxxxx.

Second, *technology* – ubiquitous computing (UbiComp) has spilled outside the traditional HCI bastions of 'work' and 'home' into every aspect of human endeavour (Bødker, 2006). This trend brought about not just new technology and interface innovations but also new technical user practices that bridge the physical and the digital city: mobility, situated technology, embodied interaction, augmented reality, urban screens, big data.

Third, *people* – must no longer be understood as passive users or consumers placed by designers, developers and marketeers at the receiving end of products and services, since they are increasingly able and empowered to also be producers (Bruns, 2008) and active agents of change (Foth, Forlano, Satchell, & Gibbs, 2011). Civic participation and engagement is thus a useful theme begging further exploration in light of cities turning more and more interactive and responsive (Foth, 2009), technology opening up new platforms and channels for citizens to be heard (Schuler, 2008), and people no longer being limited to conventional modes of citizenship (Burgess, Foth, & Klaebe, 2006) – or are they...?

In this paper, our understanding of 'civic engagement' is not as broad as what the term may entail; it does not include municipal services that are being provided or enhanced through digital means. It does not focus on the usual array of citizenship, such as voting, jury duties, consular assistance, legal rights and obligations, although it does not explicitly exclude them when they become relevant as a consequence of citizen action. This paper concentrates on what is usually termed 'civic action,' 'citizen activism,' and 'grassroots community engagement.' Prominent examples of innovations in this space that combine the aforementioned trends across people, place, and technology, include ushahidi.com – a citizen journalism and news aggregator that was initiated as a result of the 2007/2008 Kenyan post-election violence (Hirsch, 2011); the Occupy movement (Figure 1); crowdsourcing measurements of radiation levels around Fukushima following the 2011 nuclear disaster in Japan; and, the way social media and mobile devices were used to mobilise the masses from large scale, multi-country movements such as the 'Arab Spring,' to local issues such as 'Stuttgart 21,' a protest movement against a controversial railway and urban development project in southern Germany that contributed to the very first Greens Party politician to be elected Head of State of Baden-Württemberg.

What should the next generation of civic innovation look like? We call not only for new 'digital soapboxes' that urban citizens can use to have a voice and make themselves heard, but also for new strategies and approaches to close the gap between community activism that 'only' raises awareness of a particular civic issue on the one hand, and on the other hand traditional forms of top-down governance (Figure 2), polity, and decision making that may or may not bring about the societal change we need (Foth, Parra Agudelo, & Palleis, 2013). It strikes us that currently, there appears to be not just a disconnect but also an increasingly widening discrepancy in the level of interest and the pace of innovation on each side. We hope that this paper can contribute towards the discourse in interaction design that informs the thinking of the HCI and UbiComp communities when designing new and innovative interfaces for civic and urban engagement.

Figure 2: The House of Commons, London, UK

The paper is structured into three main sections. We will first provide a brief history of some of the pertinent trends and developments not just in computing but also societal terms. We will then present our analysis of three interrelated areas across social, spatial and technical domains: (a) social media; (b) media architecture, and; (c) civic activism. These perspectives inform our synthesis section where we outline some of our thoughts on designing for polity and civics, and argue for a view that is broader than just the current focus on usability, and that includes notions of citizenship, values, and ethics.

THE STORY THUS FAR

We believe that we cannot look into the future if we cannot see the past. As such, we will first provide a brief recount of some key historic developments in computing – both technical and social. This will enable us to create an informed foundation on top of which we will construct our analysis and argument.

From Information Superhighways to Digital Cities

With the advent of the commercially viable internet in the mid/late 1990s, many scholars and commentators recognised the benefits that digital communication would bring to society. The 'death of distance' was heralded that would lead to shifting trends in economic and socio-cultural terms, such as online communities (Kim, 2000; Preece, 2000), e-commerce (Hagel & Armstrong, 1997; Hearn, Mandeville, & Anthony, 1998), and distant education (Dhanarajan, 2001). This debate quickly created a conceptual dichotomy between the 'real world' (offline) and 'cyberspace' (online).

However, as more and more internet-related technological innovations unfolded, were taken up by society, and in turn, translated and appropriated into new and changing technological practices conducted as part of everyday life, scholars started to collect compelling empirical evidence that this online/offline dichotomy was an artificial one. From different disciplinary perspectives came the clear message that the increasing ubiquity of internet connectivity, digital technology, web applications, and location-based services allows for a seamless transitioning between both the online and offline aspects of our everyday lives as well as the visible and the invisible infrastructures of cities. Following on from Castells' (2000) *Information Age* trilogy, but also sharply critiquing Putnam's (2000) bleak outlook suggesting a correlation between internet uptake and a steep decline of (conventional) civic engagement, Wellman (2001) articulated a view that brought together notions of physical place and cyberspace: networked individualism. Similarly, other scholars in cultural geography and urban studies seconded this motion to re-evaluate the importance of place (Foth, Choi, Bilandzic, & Satchell, 2008; Pons-Novell & Viladecans-Marsal, 2006; Walmsley, 2000).

Mobile phones have become a well-established communication device – not only to connect with distant others, but also to coordinate social interactions in your physical vicinity, e.g. spontaneously organising collective actions (Rheingold, 2002). As Gordon and de Souza e Silva (2011) argue, place continues to matter in a networked world: *"The local still matters, and in fact, it may matter more than ever before because it can have an immediate and powerful global impact."* (p. 168).

Social and Community Informatics have always emphasised the cultural and institutional contexts of ICT use (Gurstein, 2000; Kling, Rosenbaum, & Sawyer, 2005), and increasingly examine the significance of mobility and place in people's everyday experiences (Foth, Forlano, et al., 2011; Gordon & de Souza e Silva, 2011). Similarly, contributions in the field of Computer Supported Cooperative Work (CSCW) that deal with the way technology can assist people in improving their work environments, have recognised the importance of place metaphors in media space and even postulated early on the emergence of hybrid (physical and virtual) space (Harrison & Dourish, 1996).

Everyday technology has become increasingly ubiquitous: networked, embedded and accessible anywhere, anytime. Dourish and Bell (2011) argue that the design and development of UbiComp as well as the ability to access information in places other than the conventional desktop PC, call for a better appreciation of the "messiness of everyday life," which ultimately requires social and cultural research skills such as ethnographically-informed approaches in addition to technical and design expertise. Williams et al. (2009) point out that ubicomp innovations enable HCI designers to design for a diversity of urban environments and urban citizens, rather than being limited to universal and homogeneous design outcomes. We agree with Odendaal (2006), who recommends qualitative methods to *"understand differing ways of life [and] to reveal a diversity of urban experiences"* (p. 36). In our work, we employ a design research methodology that considers *"urban experiences across different urban contexts that are created by new opportunities of real-time, ubiquitous technology"* (Foth, Choi, & Satchell, 2011).

From Smart Cities to Smart Citizens

The future of civic engagement is characterised by both technological innovation as well as new technological user practices that are fuelled by trends towards mobile, personal devices; broadband connectivity; open data; urban interfaces; and, cloud computing. These technology trends are progressing at a rapid pace, and have led global technology vendors to package and sell the 'Smart City' as a centralised service delivery platform predicted to optimise and enhance the key performance indicators of cities – and generate a profitable market (Figure 3). The top-down deployment of these large and proprietary technology platforms have helped sectors such as energy, transport, and healthcare to increase efficiencies. However, an increasing number of scholars and commentators warn of another 'IT bubble' emerging. Along with some city leaders, they argue that the top-down approach does not fit the governance dynamics and values of a liberal democracy when applied across sectors. A thorough understanding is required, of the socio-cultural nuances of how people work, live, play across different environments, and how they employ social media and mobile devices to interact with, engage in, and constitute public realms.

Figure 3: IBM Smart City Control Centre in Rio de Janeiro

Genuinely putting people, that is, a socio-culturally nuanced understanding of urban citizens, at the centre of a revised notion of the 'smart city' sounds simple, even trivial, but it is not. Similarly to how Bannon (1992) at the time called for a profound shift in attention *"from human factors to human actors,"* more and more commentators these days have started to critique the commercial and top-down-only vision of the smart city and consider alternative approaches that focus on the "smart citizens" (Foth & Brynskov, 2016, in press; Foth, Brynskov, & Ojala, 2016; Foth, Hudson-Smith, & Gifford, 2016, in press; Townsend, 2013; Waal, 2014).

We are far from witnessing another *Biedermeier* period – on the contrary: post-election violence in Kenya in 2008, the Occupy movements in New York, Hong Kong and elsewhere, the Arab Spring, Stuttgart 21, Fukushima, the Taksim Gezi Park in Istanbul, and the Vinegar Movement in Brazil in 2013. These examples of civic action shape the dynamics of governments, and in turn, call for new processes to be incorporated into governance structures. Participatory inquiries into these new processes across the triad of people, place and technology are a significant and timely investment to foster productive, sustainable, and liveable human habitats. We seek to reframe the current debates in academia and priorities in industry and government to reconceptualise the citizenry as a collective agent in tackling societal issues and struggles. This calls for new participatory approaches for co-inquiry and co-design. It is an evolving process with an explicit agenda to facilitate change – change that requires new governance infrastructures and practices for civic engagement.

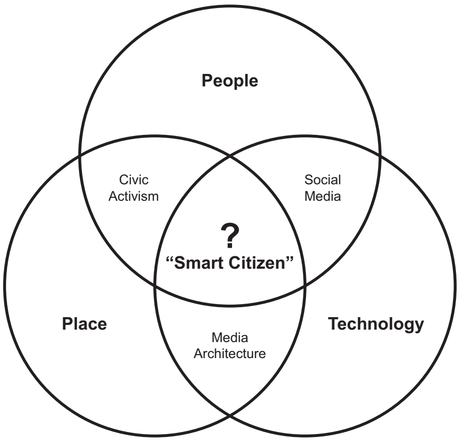


Figure 4: Our analysis across people, place, technology

ANALYSIS

The critical perspectives relating to the design of information technology and computing that we want to discuss, are positioned across a triad of people, place, and technology (Figure 4). We have identified three themes in the nexus of each area for further elaboration: social media; media architecture, and; civic activism.

Social Media

In 2006, TIME Magazine published a picture (Figure 5) of a computer on its cover as the "Person of the Year" with the words *"You. Yes, you. You control the Information Age. Welcome to your world."* The selection highlights the profound shift in the way that the World Wide Web has advanced to allow an increasing number of everyday people to not only access information, but contribute and participate in their own right. This trend has been popularised as, "Web 2.0," or "the social media revolution" – Jenkins (Jenkins, 2006) coined the term, "participatory culture." The ability for everyone to become a creator, publisher, remixer, recommender, sharer, and referrer has led to an exponential growth in content: Every 60 seconds on Facebook, 510 comments are posted, 293,000 statuses are updated, and 136,000 photos are uploaded (thesocialskinny.com). However, quantity of content does not imply quality, and thus with more sources of content, spread via more digital media channels, to more people, web users started to face the problem of information overload.

Corporations such as Facebook and Google have deployed sophisticated filters and recommendation systems designed to help us navigate the otherwise bloated social mediascape. The content displayed on Facebook's news feed is selected based on a user's profile, their location, interests, habits, online transactions – what they post, share, recommend, and "like." The popularity of social media stems from its power to create personalised spaces, walled gardens, which are tailored to individual preferences and favour content relevant to each user. An algorithm proprietary to each social media site determines what is deemed relevant: With the absence of a journalistic or editorial code of ethics, these algorithms determine the make-up of the Facebook news feed, Google's top search results, and the recommendations on whom to follow on Twitter and what to buy on Amazon. They are optimised to prioritise content that will generate more traffic. Yet, Lotan (2014) warns that, *"We're not seeing different viewpoints, but rather more of the same. A healthy democracy is contingent on having a healthy media ecosystem. As builders of these online networked spaces, how do we make sure we are optimizing not only for traffic and engagement, but also an informed public? … The underlying algorithmics powering this recommendation engine help reinforce our values and bake more of the same voices into our information streams."*

The compounding aspects of this polarisation of opinions in social media have been studied in political science and media and communication studies, e.g., echo chambers (Aiello et al., 2012) and filter bubbles (Pariser, 2011).

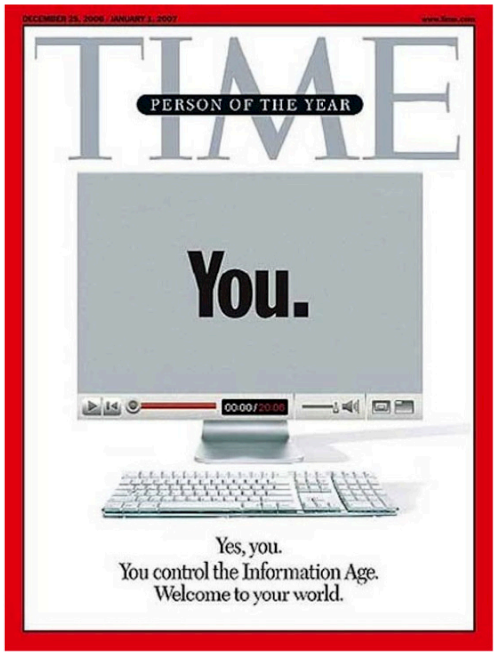


Figure 5: TIME cover from 25 Dec 2006

Although the advantages of using social media for civic engagement have been demonstrated in numerous studies (Foth, Forlano, et al., 2011; Rotman et al., 2011), we are wary that incremental improvements to the same platforms will not bring about a quantum change in the practice and impact of civic engagement. However, we see potential in the socio-cultural diversity that cities offer (Wood & Landry, 2007), and call for a focus on the touch points between 'the city' and its civic body, the citizenry. In order to provide for meaningful civic engagement, the city must provide appropriate interfaces (Foth & Brynskov, 2016, in press). Such urban interfaces can provide an innovative avenue for addressing these issues by fostering depolarisation through engagement with civic media. Urban interfaces bring unique qualities, such as their ability to reach a diversity of citizens, and the absence of personalisation algorithms. Further research is required to examine how the arrival and uptake of urban interfaces and situated civic media can be integrated into our city environments to reach, support and engage citizens. In this context, we are particularly interested in their ability to break echo chambers and burst filter bubbles.

Media Architecture

The discipline of media architecture is developing and growing as designers, architects, and planners realise the practice and promise that the combination of digital media and architecture can provide to enhance the experience of the built environment. Not only do the professionals in these disciplines need to consider how to incorporate the use of technology into the development of their profession, but they also need to understand how technology can be used to improve how people engage with the built environment.

According to Brynskov et al. (2013), *"Media Architecture is an overarching concept that covers the design of physical spaces at architectural scale incorporating materials with dynamic properties that allow for dynamic, reactive or interactive behavior. These materials are often digital, but not always, and they allow architects and (interaction) designers to create spatial contexts for situations using a variety of modalities."* (p. 1-2). Media façades are a typical example of media architecture, with other forms including physical structures, urban screens, light projections, and tangible interfaces and installations.

Media architecture is closely related to the trend of ubiquitous computing to spill over into urban environments. However, it is now time to question and analyse the purpose and impact of many media architecture projects to date, beyond the cosmetic or decorative quality of 'urban Christmas' lights. As the UK graffiti artist Banksy provokes, *"twisted little people ... go out everyday and deface this great city. Leaving their idiotic little scribblings, invading communities and making people feel dirty and used. They just take, take, take and they don't put anything back. They're mean and selfish and they make the world an ugly place to be. We call them advertising agencies and town planners"* (cited in Sliwa & Cairns, 2007, p. 78). As some like Banksy may think that media architecture – if not considered and appropriated properly – runs the risk of polluting the city with more advertising and media 'junk.'

Is there a role for media architecture to resurrect the significance and use of town halls, civic squares and public spaces of the city long lost to car parks and shopping centres? Have these places, like the Greek *agora*, ever existed in more recent times, or are they a long-lived phantom created by the romantic memories of the perpetually previous generation? Arnold (2003) argues that, *"Community is dynamic, and much angst is no doubt driven by nostalgia that fails to recognize the strengths of contemporary communities and the changing forms of contemporary communities."* (p. 78).

Figure 6: Mégaphone by momentfactory.com, Montreal (Fortin, Neustaedter, & Hennessy, 2014)

Figure 7: Discussions in Space at Fed Square, Melbourne (Schroeter, Foth, & Satchell, 2012)

Although the current practice in media architecture is often civically bleak, there are select projects that do show promise, such as the *Mégaphone* (Figure 6) (Fortin et al., 2014) and *Discussions in Space* (Figure 7) (Schroeter et al., 2012). Dade-Robertson (2013) makes the analogy between Graphical User Interfaces (GUIs) of personal computers with how he defines *Architectural User Interfaces* (AUIs) as buildings that mediate between computational information and people. In so doing, he connects the disciplines of architecture and human-computer interaction, arguing that not only does media and technology affect how people experience urban environments, architecture similarly has an effect on the development of computer technologies (Dade-Robertson, 2013). He believes that through the rise of ubiquitous computing the value of physical environments has been re-acknowledged – increasing the opportunities for architectural influence on the evolution of HCI practices. The call for architectural knowledge and input into HCI research is reinforced by Fischer et al. (2013) who claim that architecture provides spatial understandings that can assist in the development of urban HCI systems. They argue that through an architectural approach, media architecture can be refocused *"for a city beyond information and utility"* (p. 39).

We ask how can media architecture help people take control, appropriate place, and create communities. Acknowledging that media architecture is an emerging field that combines people, place, and technology in a similar way to related hybrid practices such as urban informatics (Foth, Choi, et al., 2011), it has an effect on the way the city is experienced and how people come together. We draw attention to the role that media architecture will have in facilitating communication and the interaction of city inhabitants.

We are also interested in identifying novel ways to put some of the creative process into the hands of laypeople, and in investigating the impact this may have on community engagement and citizen control. We thus ask if and how media architects as a community of practice should encourage and foster to 'open-source' our tools and approaches in order for laypeople to not only participate but become active instigators of change in their own right (Caldwell & Foth, 2014).

Civic Activism

Dutton (2009) refers to the public participatory capabilities of the internet and the web as the 'fifth estate,' as distinct from the legislative, executive, judicial branches of government, and the media. In this section, we look at three forms of citizen activism and community advocacy that operate within the fifth estate: (1) established non-government organisations (NGOs) that embrace the web to extend their campaign efforts; (2) web-based petition sites that heavily employ social media for viral marketing of 'people-powered politics' and campaigns; (3) examples of ad hoc, impromptu, and in situ (at times, hyperlocal) activism that employs hybrid – combining physical and digital – forms of engagement.

Old Dogs

Even before the advent of the internet, citizens sought to come together to form clubs, movements, societies, organisations, and associations for often non-partisan but issue-specific purposes. These non-government organisations (NGOs) often focus on broad societal issues such as development (Oxfam), environmental degradation (Greenpeace, WWF), or human rights (Amnesty International, HRC). Many of them have started to utilise the web to enhance their missions. How do people get involved? Options usually comprise: donating money, becoming a member or volunteer, join demonstrations or related events, sign petitions or letters of protest that are sent to elected constituents and politicians.

Although many NGOs transitioned from a single web *site* to a distributed web *presence* that encompasses several social media platforms, there are few examples of situated civic innovation that combine physical and digital means. Notable examples are often skewed to either the digital or the physical: Movember is a movement encouraging men to grow a moustache in the month of November to raise awareness and funds to help combat prostate cancer and depression in men. The moustache turns into a physical symbol, and social media is used to assist the campaign. The Leukaemia Foundation's *Shave for a Cure* as well as the ubiquitous red ribbon worn on 1 Dec for World AIDS Day present similar examples. A digital case was a red variation of the Human Rights Campaign (HRC) logo being adopted by millions of Facebook users in 2013 as their profile picture in – turning news feeds into a bold statement in support of same-sex marriage. At the time of writing, a rainbow of Facebook user profiles has just burst forth to celebrate the US decision to legalise same sex marriage.

New Kids on the Block

The above NGOs tend to focus on large-scale issues and so-called 'wicked problems.' Digital newcomers such as change.org, getup.org, avaaz.org, one.org, have specialised in more local and regional issues using predominantly online campaigning. Offering multiple topics allows these sites to cross-fertilise their campaigns. Some have also recognised a need for DIY campaigning, allowing users to employ the power of their mobilisation capabilities to start and maintain their own campaigns, e.g., communityrun.org.

Despite progressive uses of social media, the array of what users can do remains limited, and tends to stay online-only. Is clicking a 'Like' button or sending a petition email sufficient to solve today's societal problems, or is it just slacktivism (Lee & Hsieh, 2013; Rotman et al., 2011)? How could NGOs and citizens employ civic media to not only raise awareness, but also participate and engage in new forms of polity and civics?

Urban Guerrillas

Ideas and inspiration may be drawn from what futurists call 'weak signals,' that is, innovative small-scale movements or sub-cultures with the potential to grow and mainstream. Ad hoc gatherings in public places such as dancers in 'flashmobs' and cyclists in 'Critical Mass' use both realms of the digital (to organise, document, scale up) and the physical (to gather, perform, create spectacle). Related hybrid examples of situated engagements that are being assisted by digital means include seed bombing, permablitz, guerrilla knitting, Park(ing) Day, Dîner en Blanc, parkour, and various artistic and media performances and installations (Caldwell, Osborne, Mewburn, & Crowther, 2015) (Figure 8).

Figure 8: parkour, yarn and seed bombing, dîner en blanc

The crafting of place (or DIY placemaking or participatory city making) is a concept that encompasses a range of such urban interventions for the purpose of appropriating public spaces to assist in civic engagement, the communication of often political messages, or to simply improve the quality and experience of a place. Iveson (2013) proposes DIY urbanism as a link between the small actions and appropriations of urban space such as, guerrilla gardening, parkour, and graffiti, into a larger understanding or vision that affects the socio-cultural and civic experience of cities. What links these small actions is that urban residents imagine and create a tailored city within the city by occupying or transforming urban spaces through the injection of new meanings and functions. These inhabitants are motivated by their own purposes and often operate at the fringes or even outside existing policies and laws; they take action upon their rights as inhabitants of the city (Foth et al., 2016).

SYNTHESIS

We have unpacked and analysed some of the perspectives and issues in the related fields of social media, media architecture, and civic activism. We will now try to synthesise these findings in order to combine them to reconnect with larger theories and debates in the field. In order to do so, we briefly refer to three previous examples of categorisations that also sought to make sense of the field of HCI.

Grudin (2005) divided the development of HCI into three "faces", which roughly correspond with technical development stages at the time. Similarly, Bødker (2006) coined the notion of "waves" to articulate three distinct successions, from main frames, to CSCW, to now ubiquitous computing. Finally, and famously, the 2007 alt.chi paper by Harrison, Tatar, and Sengers (2007) review three paradigms of research in HCI, that is, human factors, classical cognitivism and information processing, and phenomenologically situated approaches. Figure 9 illustrates these three popular categorisations, yet, without making any claims whether each 'phase' correspond with their respective neighbours. For the purpose of our argument, we are more interested whether the aforementioned trends we analysed, constitute enough of a rationale to question whether HCI is now on the cusp of another wave, face, or paradigm, and if so, why, and what does it entail? One goal of this paper is to ask just that. Without pre-empting a definite answer, we will use some of the insights from our analyses in order to cast light on these questions.

Conventional dichotomies and simplistic divisions such as local / global, online / offline, private / public, large / small, mobile / static, have been created in the past to describe some of the qualities and characteristics of interfaces and their usages. However, under scrutiny and closer examination, they increasingly lose their analytical relevance, as more and more scholars recognise that the black and white nature of these dichotomies does not adequately represent the fluid and agile capabilities of many interactive applications, interfaces, and devices.

Applying mediation theory as a lifebuoy to escape the dichotomy conundrum, Verbeek (2014) suggests to embrace shades of gray, namely, *hybridity*, or more nuanced "trans-scalar" means of interaction design (Tripodi cited in Foth, Fischer, & Satchell, 2013), that is, ways in which new and emerging forms of interaction provide a range functionality, without being locked into either end of a scale. In the following, we will discuss three examples.

	Waves (Bødker, 2006)	**Faces** (Grudin, 2005)	**Paradigms** (Harrison et al., 2007)
1	Main Frame	Computer Operation	Human Factors
2	PC at Work / CSCW	Information Systems Management	Classical Cognitivism / Information Processing
3	Ubiquitous Computing	Discretionary Use	Phenomenologically Situated

Figure 9: Waves, Faces and Paradigms of HCI

Digitisation vs Fabrication

The last two or three decades of computing venturing out into every aspect of everyday life has arguably heavily focused on digitisation – atoms becoming bits. This trend brought about an avalanche of new opportunities and challenges, such as participatory culture, new forms of (digital) literacies, new frontiers in law and intellectual property, and even changes to the way many people perceive the pace of time, and the division between work and leisure.

Similarly, fabrication – the reverse process of turning bits back into atoms – will bring about many opportunities and challenges of similar magnitude. Maker cultures and the DIY (Do-It-Yourself) and DIWO (Do-It-With-Others) movements are well underway, and in tandem with technological progress and innovation with regards to fabrication devices such as 3D printers, will see profound shifts and change in many aspects of society. For example, web services, such as the eMachineShop.com, allow anyone to use CAD software, upload their design, and create custom metal and plastic parts that are shipped to their home address the next day.

It is imperative for HCI designers to closely follow and keep abreast of these developments. Reminiscent of the original *raison d'être* of HCI that called for an interdisciplinary exchange between the domains of engineers, computer scientists, human factors, and psychologists, it is now time to take stock of where we are, break out of the HCI silo that we ourselves have created, and renew the vows of interdisciplinarity by reaching out to new disciplinary frontiers. Those that seem pertinent to the quest of fabrication appear to include architecture, urban design, chemistry, and material sciences.

You vs We

If TIME Magazine's 2006 Person of the Year was 'You,' perhaps we should also ask ourselves what we, as interaction designers and computing professionals can do to change this image that is all too often attached to

Wellman's notion of networked individualism (Wellman, 2001, 2002) and often construed as egocentric and selfish. If TIME Magazine's 2016 Person of the Year was to be 'We,' what would that entail, and how do we get there?

Some promising ideas have already been proposed. First, there is Schuler's long standing advocacy for community intelligence and civic intelligence (Schuler, 2008, 2009), as well as his proposal for the world citizen parliament (Schuler, 2013). Second, proposing agonistic design, Ehn (2014) suggests a framework that interrelates innovation, design, and democracy. And third, DiSalvo (2012) suggests designerly ways to break out of the aforementioned echo chambers by questioning conventional approaches to political issues. That there is an urgent need to closely examine the ways that technology design can provoke and engage the political is illustrated by the innovative practices of the Occupy movement (Figure 1), which still fall on death ears in the way that the old chambers of power are set up (Figure 2).

Usability vs 'Citizen-ability'

Not everyone will welcome technological innovation (Selwyn, 2003). As technology becomes more pervasive, the increasing digitisation of our cities has the potential to alienate and disenfranchise citizens. This is especially true with urban interfaces and civic media that expose their message to all who pass by. The ubiquity of these technologies means that all passers-by – both users and non-users – should be considered stakeholders whose needs should influence design decisions. Therefore, rather than contextualising the design for the needs of the technically savvy end users, we have to aim to also incorporate the needs of 'non-users' into the design process (Satchell & Dourish, 2009).

Both Baumer et al. (2014), and Satchell and Dourish (2009) draw attention to the non-use of social media and computing technology. An empirically grounded understanding of the way in which urban interfaces and civic media are understood by their users and non-users alike, is required. Such interfaces can become an opportunity space that provides a platform to help disseminate community and civic information to non-users of social media who otherwise might not have access or be exposed to such information. Therefore, urban interfaces, through their very ubiquity, become a valuable platform for engaging the non-user of mainstream and social media.

In addition to the issue of non-use, there is also the question we asked at the start of this paper, that is, what comes after use and usability? A search across the ACM Digital Library for 'beyond usability' finds 10 articles on the topic, ranging from safety, intuitive use, emotions, to robotics. For the purpose of our argument, the paper by Huh et al. (2007) offers the most relevant perspective, as it asks HCI designers to take *"social, situational, cultural, and other contextual factors into account."* One of these factors that is close to our heart for the purpose of designing community and civic engagement, is citizenship, or our oddly formed term: 'citizen-ability.' What this neologism is trying to convey is the need for HCI designers to think of ways that our expertise, skills, and the infamous 'design thinking' can be employed to designing for new polities and civics.

CONCLUSIONS

It may be apt to briefly speculate about the future before concluding this paper. First, leading with the main point of our argument, we want to re-think the usefulness and merits of the usability fetish, and explore alternatives. We propose 'citizen-ability' as such an alternative, yet, it has to still be formed and shaped into a more mature and tested framework, ideally by design-in-use.

Second, we further speculate about the balance of power, and welcome the addition of the Fifth Estate (Dutton, 2009). However, our analysis of key trends indicates that the might of the Fifth Estate will greatly increase once a coalition between the digital and the physical, between the internet and the city, has been formed. De Waal (2014) postulates 'the city as interface' and we concur with him. A Sixth Estate even?

Third, what usability is for designers, is the growth imperative for economists and business, and this, we feel, has to change, too, as it is not sustainable. With all the advances in technological innovation, automation, and societal progress, we often end up being rushed, having less time for family and friends, working longer hours and wait longer until we are able to retire, we suffer from stress and other preventable diseases, and are less happy overall. This is not only ironic, but a disappointing report card for humanity. Since designers are at the core of so many human interfaces, perhaps it is time to subscribe to *happiness and wellbeing* as the primary KPIs over and above growth, financial prosperity, and – usability.

In summary, we argue that it may be useful to apply the lessons learnt from the way situated civic engagement is enacted by the urban guerrillas, to the so far mainly online focussed campaigning efforts of civil society organisations. However, the more pressing and challenging issue is to find new ways to expand the toolbox that is available to citizens to take action and bring about change (DiSalvo, 2012; Dourish, 2010). Can we offer more options than the usual array of petitioning, protesting, volunteering, and donating? Or, how can we improve the way that they are performed? How can interaction designers use their expertise, skills and craft to make a contribution to better the connection, the exchange, and the dialogue between community advocacy and activism on the one hand, and polity, governance, politics, civics, and decision making instruments on the other hand? How do we exploit and influence the role that new technology plays in this context, such as mobile devices, next generation screens, gestural and human-brain interfaces, fancy watches, and augmented reality glasses?

Interaction designers collectively created the tools that helped transform consumers turn into 'produsers' (Bruns, 2008). Let's apply these skills to augment notions of citizenship with that of *citizenability*.

REFERENCES

Aiello, L. M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., & Menczer, F. (2012). Friendship prediction and homophily in social media. *ACM Trans. Web, 6*(2), 1-33. doi:10.1145/2180861.2180866

Arnold, M. (2003). Intranets, Community, and Social Capital: The Case of Williams Bay. *Bulletin of Science, Technology & Society, 23*(2), 78-87.

Bannon, L. (1992). From human factors to human actors: the role of psychology and human-computer interaction studies in system design *Design at work* (pp. 25-44): L. Erlbaum Associates Inc.

Baumer, E. P. S., Ames, M. G., Brubaker, J. R., Burrell, J., & Dourish, P. (2014). *Refusing, limiting, departing: why we should study technology non-use*. Proc. CHI '14 Extended Abstracts on Human Factors in Computing Systems, Toronto, Ontario, Canada.

Bødker, S. (2006). *When second wave HCI meets third wave challenges*. Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles, Oslo, Norway.

Bruns, A. (2008). *Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage*. New York, NY: Peter Lang.

Brynskov, M., Dalsgaard, P., & Halskov, K. (2013). Understanding Media Architecture (Better): One Space, Three Cases *Proc. of the Workshop on Interactive City Lighting, CHI 2013*. New York: ACM.

Burgess, J., Foth, M., & Klaebe, H. (2006, Sep 25-26). *Everyday Creativity as Civic Engagement: A Cultural Citizenship View of New Media.* Paper presented at the Communications Policy & Research Forum, Sydney, NSW.

Caldwell, G. A., & Foth, M. (2014). *DIY media architecture: open and participatory approaches to community engagement*. Proceedings of the 2nd Media Architecture Biennale Conference: World Cities, Aarhus, Denmark.

Caldwell, G. A., Osborne, L., Mewburn, I., & Crowther, P. (2015). Guerrillas in the [Urban] Midst: Developing and Using Creative Research Methods – Guerrilla Research Tactics. *Journal of Urban Technology, 22*(3), 21-36. doi:10.1080/10630732.2015.1040288

Castells, M. (2000). *The Rise of the Network Society* (2nd ed.). Oxford, UK: Blackwell Publishers.

Dade-Robertson, M. (2013). Architectural User Interfaces: Themes, Trends and Directions in the Evolution of Architectural Design and Human Computer Interaction. *International Journal of Architectural Computing, 11*(1), 1-20. doi:10.1260/1478-0771.11.1.1

de Waal, M. (2014). *The City as Interface: How New Media Are Changing the City.* Rotterdam, NL: NAi010 Publisher.

Dhanarajan, G. (2001). Distance Education: Promise, Performance and Potential. *Open Learning, 16*(1), 61-68.

DiSalvo, C. (2012). *Adversarial Design*. Cambridge, MA: MIT Press.

Dourish, P. (2010). HCI and environmental sustainability: the politics of design and the design of politics *Proceedings of the 8th ACM Conference on Designing Interactive Systems* (pp. 1-10). Aarhus, Denmark: ACM.

Dourish, P., & Bell, G. (2011). *Divining a digital future: Mess and mythology in ubiquitous computing*. Cambridge, MA: MIT Press.

Dutton, W. H. (2009). The Fifth Estate Emerging through the Network of Networks. *Prometheus, 27*(1), 1-15. doi:10.1080/08109020802657453

Ehn, P. (2014). *Utopias lost and futures-in-the-making: marginal notes on innovation, design and democracy*. Proceedings of the 13th Participatory Design Conference, Volume 2, Windhoek, Namibia.

Fischer, P. T., Zollner, C., Hoffmann, T., Piatza, S., & Hornecker, E. (2013). Beyond information and utility: Transforming public spaces with media facades. *IEEE Computer Graphics and Applications, 33*(2), 38-46. doi:10.1109/mcg.2012.126

Fortin, C., Neustaedter, C., & Hennessy, K. (2014). *The appropriation of a digital "speakers" corner: lessons learned from the deployment of mégaphone*. Proceedings of Designing Interactive Systems, Vancouver, BC, Canada.

Foth, M. (Ed.) (2009). *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*. Hershey, PA: IGI Global.

Foth, M., & Brynskov, M. (2016, in press). Participatory Action Research for Civic Engagement. In E. Gordon & P. Mihailidis (Eds.), *Civic Media: Technology, Design, Practice*. Cambridge, MA: MIT Press.

Foth, M., Brynskov, M., & Ojala, T. (Eds.). (2016). *Citizen's Right to the Digital City: Urban Interfaces, Activism, and Placemaking*. Singapore: Springer.

Foth, M., Choi, J. H.-j., Bilandzic, M., & Satchell, C. (2008, Oct 7-9). *Collective and Network Sociality in an Urban Village.* Proc. MindTrek, Tampere, Finland.

Foth, M., Choi, J. H.-j., & Satchell, C. (2011). Urban Informatics *Proceedings CSCW '11*. New York: ACM.

Foth, M., Fischer, F., & Satchell, C. (2013). From movie screens to moving screens: Mapping qualities of new urban interactions. In J. Geiger, O. Khan, & M. Shepard (Eds.), *Proceedings of MediaCity 4* (pp. 194-204). Buffalo, NY: University at Buffalo.

Foth, M., Forlano, L., Satchell, C., & Gibbs, M. (Eds.). (2011). *From Social Butterfly to Engaged Citizen: Urban Informatics, Social Media, Ubiquitous Computing, and Mobile Technology to Support Citizen Engagement*. Cambridge, MA: MIT Press.

Foth, M., Hudson-Smith, A., & Gifford, D. (2016, in press). Smart Cities, Social Capital, and Citizens at Play: A Critique and a Way Forward. In F. X. Olleros & M. Zhegu (Eds.), *Research Handbook on Digital Transformations*. Cheltenham, UK: Edward Elgar.

Foth, M., Parra Agudelo, L., & Palleis, R. (2013, Sep 8). Digital Soapboxes: Towards an Interaction Design

Agenda for Situated Civic Innovation. In HiCUE 2013: Human Interfaces for Civic and Urban Engagement, UbiComp Adjunct Proceedings, Zürich, Switzerland. New York: ACM, pp. 725-728. doi: 10.1145/2494091.2495995

Gordon, E., & de Souza e Silva, A. (2011). *Net Locality: Why Location Matters in a Networked World*. Chichester, UK: Wiley-Blackwell.

Grudin, J. (2005). Three Faces of Human-Computer Interaction. *IEEE Ann. Hist. Comput., 27*(4), 46-62. doi:A68e0687-6a30-4c46-9df7-7e0a4ae9fe20

Gurstein, M. (Ed.) (2000). *Community Informatics: Enabling Communities with Information and Communication Technologies*. Hershey, PA: IGI.

Hagel, J., & Armstrong, A. G. (1997). *Net Gain: Expanding Markets Through Virtual Communities*. Boston: Harvard Business School Press.

Harrison, S., & Dourish, P. (1996). *Re-Place-ing Space: The Roles of Place and Space in Collaborative Systems*. Proc. CSCW conference, Boston, MA.

Harrison, S., Tatar, D., & Sengers, P. (2007). The Three Paradigms of HCI. *Proceedings of CHI, alt.chi*. New York: ACM.

Hearn, G., Mandeville, T. D., & Anthony, D. (1998). *The communication superhighway: social and economic change in the digital age*. Sydney: Allen & Unwin.

Hirsch, T. (2011). More than Friends: Social and Mobile Media for Activist Organizations. In M. Foth, L. Forlano, C. Satchell, & M. Gibbs (Eds.), *From Social Butterfly to Engaged Citizen: Urban Informatics, Social Media, Ubiquitous Computing, and Mobile Technology to Support Citizen Engagement*. Cambridge, MA: MIT Press.

Huh, J., Ackerman, M. S., Erickson, T., Harrison, S., & Sengers, P. (2007). *Beyond usability: taking social, situational, cultural, and other contextual factors into account*. Proc. CHI '07 Extended Abstracts on Human Factors in Computing Systems, San Jose, CA, USA.

Iveson, K. (2013). Cities within the City: Do-It-Yourself Urbanism and the Right to the City. *International Journal of Urban and Regional Research, 37*(3), 941-956. doi:10.1111/1468-2427.12053

Jenkins, H. (2006). *Fans, Bloggers, and Gamers: Exploring Participatory Culture*. New York: New York University Press.

Kim, A. J. (2000). *Community Building on the Web: Secret Strategies for Successful Online Communities*. Berkeley, CA: Peachpit.

Kling, R., Rosenbaum, H., & Sawyer, S. (2005). *Understanding and Communicating Social Informatics*. Medford, NJ: Information Today, Inc.

Lee, Y.-H., & Hsieh, G. (2013). *Does slacktivism hurt activism?: the effects of moral balancing and consistency in online activism*. Proceedings of CHI, Paris, France.

Lotan, G. (2014). Israel, Gaza, War & Data: Social networks and the art of personalizing propaganda. Retrieved from https://medium.com/i-data/israel-gaza-war-data-a54969aeb23e

Odendaal, N. (2006). Towards the Digital City in South Africa: Issues and Constraints. *Journal of Urban Technology, 13*(3), 29-48.

Pariser, E. (2011). *The filter bubble : what the Internet is hiding from you*. New York: Penguin Press.

Pons-Novell, J., & Viladecans-Marsal, E. (2006). Cities and the Internet: The end of distance? *Journal of Urban Technology, 13*(1), 109-132.

Preece, J. (2000). *Online communities: designing usability, supporting sociability*. Chichester: John Wiley.

Putnam, R. D. (2000). *Bowling Alone: the Collapse and Revival of American Community*. New York: Simon & Schuster.

Rheingold, H. (2002). *Smart Mobs: The Next Social Revolution*. Cambridge, MA: Perseus.

Rotman, D., Sarah Vieweg, Sarita Yardi, Ed Chi, Jenny Preece, Ben Shneiderman, Peter Pirolli, and Tom Glaisyer. 2011. From slacktivism to activism: participatory culture in the age of social media. In CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11). ACM, New York, NY, USA, 819-822. doi: 10.1145/1979742.1979543.

Satchell, C., & Dourish, P. (2009). Beyond the User: Use and Non-Use in HCI. *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group (OZCHI '09)* (pp. 9-16). New York: ACM.

Schroeter, R., Foth, M., & Satchell, C. (2012). People, content, location: Sweet spotting urban screens for situated engagement. *Proceedings of the Designing Interactive Systems Conference* (pp. 146-155). Newcastle Upon Tyne, UK: ACM.

Schuler, D. (2008). *Liberating voices: A pattern language for communication revolution*. Cambridge, MA: MIT Press.

Schuler, D. (2009, Jun 25-27). *Communities, Technology, and Civic Intelligence.* Paper presented at the 4th International Conference on Communities and Technologies, University Park, PA.

Schuler, D. (2013). Creating the world citizen parliament: seven challenges for interaction designers. *Interactions, 20*(3), 38-47. doi:10.1145/2451856.2451867

Selwyn, N. (2003). Apart from technology: understanding people's non-use of information and communication technologies in everyday life. *Technology in Society, 25*(1), 99-116.

Sliwa, M., & Cairns, G. (2007). Exploring Narratives and Antenarratives of Graffiti Artists: Beyond Dichotomies of Commitment and Detachment. *Culture and Organization, 13*(1), 73-82. doi: 10.1080/14759550601167321

Townsend, A. M. (2013). *Smart cities: big data, civic hackers, and the quest for a new utopia*. New York: W.W. Norton & Company.

Verbeek, P.-P. (2014). *On Icarus' Wings: Craft and the Art of Hybridization*. Keynote presented at the Designing Interactive Systems (DIS) conference, Vancouver, BC.

Waal, M. d. (2014). *The city as interface : how digital media are changing the city*. Rotterdam: nai010 publishers.

Walmsley, D. J. (2000). Community, Place and Cyberspace. *Australian Geographer, 31*(1), 5-19.

Wellman, B. (2001). Physical Place and Cyberplace: The Rise of Personalized Networking. *International Journal of Urban and Regional Research, 25*(2), 227-252.

Wellman, B. (2002). Little Boxes, Glocalization, and Networked Individualism. In M. Tanabe, P. van den Besselaar, & T. Ishida (Eds.), *Digital Cities II: Second Kyoto Workshop on Digital Cities* (Vol. LNCS 2362, pp. 10-25). Heidelberg, Germany: Springer.

Williams, A., Robles, E., & Dourish, P. (2009). Urbane-ing The City: Examining and Refining The Assumptions Behind Urban Informatics. In M. Foth (Ed.), *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City* (pp. 1-20). Hershey, PA: IGI Global.

Wood, P., & Landry, C. (2007). *The Intercultural City: Planning for Diversity Advantage*. London: Earthscan.

THE CONVERSATION

Australia needs an innovation 'skunkworks'

December 1, 2015 7.15pm GMT

John Howland and Dr Mark Bilandzic, winners of the Digital Media mashup award in the Libraryhack 2011 at The Edge, State Library of Queensland. Libraryhack, Author provided

Marcus Foth

Marcus Foth is a Friend of The Conversation.

Professor, Urban Informatics, Queensland University of Technology

The Conversation's partners

The Conversation UK receives funding from Hefce, Hefcw, SAGE, SFC, RCUK, The Nuffield Foundation, The Ogden Trust, The Royal Society, The Wellcome Trust, Esmée Fairbairn Foundation and The Alliance for Useful Evidence, as well as sixty five university members.

View the full list

Malcolm Turnbull has been heralded as the new "innovation PM". Expectations are high that he must now translate his rhetoric around agility, disruption, entrepreneurship into concrete economic policies.

Both Glenn Withers, Professor of Economics at Australian National University, and myself have argued that we need not just STEM (Science, Technology, Engineering and Mathematics), but also researchers from the social sciences, arts, design and the humanities contributing to innovation.

Withers argues:

> *The focus on STEM education in recent years is very important. But it must be equally complemented by humanities and social science enhancement. The latter disciplines are necessary to understand ourselves and the cultures and societies in which Australia wants to operate and engage, and to build creative and cultural industries.*

Several commentators have called for better support of innovation, such as Mark Dodgson, Director, Technology and Innovation Management Centre, The University of Queensland, Tony Peacock, Chief Executive of the Cooperative Research Centres Association, Glyn

Davis, Vice-Chancellor of the University of Melbourne, and Jenny Stewart, Professor of Public Policy, UNSW Australia.

Their recommendations centre around three priorities:

tackle the ailing rates of collaboration between researchers and industry in Australia, which are the worst in the OECD

support Universities Australia's Keep It Clever campaign calling for increased public and private investment into our universities

widen the scope and scale of commercialisation of research intellectual property that will provide fertile ground for start-ups emerging from university-led R&D.

Stand-up before start-up, but where?

However, the Discovery, Incubation and Acceleration (D-I-A) model that underpins many innovation systems is often plagued by the "valley of death". Institutional forms of discovery tend to be risk-averse compromising the full capacity to innovate. And the strenuous leap from discovery into incubation is currently for the few, not the many.

In addition to better start-up support programs, Indian Prime Minister Narendra Modi in August launched the Start-Up India, Stand-Up India campaign. This echoes policy recommendations by the Foundation for Young Australians to first "promote attitudes and skills for mobilizing entrepreneurs".

Right now, the public debate tends to be narrowly focused on institutionalised innovation in industry and universities. And it is constrained by utopian tech visions trying to import success from abroad in a copy and paste mentality: Silicon Valley in California; Silicon Alley in New York; and Silicon Roundabout in London.

Tan Yigitcanlar, Associate Professor of Urban and Regional Planning, Queensland University of Technology, argues:

> *In Australia, the Australian Technology Park in Sydney, Parkville Knowledge Precinct in Melbourne, and Kelvin Grove Urban Village in Brisbane are certainly emerging urban knowledge precincts.*

But Yigitcanlar also rightly asks: "Why confine innovation to a precinct?" The trend is to scale up, from centre to campus to precinct to district to city to region to nation.

Yet, in addition to challenges of growth, Australia should not forget the mechanisms of creativity and innovation at play at and between each scale. This is an ecology.

Highly specialised, secure, white lab coat innovation centres such as the Translational Research Institute in Brisbane, are important. But they don't help everyday Australians to stand up. This requires sandboxes, tinkering spaces, experimental and messy studios, garages and workshops where people from all walks of life come together to create and innovate.

Australia needs a 'skunkworks'

The term I like best to describe these open and accessible spaces where creativity and innovation can stand up is Skunkworks. I use it to refer to those types of space that attract, house, support and unleash innovators, makers, thinkers and doers to stand up.

Some of the great examples overseas that we currently study include The Old Truman Brewery in London, which offers a blended experience of hospitality, exhibition and incubator spaces in a heritage precinct. Game developers pop over from Campus London and mix with fashion designers organising an ad hoc catwalk show at night. They are surrounded by the best baristas and food trucks – not casinos.

Perhaps this is what Withers meant when he suggested we use “our liveable cities to attract and retain talent”. Cities are smart when they enable the smart citizens.

In our research of innovation spaces, we found more than just the institutional spaces of discovery and incubation that are currently in the spotlight. Great examples of spaces for experimentation and civic innovation include CityStudio Vancouver, New Urban Mechanics in Boston, and the Urban Innovation Centre in London.

Creativity and innovation are abundant in those “in between spaces” such as coworking spaces, hacker spaces, maker spaces and living labs.

Michael Doneman is the Director of Edgeware Creative Entrepreneurship and one of the master minds behind the revival of The Old Ambulance Station in Nambour as “a creative space for creative businesses”. He sees a convergence of these spaces occurring that he calls Bauhaus 2.0.

In so many ways, it is libraries that are leading the charge, evolving into spaces for incubation and innovation. The Edge at the State Library of Queensland (SLQ) in Brisbane is being complemented by a new Business Studio.

A prime example of the new Bauhaus 2.0, SLQ now offers coworking space and tinkering and experimental programs and workshops that foster connected learning with a clear (but optional) pathway into business incubation.

The skunkworks spaces in Australia’s innovation system should avoid entrepreneurial gatekeepers and one-track tech minds, encourage interdisciplinary contributions from outside STEM, and embrace the experimental messiness of the creative imagination that is the backbone of innovation.

These spaces should be free, open and ubiquitous. And they should recognise and support other forms of innovation that may not lead to a conventional start-up, such as social innovation, civic innovation, and change for good.

Innovation
innovation hubs
Silicon Valley
STEM
Start-ups
Why innovation matters

October 2014

Debating the Sharing Economy

Juliet Schor

The “sharing economy” has attracted a great deal of attention in recent months. Platforms such as Airbnb and Uber are experiencing explosive growth, which, in turn, has led to regulatory and political battles. Boosters claim the new technologies will yield utopian outcomes—empowerment of ordinary people, efficiency, and even lower carbon footprints. Critics denounce them for being about economic self-interest rather than sharing, and for being predatory and exploitative. Not surprisingly, the reality is more complex. This essay, based on more than three years of study of both non-profit and for-profit initiatives in the “sharing economy,” discusses what’s new and not so new about the sector and how the claims of proponents and critics stack up. While the for-profit companies may be “acting badly,” these new technologies of peer-to-peer economic activity are potentially powerful tools for building a social movement centered on genuine practices of sharing and cooperation in the production and consumption of goods and services. But achieving that potential will require democratizing the ownership and governance of the platforms.

Introduction

Earlier this year, hundreds of people gathered in a downtown San Francisco venue to celebrate—and debate—the "sharing economy." The term covers a sprawling range of digital platforms and offline activities, from financially successful companies like Airbnb, a peer-to-peer lodging service, to smaller initiatives such as repair collectives and tool libraries. Many organizations have been eager to position themselves under the "big tent" of the sharing economy because of the positive symbolic meaning of sharing, the magnetism of innovative digital technologies, and the rapidly growing volume of sharing activity.

While boosterism has been the rule in this sector, a strong contingent at the conference questioned whether the popular claim that the sharing economy is fairer, lower-carbon, and more transparent, participatory, and socially-connected is anything more than rhetoric for the large, monied players. Janelle Orsi, an activist lawyer, opened with a provocative challenge: "How are we going to harness the sharing economy to spread the wealth?" The Airbnbs of the world and their venture capitalist backers are siphoning off too much value, she and others argued. Discussions of labor exploitation, race to the bottom dynamics, perverse eco-impacts, unequal access for low-income and minority communities, and the status of regulation and taxation engaged attendees throughout the next two days.

Will the sector evolve in line with its stated progressive goals or devolve into business as usual?

Over the last year, these and related debates have been raging within and outside the sharing community. Will the sector evolve in line with its stated progressive, green, and utopian goals, or will it devolve into business as usual? This moment is reminiscent of the early days of the Internet, when many believed that digital connection would become a force for empowerment. The tendency of platforms to scale and dominate (think Google, Facebook, and Amazon) offers a cautionary tale. So, too, does the history of Zipcar. Once the face of the sharing economy, it is now a sub-brand of Avis. Will other sharing platforms follow similar trajectories as they grow? Or will the sharing economy be the disruptive, world-changing innovation its proponents expect? And if it is, will it change the world for the better? It is too early for definitive answers to these questions, but important to ask them.[1]

While many of the most visible platforms in the sharing economy began in the United States, sharing has become a global phenomenon, both because of the expansion of platforms to other countries, and because the idea of sharing has caught on around the world. Platforms are proliferating throughout Europe, where cities are becoming centers of "sharing" practices. Paris, for example, has become the annual home of the "OuiShare" fest. The Arab world has a raft of new sharing innovations, Colombia has become a sharing hub in Latin America, and Seoul is a center of sharing. Last year, the government of Ecuador launched Buen Conocer, an initiative to radically reimagine the nation according to principles of sharing—open networks, open production, and an economy of the commons. While the politics of these sharing efforts differ across the globe, what is common is the desire among participants to create fairer, more sustainable, and more socially connected societies.

I became interested in the sharing economy in 2008 while I was writing a book about a transition to a small-scale, ecologically sustainable economy.[2] At that time, I predicted a decline in full-time employment, as well as the need to reduce working hours as a method of controlling carbon emissions. I proposed a new household model in which people would have diverse sources of income, and would access goods and services through varied low-cost channels. With enough of a safety net and sufficient public goods, such a world could yield greater freedom, autonomy, and quality of life. If it were able to provide decent earnings and reasonably low prices, the sharing economy could be an important component of that new model. Today, however, with the corporatization of a number of the leading players, the role of the sharing economy in a just and sustainable transition is an open question.

Coming up with a solid definition of the sharing economy that reflects common usage is nearly impossible.

It is timely to step back and take stock of what has happened and how the arguments both for and against the sharing economy stack up. Because my research has focused on the United States, this essay will do so as well, returning to the global dimensions of sharing in the conclusion. I begin with a brief review of what the sharing economy is, where it came from, and why people are participating in it. I will then consider the sharing economy's impacts on ecological well-being and social connection. I conclude with the question of whether these new technologies and practices can lead to new forms of organizing that may be part of a citizens movement for a fairer and more sustainable economy.

What is the Sharing Economy?

Coming up with a solid definition of the sharing economy that reflects common usage is nearly impossible. There is great diversity among activities as well as baffling boundaries drawn by participants. TaskRabbit, an "errands" site, is often included, but Mechanical Turk (Amazon's online labor market) is not. Airbnb is practically synonymous with the sharing economy, but traditional bed and breakfasts are left out. Lyft, a ride service company, claims to be in, but Uber, another ride service company, does not. Shouldn't public libraries and parks count? When I posed these questions to a few sharing innovators, they were pragmatic, rather than analytical: self-definition by the platforms and the press defines who is in and who is out.

Sharing economy activities fall into four broad categories: recirculation of goods, increased utilization of durable assets, exchange of services, and sharing of productive assets. The origins of the first date to 1995 with the founding of eBay and Craigslist, two marketplaces for recirculation of goods that are now firmly part of the mainstream consumer experience. These sites were propelled by nearly two decades of heavy acquisition of cheap imports that led to a proliferation of unwanted items.[3] In addition, sophisticated software reduced the traditionally high transaction costs of secondary markets, and at eBay, reputational information on sellers was crowdsourced from buyers, thereby reducing the risks of transacting with strangers. By 2010, many similar sites had launched, including ThredUp and Threadflip for apparel, free exchange sites like Freecycle and Yerdle, and barter sites such as Swapstyle.com. Online exchange now includes "thick," or dense, markets in apparel, books, and toys, as well as thinner markets for sporting equipment, furniture, and home goods.

The second type of platform facilitates using durable goods and other assets more intensively. In wealthy nations, households purchase products or hold property that is not used to capacity (e.g., spare rooms and lawn mowers). Here, the innovator was Zipcar, a company that placed vehicles in convenient urban locations and offered hourly rentals. After the 2009 recession, renting assets became more economically attractive, and similar initiatives proliferated. In transportation, these include car rental sites (Relay Rides), ride sharing (Zimride), ride services (Uber, UberX, Lyft), and bicycle sharing (Boston's Hubway or Chicago's Divvy Bikes). In the lodging sector, the innovator was Couchsurfing, which began pairing travelers with people who offered rooms or couches without payment back in 1999. Couchsurfing led to Airbnb, which has reported more than 10 million stays.[4]

Neighbordoohd-based efforts can provide people with low-cost access to goods and space.

There has also been a revival of non-monetized initiatives such as tool libraries, which arose decades ago in in low-income communities. These efforts are typically neighborhood-based in order to enhance trust and minimize transportation costs for bulky items. New digital platforms include the sharing of durable goods as a component of neighborhood building (e.g., Share Some Sugar, Neighborgoods). These innovations can provide people with low-cost access to goods and space, and some offer opportunities to earn money, often to supplement regular income streams.

The third practice is service exchange. Its origins lie in time banking, which, in the United States, began in the 1980s to provide opportunities for the unemployed.[5] Time banks are community-based, non-profit multilateral barter sites in which services are traded on the basis of time spent, according to the principle that every member's time is valued equally. In contrast to other platforms, time banks have not grown rapidly, in part because of the demanding nature of maintaining an equal trading ratio.[6] There are also a number of monetized service exchanges, such as Task Rabbit and Zaarly, which pair users who need tasks done with people who do them, although these have encountered difficulties expanding as well.

The fourth category consists of efforts focused on sharing assets or space in order to enable production, rather than consumption. Cooperatives are the historic form these efforts have taken. They have been operating in the US since the nineteenth century, although there has been a recent uptick in new ones. Related initiatives include hackerspaces, which grew out of informal computer hacking sessions; makerspaces, which provide shared tools; and co-working spaces, or communal offices. Other production sites include educational platforms such as Skillshare.com and Peer-to-Peer University that aim to supplant traditional educational institutions by democratizing access to skills and knowledge and promoting peer instruction.[7]

In what follows, I will use a number of terms, including providers, consumers, participants, and users. Consumers are those who are buying services, while providers, or suppliers, are offering them. Participants can be on either side of a transaction. Users is also often employed this way. For example, Airbnb calls hosts and guests users, but in other platforms, e.g., Lyft or Uber, users would be riders, rather than drivers. On the other hand, we have found in our research that quite a few people who are providers on a site also use it as consumers, so the distinction is often more useful for transactions than persons.

The operation and the long-term impacts of these platforms are shaped by both their market orientation (for-profit vs. non-profit) and market structure (peer-to-peer vs. business-to-peer). These dimensions shape the platforms' business models, logics of exchange, and potential for disrupting conventional businesses. Examples of each type are shown in Figure 1.

The introduction of venture capitalists into the space has changed the dynamics of sharing enterprises.

Figure 1

		Type of Provider	
		Peer to Peer	Business to Peer
Platform Orientation	Non-Profit	Food Swaps, Time Banks	Makerspaces
Platform Orientation	For-Profit	Relay Rides, Airbnb	Zipcar

While all sharing economy platforms effectively create "markets in sharing" by facilitating exchanges, the imperative for a platform to generate a profit influences how sharing takes place and how much revenue devolves to management and owners. For-profit platforms push for revenue and asset maximization. The most successful platforms—Airbnb and Uber, valued at $10 and $18 billion respectively—have strong backing from venture capitalists and are highly integrated into existing economic interests.[8] The introduction of venture capitalists into the space has changed the dynamics of these initiatives, particularly by promoting more rapid expansion.

While some of the platforms present a gentle face to the world, they can also be ruthless. Uber, which is backed by Google and Goldman Sachs, has been engaging in anti-competitive behavior, such as recruiting its competitors' drivers. While its representatives articulate a neoliberal rhetoric about the virtue of "free markets," the company is apparently hedging its bets on what "free" markets will deliver for it by hiring Obama campaign manager David Plouffe to bring some old-fashioned political capital to its defense. By contrast, many of the initiatives in the sharing space, such as tool libraries, seed banks, time banks, and food swaps, are non-profits. They do not seek growth or revenue maximization, but instead aim to serve needs, usually at a community scale.

While the for-profit vs. non-profit divide is the most important one, the divide between P2P (peer-to-peer) and B2P (business-to-peer) platforms is also significant. P2P entities earn money by commissions on exchanges, so revenue growth depends on increasing the number of trades. In contrast, B2P platforms often seek to seek to maximize revenue per transaction, as traditional businesses often do. Consider the differences between Zipcar (B2P) and RelayRides (P2P). On RelayRides, owners earn income from renting their own vehicles, choosing trades based on their needs, and setting rates and availability. Zipcar functions like an ordinary short-term car rental company. With a P2P structure, as long as there is competition, the "peers" (both providers and consumers) should be able to capture a higher fraction of value. Of course, when there is little competition, the platform can extract rents, or excess profits, regardless.

Beyond novelty and the pull of new technologies, participants tend to be motivated by economic, environmental, and social factors.

Sharing platforms, particularly non-profits that are operating to provide a public benefit, can also function as "public goods." A tool library is like a public library in many ways, although it is not organized by a government, not typically supported by public funds, and not necessarily governed by a democratic process. Many public goods have a G2P structure (government-to-peer), rather than P2P. But P2P structures can be, and frequently are, democratically organized.[9]

Why Share?

Motives for participating in the new sharing economy differ, which is not surprising given the diversity of platforms and activities. Some participants are drawn by the trendiness or novelty of the platforms. It is, however, important to recognize that the novelty about which many participants (and platforms) talk can be an expression of classism and racism. Sharing is not just a relic of pre-modern societies; such practices remain more common in working-class, poor, and minority communities. The discourse of novelty employs a false universalism that can be alienating to people who have maintained non-digital sharing practices in their daily lives.

Beyond novelty and the pull of new technologies, participants tend to be motivated by economic, environmental, and social factors.[10] Sharing economy sites are generally lower in cost than market alternatives. Particularly with P2P sites, value can be redistributed across the supply chain to producers and consumers and away from "middlemen," in part because producers' costs are lower. An Airbnb host, for example, can deliver a room more cheaply than a hotel. The platforms' fees are also lower than what established businesses extract in profits. (Airbnb's maximum fee is 15%.) Service and labor exchange platforms, whether they are time banks or for-profit platforms like Task Rabbit, extract far less value than traditional agencies that arrange child care, concierge services, or home health care aides. The platforms also allow people to earn money in ways that had not previously been safely or easily available.

Many sites advertise themselves as green and present sharing as a way to reduce carbon footprints. It is a truism among "sharers" that sharing is less resource intensive than the dominant ways of accessing goods and services (e.g., hotels, taxis, shopping malls) because of the assumed reduction in demand for new goods or facilities. The actual environmental impacts of the sites are far more complicated, however, as will be discussed in the following section.

There are almost no comprehensive studies of the ecological impact of the sharing sector.

The desire to increase social connections is also a common motivation. Many sites advertise this feature of their activities, and participants often articulate a desire to meet new people or get to know their neighbors. While heartwarming anecdotes about making new friends are plentiful, many platforms fail to deliver durable social ties. For instance, a recent study of carsharing found that the two parties to the transaction often never met on account of remote access technology.[11]

Finally, a commitment to social transformation is an important motivator. My Connected Consumption Research Team has found that many respondents emphasize the value of sharing and collaboration, and some are highly critical of capitalism, the operation of the market, and the business-as-usual economy.[12] Ideological motivation, however, varies by site, with less exhibited by earners on platforms such as Airbnb and RelayRides and more by participants in time banks and food swaps.

How Green is the Sharing Economy?

Most sharing economy websites advertise their green credentials, and many users care about their ecological impact. The ecological benefits of sharing are often seen as obvious: secondary markets reduce demand for new goods, so footprints go down. Staying in existing homes reduces the demand for new hotels just as toolsharing reduces new tool purchases. However, despite the widespread belief that the sector helps to reduce carbon emissions, there are almost no comprehensive studies of its impact. At this point, they are long overdue.

An exception is a recent study of carsharing.[13] It found a measurable reduction in greenhouse gas emissions, but only because of substantial reductions from a small fraction of households. For the majority, carsharing, by expanding access to cars, increased emissions.

The ordinary assumptions about ecological impacts are generally about the first, visible shifts made by a consumer—purchasing used products rather than new ones, or staying in a private home rather than a hotel. To assess overall ecological impacts, however, we have to consider ripple effects. What does the seller or the host do with the money earned? She may use the money to buy high-impact products. Does the appearance of a market for used goods lead people to buy more new things that they intend to sell later? If travel becomes less expensive, do people do more of it? All of these effects increase ecological and carbon footprints.

There is also the question of impacts at the level of the economy as a whole. The platforms are creating new markets that expand the volume of commerce and boost purchasing power. The larger, for-profit companies are claiming to generate substantial business and income for their providers. If so, they are likely creating economic activity that would not have existed otherwise—more travel, more private automobile rides—and not just shifting purchasing from one type of provider to another. My students and I have found that Airbnb users are taking more trips now and that the availablity of cheap ride services is diverting some people from public transportation. That means the platforms result in higher carbon emissions, because their services use energy. The companies can't have it both ways—creating new economic activity and reducing carbon emissions—because the two are closely linked.

Do these sites actually build social capital? The evidence is mixed.

Does the Sharing Economy Build Social Capital?

While the discourse of novelty in this sector is overrated, there is something new afoot: what I call "stranger sharing." Although there are exceptions (e.g., elite travelers in ancient Greece), people have historically limited sharing to within their own social networks. Today's sharing platforms facilitate sharing among people who do not know each other and who do not have friends or connections in common. Stranger sharing entails higher degrees of risk, and many of today's exchanges are quite intimate—sharing one's home or car, going into strangers' homes to do work, or eating food prepared by unknown cooks. The platforms reduce risk by posting information on users via feedback and ratings. This points to a second novel dimension—the use of digital technology to reduce transactions costs, create opportunities in real time, and crowdsource information. The uniqueness of this new sharing economy is that it mobilizes technology, markets, and the "wisdom of crowds" to bring strangers together.

Many sites in the sharing space advertise social connection as a core outcome of their activity. But do these sites actually build friendships, networks, and social trust? The evidence is mixed. Stanford sociologist Paolo Parigi and his colleagues have found that Couchsurfing does, in fact, lead to new friendships. However, the ability of the platform to create such connections, especially close ones, has declined since its inception in 2003. Users have become "disenchanted" as the relationships they form are now more casual and less durable.[14] Other studies have found that social connection can be elusive, with time bank participants expressing disappointment in the degree of social connection they gained and RelayRides users describing their interactions as "anonymous" and "sterile."[15]

The role of ratings and reputational information is at the center of questions about social capital. The conventional wisdom is that the provision of crowdsourced information on users is what leads people to feel safe about interacting in intimate ways with strangers.[16] Parigi's research, however, uncovered a paradox: the more reputational information the site provided about people, the less users formed strong bonds. Venturing into unknown territory with strangers may be more of the appeal of some sites than their ability to master a utilitarian calculus of risk and reward.

Sharing economy sites can reproduce class, gender, and racial biases and hierarchies.

Sharing economy sites can also reproduce class, gender, and racial biases and hierarchies. In our research at a food swap, my team and I found that cultural capital, a type of class privilege, limited the trades members were willing to make. Only participants with the "right" offerings, packaging, appearance, or "taste" received offers or, in some cases, even felt comfortable returning. In our time bank research, we found that some people screen potential trading partners by grammar and education, and that many highly educated people were unwilling to offer their most valuable skills (like programming or web design), preferring instead to act as amateur electricians or manual workers.[17] A recent study also reported evidence of racial discrimination among Airbnb users, finding that non-black hosts were able to charge 12% more than blacks for comparable properties.[18]

Exploiting Labor?

The debut of the sharing economy was marked by plenty of language about doing good, building social connections, saving the environment, and providing economic benefits to ordinary people. It was a feel-good story in which technological and economic innovation ushered in a better economic model. Especially in the aftermath of the financial crash, this positive narrative was hard to resist. Social activists flocked to these initiatives, hoping to piggyback on their popularity. Maybe, they thought, digital P2P platforms could be a pathway to a true grassroots, inclusive, fair, and low-impact economy.

But within a few years, and particularly since the for-profit platforms began to take large sums of outside investment from venture capitalists, the situation has become more contested. A backlash has begun, from politicians, regulators, and commentators, as well as the businesses being "disrupted" by these technologies. Local officials are investigating platforms and restricting activity. Critical articles are proliferating. Workers are organizing against some of the more aggressive platforms.

Dean Baker, a progressive economist, claims the new sharing is "largely based on evading regulations and breaking the law" and subjects consumers to a substandard, possibly unsafe product.[19] Anthony Kalamar has called out "sharewashing," in which platforms shift risk onto employees under the guise of "sharing."[20] Tom Slee, writing in *Jacobin*, has challenged Airbnb's claim that its users are single individuals earning small amounts of extra money, finding that half the revenue generated in New York City accrues to hosts with multiple listings.[21]

The central theme of the critics is that for-profit platforms have coopted what began as a progressive, socially transformative idea. Are they right? Regarding regulation, insurance, and taxation, the platforms are mobilizing political support, and, my experience suggests that they seem to be generally accepting of the idea that some regulation is necessary. Because most of the action is at the local and state level, there is a great deal of variation. But the trend seems to be towards a light regulatory touch that will allow the platforms to operate and grow.

Have for-profit platforms co-opted what began as a progressive, socially transformative idea?

There is less clarity about how the platforms are affecting labor conditions. Critics see them as architects of a growing "precariat," a class on the precarious edge of economic security, and argue that the impetus for sharing is not trust, but desperation.[22] From the perspective of drivers, errand-runners, and hosts, they describe a race to the bottom, with risk-shifting from companies to individual "micro-entrepreneurs."

Part of the difficulty in assessing the impact of these new earning opportunities is that they are being introduced during a period of high unemployment and rapid labor market restructuring. Working conditions and protections are already being eroded, real wages are declining, and labor's share of national income in the US has declined to historic lows. If the labor market continues to worsen for workers, their conditions will continue to erode, and it will not be because of sharing opportunities. Alternatively, if labor markets improve, sharers can demand more of the platforms because they have better alternatives. The two effects will work in opposite directions: with destruction of demand for legacy businesses and growth for sharing companies.

We also need to consider the diversity of industries in which sharing platforms are operating. Some sectors are characterized by high rents that are easy to capture with disruptive technologies. Consider taxis. The biggest impact is likely the erosion in the value of medallions, the licenses they must possess to operate, because these medallions yield pure rents. While drivers in conventional operations may be capturing some of this excess profit, they are already facing adverse market conditions and, in many places, earning low hourly wages, as they are forced to pay high leasing and other fees to the owners of the medallions and vehicles. Union members fare better, but could they do better with Uber? Many have switched in hopes that they can. So far, though, the results are mixed, in part because they face increasing competition from platforms like UberX and Lyft, on which drivers use their own cars. And early high returns have been reduced by Uber's fare cuts, which have led to driver protests and organizing efforts.

An online platform with a good rating system should improve labor conditions. Consider the market for home health aides, where agencies currently take an enormous fraction of hourly fees, sometimes more than half.[23] A P2P matching platform would take a lower fraction, enabling low-paid workers to earn considerably more and have more autonomy over which jobs they accept. Where owners, agencies, or other actors are extracting rents, P2P platforms should do what they claim—distribute value to consumers and producers and away from gatekeepers and rent extractors.

Ultimately, the question is about how much value providers on these platforms can capture. This depends partly on whether they can organize themselves, a question the next section will explore. But there is another dimension, which is whether there is competition among platforms. Will they come to monopolize a given space, as we have seen in the areas of search, social media, and retail (Google, Facebook, Amazon)? Or are these P2P enterprises different? What they are offering is software, insurance, ratings, and a critical mass of participants. These are functions that can be replicated. For example, if the volume of users continues to grow, then critical mass may be achievable on multiple platforms. The ratings systems are not yet very good, and there are already start-ups attempting to delink ratings from individual platforms. Insurance can also be unbundled. At the May conference, venture capitalist Brad Burnham predicted a coming round of cost-squeezing akin to the cost-squeezing that the start-ups are inflicting on legacy businesses. On the other hand, the more the platforms are backed by and integrated with the large corporations that dominate the economy, the more monopolized the sector will be, and the less likely value will flow to providers and consumers.

The question of whether providers should organize is now firmly on the table.

Organized Sharers?

An alternative to the co-optation path is one in which sharing entities become part of a larger movement that seeks to redistribute wealth and foster participation, ecological protection, and social connection. This will only happen via organization, even unionization, of users. Indeed, the question of whether providers should organize is now firmly on the table, although it is too early to know how things will evolve.[24]

Airbnb has begun to encourage its users to organize. In 2013, the global head of "community" at the company co-founded Peers.org, an attempt to build a social movement of sharers. Not long after, Airbnb created its own organizing platform for guests, hosts, and employees, which has led to the creation of numerous local groups of users who are coming together on and offline for a variety of purposes, including sharing advice and affecting public policy. The company wants these groups to push for favorable regulation. But they may develop agendas of their own, including making demands of the company itself, such as setting price floors for providers, pushing risk back onto the platforms, or reducing excessive returns to the entrepreneurs and the venture capitalists. On the labor exchanges, where the need for organization is perhaps most acute, providers could push for minimum wages.

Existing platforms could also potentially become user-governed or cooperatively owned, an outcome some voices within the community are advocating. The platforms' discourse borrows heavily from the peer production world and emphasizes the ability of these technologies to empower ordinary individuals. As many online communities have shown, the online environment can be conducive to organizing against unpopular policies, software changes, and practices. The fact that users create so much of the value in these spaces militates in favor of their being able to capture it, should they organize to do so. To date, that type of movement has not developed, but it still might.

Alternately, organizations that are part of the solidarity sector, such as unions, churches, civil society groups, and cooperatives, could create platforms for their members. They could build alternatives to the for-profits, particularly if the software to operate these exchanges is not too expensive. These platforms could be user-governed and/or owned. For example, a taxi cooperative in Portland, Oregon, has adopted the technology used by ride sharing companies and will effectively morph into a driver-owned Lyft or Uber. In general, mounting a competitive challenge to business-as-usual should be easier when production is P2P because the platform is a broker, not a producer. This is one of the reasons the sharing platforms have grown so rapidly, while efforts to create worker cooperatives have yielded so few new enterprises and jobs. In the end, though, it is not just about economics. The key to making sharing economies socially just is to emphasize an explicit politics of sharing, as well as nurturing collective, public forms of sharing.

We are at a critical juncture in which users' organizing can make a critical difference in realizing the potential of the sharing model.

Conclusion: Creating a Movement

So what are we to make of the sharing economy? There is little doubt that the pro-sharing discourse is blind to the dark side of these innovations. At the same time, the critics are too cynical. There is potential in this sector for creating new businesses that allocate value more fairly, that are more democratically organized, that reduce eco-footprints, and that can bring people together in new ways. That is why there has been so much excitement about the sharing economy. The emergence of P2P communities that share goods, space, and labor services can be the foundation of a new household model in which people are less dependent on employers and more able to diversify their access to income, goods, and services. But the early stage goodwill from the big platforms will dissipate as they become incorporated into the business-as-usual economy. We are at a critical juncture in which users' organizing for fair treatment, demands for eco-accountability, and attention to whether human connections are strengthened through these technologies can make a critical difference in realizing the potential of the sharing model. There is an enormous amount of new economic value being created in this space. It is imperative that it flow equitably to all participants. After all, that is what we ordinarily call sharing.

Ultimately, the ability of the new sharing practices to help catalyze a social transition may also depend on the form these initiatives take around the world. As the sharing economy expands in Europe, its practices are likely to be embedded in political, regulatory, and social contexts which are more attuned to the stated values of fairness, sustainability, openness, and cooperation. In Latin America, the leftward shift toward social solidarity, poverty alleviation, and democracy also suggests a context more conducive to a cooperative and community-oriented sharing movement, as we have seen in Ecuador. For those of us in countries where the pressures to commodify and concentrate value from these platforms are most intense, these developments can reveal possibilities.

Outside the US, the impetus to share in transportation, housing, foods, and goods is more integrally tied to city-level goals of carbon emission reduction, informational transparency and genuine democracy. By embedding sharing practices within those larger municipal level movements, the likelihood that the sharing movement can achieve its stated goals is greater. My hunch is that the more that US sharing activists connect with other sharers around the globe, the more success we will have in pushing the goals of eco-accountability, value distribution, and social solidarity. This also means an openness to and ideally connection with other social movements that are already active on these issues. Ultimately, a cross-fertilization could both create accountability for the sharing platforms and organizations and embed sharing practices and cooperative economic activity into the DNA of the social movements.

Now our task is to build a movement to harness that power.

The sharing economy has been propelled by exciting new technologies. The ease with which individuals, even strangers, can now connect, exchange, share information, and cooperate is truly transformative. That's the promise of the sharing platforms about which virtually everyone agrees. But technologies are only as good as the political and social context in which they are employed. Software, crowdsourcing, and the information commons give us powerful tools for building social solidarity, democracy, and sustainability. Now our task is to build a movement to harness that power.

Endnotes

1. Over the past three years, I have been conducting qualitative research on the sharing economy funded by the MacArthur Foundation (http://clrn.dmlhub.net/projects/connected-consumption). My research team has conducted seven cases: three non-profits (a time bank, a food swap, and maker-space), three for-profits (Airbnb, RelayRides, and Task Rabbit), and one hybrid (open-access education). We have done more than 150 interviews and more than 500 hours of participant observation.
2. Juliet B. Schor, *True Wealth: How and Why Millions of Americans Are Creating a Time-rich, Ecologically-light, Small-scale, High-satisfaction Economy* (New York: The Penguin Press 2011).
3. Ibid.
4. Ryan Lawler, "Airbnb Tops 10 Million Guest Stays Since Launch, Now Has 550,000 Properties Listed Worldwide," December 19, 2013, available at http://techcrunch.com/2013/12/19/airbnb-10m/.
5. Edgar Cahn and Jonathan Rowe, *Time Dollars* (Emmaus, PA: Rodale Press, 1992).
6. Emilie Dubois, Juliet Schor, and Lindsey Carfagna, "New Cultures of Connection in a Boston Time Bank," in *Practicing Plenitude,* eds. Juliet B. Schor and Craig J. Thompson (New Haven: Yale University Press, 2014).
7. It is worth noting the historical and global connections between the sharing platforms and other types of P2P activity. The collaborative software movement, which harnesses the unpaid work of software engineers to write code and solve problems collectively, paved the way for file-sharing, video posting, and crowdsourcing information, as seen in Wikipedia or citizen science. The global "commons" movement is encouraging peer production and the information commons, as well as the protection of ecological commons.
8. Andrew Ross Sorkin, "Why Uber Might Well be Worth $18 Billion," *New York Times*, June 9, 2014, available at http://dealbook.nytimes.com/2014/06/09/how-uber-pulls-in-billions-all-via-iphone/; Evelyn Rusli, Douglas MacMillan, and Mike Spector, "Airbnb Is in Advanced Talks to Raise Funds at a $10 Billion Valuation," *Wall Street Journal*, March 21, 2014, available at http://online.wsj.com/news/articles/SB10001424052702303802104579451022670668410.

9. For more on sharing and public goods, see Julian Ageyman, Duncan McLaren, and Adrianne Schaefer-Borrego, "Sharing Cities," Briefing for the Friends of the Earth Big Ideas Project, September 2013, available at http://www.foe.co.uk/sites/default/files/downloads/agyeman_sharing_cities.pdf.

10. Technophilia also spurs participation. People enjoy the sophisticated interfaces offered by many sites and like using the internet to do things quickly and easily. Many users have been "digitally primed" by years of sharing files or contributing information online.

11. Anny Fenton, "Making Markets Personal: Exploring Market Construction at the Micro Level in the Car-sharing and Time Bank Markets," Unpublished paper, Harvard University, 2013.

12. Dubois, Schor, and Carfagna, op. cit.

13. Elliott W. Martin and Susan A. Shaheen, *Greenhouse Gas Impacts of Car Sharing in North America*, Mineta Transportation Institute Report 09-11 (San Jose, CA: Mineta Transportation Institute, 2010).

14. Paolo Parigi and Bogdan State, "Disenchanting the World: The Impact of Technology on Relationships," Unpublished paper, Stanford University, n.d.

15. Dubois, Schor, and Carfagna, op. cit.; Fenton, op. cit.

16. Recent studies have found inaccuracies in ratings systems, especially the tendency to overrate positive features and under-report bad experiences. A colleague and I review recent studies in Juliet B. Schor and Connor Fitzmaurice, "Collaborating and Connecting: The Emergence of a Sharing Economy," in *Handbook on Research on Sustainable Consumption*, eds. Lucia Reisch and John Thogersen (Cheltenham, UK: Edward Elgar), 2015.

17. Juliet B. Schor et al., "Paradoxes of Openness and Distinction in the Sharing Economy," Unpublished paper, Boston College, 2014.

18. Benjamin Hardin and Michael Luca, "Digital Discrimination: The Case of Airbnb," Harvard Business School Working Papers, 2014.

19. Dean Baker, "Don't Buy the 'Sharing Economy' Hype: Aibnb and Uber Are Facilitating Ripoffs," *The Guardian*, May 27, 2014, available at http://www.theguardian.com/commentisfree/2014/may/27/airbnb-uber-taxes-regulation.

20. Anthony Kalamar, "Sharewashing is the New Greenwashing," *OpEd News*, May 13, 2013, available at http://www.opednews.com/articles/Sharewashing-is-the-New-Gr-by-Anthony-Kalamar-130513-834.html.

21. Tom Slee, "Sharing and Caring," *Jacobin Magazine*, January 24, 2014, available at https://www.jacobinmag.com/2014/01/sharing-and-caring/.

22. Kevin Roose, "The Sharing Economy Isn't About Trust, It's About Desperation," *New York Magazine*, April 24, 2014, available at http://nymag.com/daily/intelligencer/2014/04/sharing-economy-is-about-desperation.html.

23. Jane Gross, "Home Health Aides: What They Make, What They Cost," *New York Times*, December 30, 2008, available at http://newoldage.blogs.nytimes.com/2008/12/30/home-health-aides-what-they-make-what-they-cost/.

24. Uber and Lyft drivers have begun unionization efforts in various cities. Union representatives were at the San Francisco conference, and unionization has emerged as a topic of conversation among sharing innovators. This year, the National Freelancers' Union opened a benefits desk offering insurance, 401(k) plans and other benefits for "independent" laborers in a variety of companies, including Lyft.

About the Author

Juliet Schor is Professor of Sociology at Boston College and is currently serving as the Matina S. Horner Distinguished Visiting Professor at the Radcliffe Institute, Harvard University. Her PhD is in economics, which she taught at Harvard from 1984 to 1995. Schor is the author of the national best-seller *The Overworked American: The Unexpected Decline of Leisure*, *The Overspent American: Why We Want What We Don't Need*, and *True Wealth: How and Why Millions of Americans are Creating a Time-Rich, Ecologically Light, Small-Scale, High-Satisfaction Economy*. Schor's current research topics include consumer culture, sustainable lifestyles, and the relationship between time use and carbon emissions. She is a member of the MacArthur Foundation Connected Learning Research Network, for which she is leading a six-year project on the sharing economy.

About the Publication

Published as an Essay by the Great Transition Initiative.

Cite as Juliet Schor, "Debating the Sharing Economy," Great Transition Initiative (October 2014).

Available at http://greattransition.org/publication/debating-the-sharing-economy.

About the Great Transition Initiative

The Great Transition Initiative is an international collaboration for charting pathways to a planetary civilization rooted in solidarity, sustainability, and human well-being.

As a forum for collectively understanding and shaping the global future, GTI welcomes diverse ideas. Thus, the opinions expressed in our publications do not necessarily reflect the views of GTI or the Tellus Institute.

Smarter regulation for the sharing economy

Koen Frenken, Toon Meelen, Martijn Arets and Pieter van de Glind
Wednesday 20 May 2015 07.00 BST

Technology-based sharing platforms are touching more and more people's lives, but policy and regulation can struggle to cope with new technologies. We need to rethink our notion of the sharing economy in order to make better regulation

The technology-enabled sharing economy is an increasingly significant social and economic phenomenon. Increasing numbers share their homes (Airbnb), cars (Relayrides) and rides (BlaBlaCar), and people are also now sharing boats, tools, gardens, pets and clothes.

This week in Paris entrepreneurs, innovators, regulators and academics will come together at OuiShare Fest, a major three day celebration of the new sharing economy. But the sharing economy is controversial, because it can lead to moreprecarious jobs and induce unfair competition. Many of its practices are illegal in some jurisdictions. We believe that smart regulation can solve many of these current controversies. However, we first need to understand what the sharing economy is - and is not.

Defining the sharing economy

Discussions about the sharing economy lack clear definitions. This bogs public debate down in platitudes and slogans, voiced by both proponents ("you cannot stop innovation") and opponents ("this isn't capitalism, it's worse"). Two of us have previously defined the sharing economy as consumers granting each other temporary access to under-utilised physical assets ("idle capacity"), possibly for money. By parsing this definition into three elements, we can clearly distinguish the sharing economy from other economic forms.

1. Sharing is about **consumer-to-consumer** (C2C) platforms and not about renting or leasing a good from a company (business-to-consumer). In the latter case we would speak of product-service economy, where a consumer gains access to a product whilst the

service provider retains ownership. An example is car-rental (see Figure 1).

2. Sharing is about consumers providing each other **temporary access** to a good, and not about the transfer of ownership of the good. Thus, the sharing economy does not include the second-hand economy, in which goods are sold or given away between consumers (as occurs on online platforms such as Ebay or Facebook).

3. Sharing is about more efficient use of **physical assets** and not about private individuals delivering each other a service. After all, physical goods can go unused, but people cannot. internet platforms that bring consumers together to provide each other with services represent the on-demand economy. An example of such a platform is Taskrabbit, through which you can hire people to carry out work around the house.

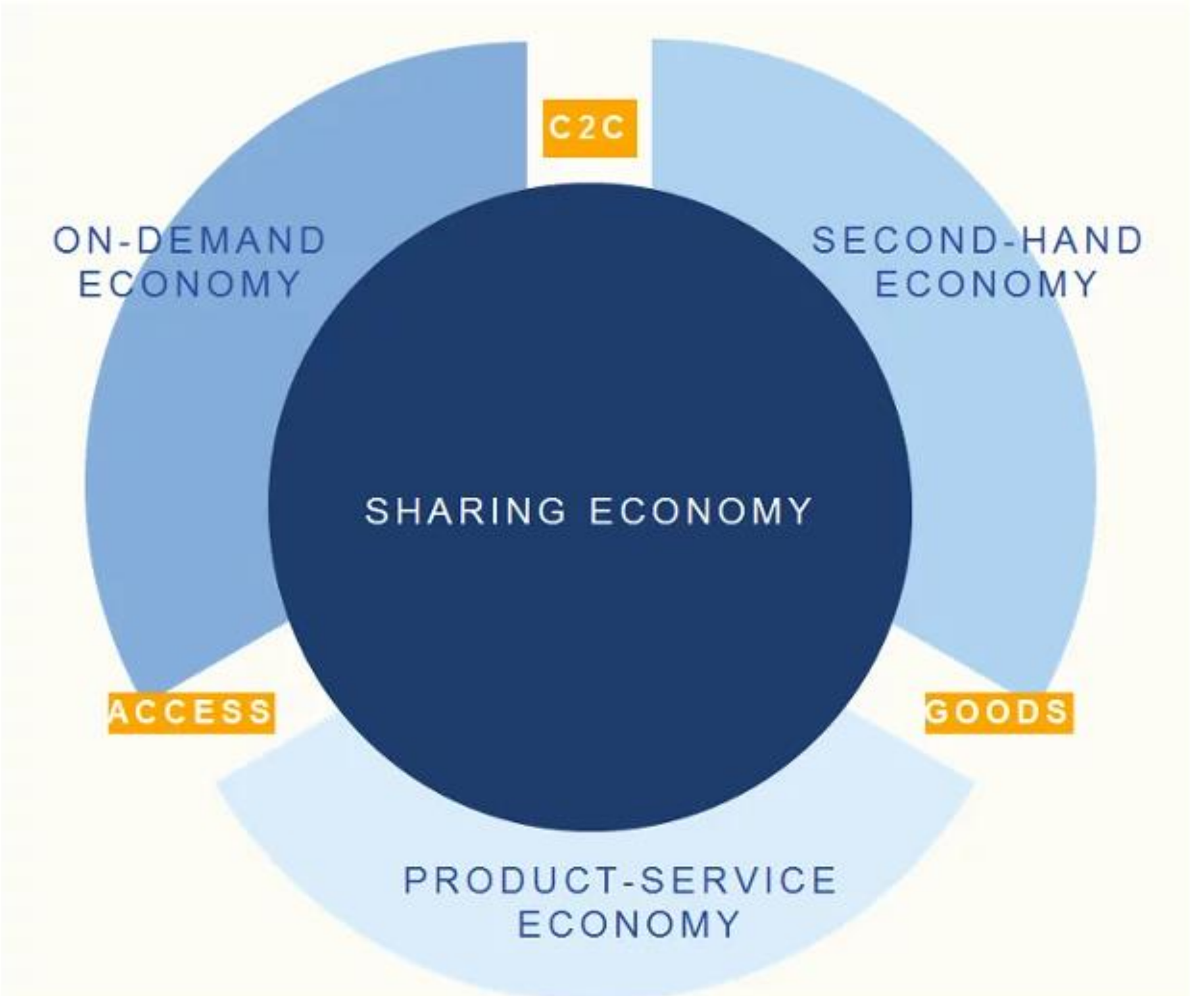


Figure 1 - The sharing economy and other related economic forms

Take the example of a drill. Consumers can avoid buying a new drill by using one out of the four platforms: you can buy a second-hand drill from an individual (Ebay), you can rent a drill from a company (Home Depot), you can hire someone on-demand to drill the hole in your wall (Taskrabbit), or you can ask an individual to share a drill (Peerby). For cars, the same logic applies. You can buy a second-hand car using a site (Ebay), you can rent a car at a car-rental company (Car2Go), you can hire on-demand an individual to drive you (UberX), or you can rent a car from a private individual (Relayrides).

Increasing criticism

There is little doubt that the sharing economy can make a contribution to a more sustainable economy. While sharing used to be limited to a small circle of family and friends, the internet platforms allow us to share with anyone in society. This means that physical assets can be used more efficiently, and less energy and materials are needed in our economy. At the same time, the sharing economy has come under fire. Critics have pointed to undesirable effects, such as platform monopolies, privacy violations, exploitation of labour, and unfair competition.

Our scheme can be used to place the various criticisms in the context in which they engage. The first two critiques (monopoly, privacy) are not specific to the sharing economy, but are problems associated with all sorts of internet-based businesses, including search engines (Google), social media (Facebook) and data storage (Dropbox). All platforms that enable interactions between people are characterized by network effects: the more people participate, the more attractive it becomes for others to join. For this reason, monopolies emerge and privacy-sensitive information gets concentrated in the hands of the businesses concerned. Concerning new forms of exploitation, many point to the increased flexibility and precariousness of work for freelancers who work for a platform (for example, UberX drivers). This criticism applies to the on-demand economy and not to the sharing economy as we have defined it.

The last criticism is that of unfair competition between the new platforms and existing businesses. Individuals providing their goods or services to others are competing with existing providers but are not necessarily bound by the same rules and regulations. Despite growing criticism, most platforms don’t seem eager to take the responsibility and to respect regulations. Notably, Uber has continued offering UberX despite court bans in various countries, and Airbnb only barely cooperates with municipalities in their aim to collect taxes and to help combat illegal hotels.

The need for smart regulation

Regulation will determine whether businesses like Uber and Airbnb will be integrated into the sharing economy or not. An illustration is the regulatory process that the Dutch municipality ofAmsterdam has initiated with Airbnb. The municipality wants to ensure that people only occasionally rent out their house whilst away (sharing economy), rather than run a permanent, unregulated hotel (not sharing economy). It has chosen to allow its residents to rent out their homes for up to 60 days per year. Customers are also supposed to pay tourist tax via Airbnb, although its enforcement leaves much to be desired (not least because Airbnb does not want to share its data with the municipality). The philosophy of the policy is important here: it is trying to regulate the rental of homes in such a way that it becomes part of the sharing economy as previously defined. Without this regulation, Airbnb would create an incentive for illegal renting with negative consequences for the local residents (higher rents, nuisance and speculation).

In the case of Uber, regulation is still largely lacking. Many options for regulating UberX are possible. The debate tends however to be narrowly focused on what requirements UberX drivers must meet compared with regular drivers - that is on regulating Uber as a provider of taxi services. But where a company like Uber presents itself so prominently as a “sharing” service, regulators could encourage the company to *organize* its services as such. In Uber’s case, that would mean doing away with the on-demand taxi model it has now and becoming a supply-driven ridesharing platform such as BlaBlaCar. On this type of platform, drivers post trips that they are going to make anyway, and other people can join these rides.

The risk of monopoly in the sharing economy is a real one. As with any other market, regulations are required to ensure sufficient competition between platforms. The key here is to empower users vis-à-vis the platform. For example, new rules are needed to let users switch easily between platforms, and to take with them their personal reviews and ratings from one platform to the other. This stops users from getting locked in in one platform which can then extract most of the value generated by platform interactions and transactions. Most of all, the ownership of data generated by users through use of sharing platforms should be a major issue of regulatory concern.

Koen Frenken is Professor of Innovation Studies at the Copernicus Institute of Sustainable Development at Utrecht University, The Netherlands. On Twitter he is@kfrenken.

Toon Meelen is a PhD candidate at the Innovation Studies group at the Copernicus Institute of Sustainable Development at Utrecht University, The Netherlands. His main research interests are the sharing economy and the diffusion of electric vehicles.

Martijn Arets is collaborative economy analyst and expedition leader at Crowd Expedition, author of 'Brand Expedition' and 'Crowdfunding, beyond the Hype', Dutch entrepreneur of 2011 and one of the '40 Young Potential European Leaders under 40'. On Twitter he is @martijnarets.

Pieter van de Glind is a co-founder of shareNL, the Dutch network and knowledge platform for the collaborative economy. On Twitter he is @P_vd_Glind.

Conceptualizing Smart City with Dimensions of Technology, People, and Institutions

Taewoo Nam & Theresa A. Pardo
Center for Technology in Government
University at Albany, State University of New York, U.S.
{tnam,tpardo}@ctg.albany.edu

ABSTRACT
This conceptual paper discusses how we can consider a particular city as a smart one, drawing on recent practices to make cities smart. A set of the common multidimensional components underlying the smart city concept and the core factors for a successful smart city initiative is identified by exploring current working definitions of smart city and a diversity of various conceptual relatives similar to smart city. The paper offers strategic principles aligning to the three main dimensions (technology, people, and institutions) of smart city: integration of infrastructures and technology-mediated services, social learning for strengthening human infrastructure, and governance for institutional improvement and citizen engagement.

Categories and Subject Descriptors
H.4.2 [**Information Systems Applications**]:
Type of systems – *e-government applications*

General Terms
Management, Performance, Human Factors, Standardization, Theory

Keywords
Smart city, Smart technology, Service integration, Infrastructure integration, Governance

1. MOTIVATION OF RESEARCH
The city, as a government unit, is growing increasingly larger, more complex and more important as the population ranks of urban areas swell with ever increasing speed. According to the United Nations Population Fund (see www.unfpa.org), 2008 marked the year when more than 50 percent of all people, 3.3 billion, lived in urban areas. By 2030 this number is expected to increase to 5 billion. With the rapid increase of the urban population worldwide, cities face a variety of risks, concerns, and problems; for example, physical risks such as deteriorating conditions in air and transportation, and economic risks such as unemployment. The unprecedented rate of urban growth creates an urgency to finding smarter ways to manage the accompanying challenges. Some cities are identified to successfully operate in a smarter way to solve concerns. Recent practices to make cities better for living have become successful cases for new city development strategies. We need to learn from the successfully progressive practices of the cities listed below or more.

Table 1. The List of Smart Cities

Region	Cities
Asia	Bangalore (India); Chongqing (China); Doha (Qatar); Gangnam District, Seoul (Korea); Hong Kong; HwaSeong-DongTan (Korea); Hyderabad (India); Ichikawa (Japan); Jaipur, Rajasthan (India); Jia Ding (China); Kabul (Afghanistan); Mitaka (Japan); Shanghai (China); Seoul (Korea); Singapore; Suwon (Korea); Taipei (Taiwan); Taoyuan County (Taiwan); Tel Aviv (Israel); Tianjin (China); Yokosuka (Japan)
Africa	Cape Town (South Africa); Nelson Mandela Bay (South Africa)
Europe	Besançon (France); Birmingham (UK); Dundee, Scotland (UK); Eindhoven (Netherlands); Glasgow, Scotland (UK); Hammarby Sjostad (Sweden); Issy-les-Moulineaux (France); Karlskrona (Sweden); Malta (Malta); Manchester (UK); Reykjavik (Iceland); Sopron (Hungary); Stockholm (Sweden); Tallinn (Estonia); Sunderland (UK); Trikala (Greece)
North America	US: Albany (New York); Ashland (Oregon); Arlington County (Virginia); Bettendorf (Iowa); Bristol (Virginia); Chattanooga (Tennessee); Cleveland (Ohio); Corpus Christi (Texas); Dakota County (Minnesota); Danville (Virginia); Dublin (Ohio); Florida High Tech Corridor; LaGrange (Georgia); Northeast Ohio; Loma Linda (California); Riverside (California); San Francisco; Spokane (Washington); Westchester County (New York); Winston-Salem (Carolina) Canada: Burlington (Ontario); Calgary (Alberta); Edmonton (Alberta); Fredericton (New Brunswick); Kenora (Ontario); Moncton (New Brunswick); Ottawa (Ontario); Quebec City (Quebec); Stratford (Ontario); Toronto (Ontario); Vancouver (British Columbia); Waterloo (Ontario); Western Valley (Nova Scotia); Windsor-Essex (Ontario); Winnipeg (Manitoba)
Middle/South America	Barceloneta (Puerto Rico); Curitiba, Paraná (Brazil); Pirai (Brazil); Porto Alegre (Brazil)
Oceania	Ballarat (Australia); Gold Coast City (Australia); Ipswich, Queensland (Australia); State of Victoria (Australia); Whittlesea, Victoria (Australia)

Source: https://www.intelligentcommunity.org/index.php?submenu=Awards&src=gendocs&ref=Smart21&category=Events&link=Smart21

Intelligent Community Forum (ICF) annually announces cities awarded as Smart21 Communities, which earns high score in terms of five successful factors to be an intelligent community

Dg.o'11, June 12–15, 2011, College Park, MD, USA.

(i.e., broadband connectivity, knowledge workforce, digital inclusion, innovation, and marketing and advocacy). Table 1 describes the cumulative list of cities (in an alphabetical order) awarded by ICF from 2007 to 2011. Practices in the cities listed up deserve attention. Quebec City in Canada was a city highly dependent upon its provincial government because of its weak industrial base until early 1990s. The city government kicked off a public-private partnership to support the growing multimedia sector and high-tech entrepreneurship. For sustainable urban growth, the City of Riverside in California is improving traffic flow and replacing aging water, sewer and electric infrastructure by tech-based transformation. Estonia overcame post-Soviet economic ruin, and its capital city Tallinn played as a center to economic development, harnessing information and communication technologies (ICTs). The city developed a large-scale digital skills training program, extensive e-government, and an award-winning smart ID card. By fostering high-tech parks, Tallinn gains about 80 percent of Estonia's foreign direct investment. Taoyuan County in Taiwan is home to the international airport. It faced powerful competition from other cities. The Aerotropolis initiative makes its economy more robust and improve the quality of living through ICTs. A series of the county government's projects has evolved from *E*-Taoyuan to *M*-Taoyuan to *U*-Taoyuan.

A common fact underlies the practices: that is, those cities are meeting a growing demand for more livable cities. The cities are being labeled with a common phrase: *smart city*. The concept of smart city is not novel, but in the recent years it has taken on a new dimension of using ICTs to build and integrate critical infrastructures and services of a city. The initiatives of making a city smart have recently emerged as a model to mitigate and remedy current urban problems and make cities better as places to live. Hence some view smart city as an icon of a sustainably livable city. Yet, so far we see academics have seldom tackled the practical concept. Considering that, we take an analytic look at the conceptual identity of smart city.

We see commentators confused between visions and basic components of smart city. While a majority of discussions present rosy visions and ideal images of smart city (e.g., smart transportation, smart mobility, smart environment, smart energy, smart safety, and so on), little research has tackled enabling factors of a smart city initiative (what really makes cities smart). Concepts and success factors of smart city have not been discussed with a comprehensive understanding. The discussion of smart city has been made without solid conceptualization.

In this sense, we recognize a research gap in the current literature of smart city. Considering the gap, we raise various conceptual questions. What are main characteristics of smart city? In what aspects do people label some particular cities as smart city? Why is smart city being recognized as a novel concept, making distinction from other similar ones? What leads to the success of a smart city initiative? This paper seeks to answer these inquiries, fill the research gap, and conceptualize smart city for both academics' and practitioners' use of that concept.

This paper aims to suggest a framework connecting conceptual variants of the smart city label, key elements for being a smart city, and strategic principles for making a city smart. The paper after this introduction is organized into five sections. Section 2 defines smart city by specifying the meanings of smartness in the urban context, exploring current definitions of smart city, and understanding smart city as a broad concept comprising its conceptual relatives. Section 3 derives prerequisites or central components of smart city from the recent literature. Section 4 discusses what strategic principles contribute to the success of smart city initiatives. The last section addresses concluding remarks.

2. DEFINING SMART CITY

The definitions of smart city are various. As the concept is being known popularly but used all over the world with different names and in different circumstances, there are a range of conceptual variants generated by replacing *smart* with other alternative adjectives. Hollands [41] recognized smart city as an "urban labeling" phenomenon, particularly in terms of what the label ideologically reveals as well as hides. The label smart city is a fuzzy concept and is used in ways that are not always consistent. There is neither a single template of framing smart city nor a one-size-fits-all definition of smart city. This section seeks to dismantle "the diversifying terrain of smart cities" [12].

2.1 The Meanings of "Smart" in the Smart City Context

Tracing the genealogy of the word *smart* in the label smart city can contribute to an understanding of how the term *smart* is being loaded. In marketing language, smartness is centered on a *user perspective* [50]. Because of the need for appeal to a broader base of community members, smart serves better than the more elitist term intelligent. Smart is more *user-friendly* than intelligent, which is limited to having a quick mind and being responsive to feedback. Smart city is required to adapt itself to the user needs and to provide customized interfaces [62].

In the urban planning field, the smartness in smart growth is treated as a *normative* claim and *ideological* dimension. Being smarter entails *strategic* directions. Governments and public agencies at all levels are embracing the notion of smartness to distinguish their new policies, strategies, and programs for targeting sustainable development, sound economic growth, and better quality of life for their citizens [19]. They associate smart with achieving policy success in their jurisdictions.

The smartness in smart technologies also merits attention. The technologies had permeated into the commercial application of intelligent-acting products and services, artificial intelligence, and thinking machines [51,66]. Smartness in the technology context implies the automatic computing principle like self-configuration, self-healing, self-protection, and self-optimization [75]. Smart homes, smart buildings, and larger smart ensembles like airports, hospitals or university campuses are equipped with a multitude of mobile terminals and embedded devices as well as connected sensors and actuators [50]. *Smart ecosystem* is a conceptual extension of smart space from the personal context to the larger community and the entire city [88].

2.2 Working Definitions of Smart City

Table 1 presents working definitions of smart city. Washburn *et al.* [80] conceptualizes smart city by laying an explicit emphasis on the use of smart computing technologies. They viewed current urban crises as an imperative of a smart city initiative. Deteriorating conditions of cities in a crisis include scarcity of resources, inadequate and poor infrastructure, energy shortages and price instability, global environmental concerns, and human health concerns. Giffinger *et al.* [35] highlighted the performance of smart city in economy, people, governance, mobility, environment, and living. The Smarter Cities project of the Natural Resources Defense Council (see http://smartercities.nrda.org) conceptualizes smart city by highlighting positive outcomes made by being smarter.

Some definitions stress technologies. The key part of R. Hall's [39] definition is "city that monitors and integrates conditions of all of its critical infrastructures." One of core mechanisms in smart city is a self-monitoring and self-response system. IBM's view of smart city envisions its three main characteristics: instrumented, interconnected, and intelligent [40]. Instrumentation means sourcing of real-time real-world data from both physical and virtual sensors. Such data may be interconnected across multiple processes, systems, organizations, industries, or value chains. The combination of instrumented and interconnected systems effectively connects the physical world to the virtual world.

Other definitions highlight different aspects. Rios's [73] approach is based on an architectural lens. He sees smart city as a city that gives inspiration, shares culture, knowledge, and life, and motivates its inhabitants to create and flourish in their own lives. Partridge's [69] observation of Brisbane in Australia sheds light on social inclusion and equal participation as enhanced opportunities created by smart city initiatives.

Table 2. Definitions of Smart City

	Definition
[80]	"The use of Smart Computing technologies to make the critical infrastructure components and services of a city—which include city administration, education, healthcare, public safety, real estate, transportation, and utilities—more intelligent, interconnected, and efficient."
[35]	"A city well performing in a forward-looking way in economy, people, governance, mobility, environment, and living, built on the smart combination of endowments and activities of self-decisive, independent and aware citizens."
NRDC	A city striving to make itself "smarter" (more efficient, sustainable, equitable, and livable)
[39]	A city that monitors and integrates conditions of all of its critical infrastructures, including roads, bridges, tunnels, rails, subways, airports, seaports, communications, water, power, even major buildings, can better optimize its resources, plan its preventive maintenance activities, and monitor security aspects while maximizing services to its citizens.
[40]	An instrumented, interconnected, and intelligent city. *Instrumentation* enables the capture and integration of live real-world data through the use of sensors, kiosks, meters, personal devices, appliances, cameras, smart phones, implanted medical devices, the web, and other similar data-acquisition systems, including social networks as networks of human sensors. *Interconnected* means the integration of those data into an enterprise computing platform and the communication of such information among the various city services. *Intelligent* refers to the inclusion of complex analytics, modeling, optimization, and visualization in the operational business processes to make better operational decisions.
[73]	"A city that gives inspiration, shares culture, knowledge, and life, a city that motivates its inhabitants to create and flourish in their own lives"
[69]	"A city where the ICT strengthen the freedom of speech and the accessibility to public information and services"

The smart city concept has been expressed with some metaphors. Importantly, smart city has been viewed as a large organic system. Dirks and Keeling [23] stress the organic integration of systems. The interrelationship between a smart city's core systems is taken into account to make *the system of systems* smarter. No system operates in isolation. A smarter city infuses information into its physical infrastructure to improve conveniences, facilitate mobility, add efficiencies, conserve energy, improve the quality of air and water, identify problems and fix them quickly, recover rapidly from disasters, collect data to make better decisions, deploy resources effectively, and share data to enable collaboration across entities and domains. However, infusing intelligence into each subsystem of a city, one by one—transportation, energy, education, health care, buildings, physical infrastructure, food, water, public safety, etc.—is not enough to become a smarter city. A smarter city should be treated as an *organic whole*—as a *network*, as a *linked system* [49].

While systems in industrial cities were mostly skeleton and skin, postindustrial cities are like organisms that develop an *artificial nervous system*, which enables them to behave in intelligently coordinated ways [65]. The new intelligence of cities, then, resides in the increasingly effective combination of digital telecommunication networks (the nerves), ubiquitously embedded intelligence (the brains), sensors and tags (the sensory organs), and software (the knowledge and cognitive competence). There is a growing web of direct connections to the mechanical and electrical systems of buildings, household appliances, production machinery, process plants, transportation systems, electrical grids and other energy supply networks, water supply and waste removal networks, systems that provide life safety and security, and management systems for just about every imaginable human activity.

2.3 Conceptual Relatives

To build the set of common multidimensional components we need to take a close look at many conceptual cousins of smart city and trace the roots of the terms popularly used. A variety of the labels can be largely categorized into three dimensions: technology, people, and community. The conceptual variants are mutually connected with substantial confusion in definitions and complicated usages rather than independent on each other.

Table 3. Conceptual Relatives of Smart City

Dimensions	Concepts	Studies
Technology	Digital city	[6,46,47,81,82,88]
	Intelligent city	[11,52-4,61]
	Ubiquitous city	[4,5,56]
	Wired city	[24]
	Hybrid city	[77]
	Information city	[18,74,76]
People	Creative city	[31,38,55,78]
	Learning city	[16,20,21,71]
	Humane city	[77]
	Knowledge city	[23,26,27,49,84-7]
Community	Smart community	[14,15,19,20,21,28-30,43,44,49,58,66]

2.3.1 Technology Dimension

There are various cousins of the smart city concept that draws from a technology perspective. A *digital city* refers to "a connected community that combines broadband communications infrastructure; a flexible, service-oriented computing infrastructure based on open industry standards; and, innovative services to meet the needs of governments and their employees, citizens and businesses" [88]. Its goal is to create an environment for information sharing, collaboration, interoperability and seamless experiences for all inhabitants anywhere in the city. Williams [82] views it as a sharing of networks. Through digital technologies and wide-area infrastructures/applications, those networks connect organizations, social groups and enterprises located in a city area [5,6]. For example, Widmayer [81] viewed Chicago as a *digital metropolis* consisting of large networks.

The notion of an *intelligent city* emerges at the crossing of the knowledge society (a society in which knowledge and creativity have great emphasis and intangible, human and social capital are considered the most valuable asset) with the digital city [66]. Malek [61] defined an intelligent city as a city that has all the infrastructure and infostructure of information technology, the latest technology in telecommunications, electronic and mechanical technology. According to Komninos and Sefertzi [54], initiatives for intelligent cities make conscious efforts to use information technology to transform life and work within its region in significant and fundamental rather than incremental ways.

There is a conceptual and practical distinction between digital city and intelligent city. The label intelligent city is usually used to characterize a city that has the ability to support learning, technological development, and innovation procedures. In this sense, every digital city is not necessarily intelligent, but every intelligent city has digital components. Both terms are different in the linkage between real city and virtual city. Digital city involves every function of the city such as work, housing, movement, recreation, and environment. Intelligent city primarily involves functions of research, technology transfer, product development, and technological innovation, as a hotbed of innovative industries [54], analogous to knowledge city.

In a *virtual city*, city functions are implemented in a cyberspace [12]. Given the experiential blurring between cyberspace and material space [89], the category of the smart city concept comprises the notion of a *hybrid city* [77], which consists of a reality with its physical entities and real inhabitants and a parallel virtual city of counterparts of real entities and people. Today some cities are experienced as and constituted within virtual and material spaces simultaneously. However, physical distance and location still have importance for consideration [12,63]. Hyperbolic claims that distance will be dead soon belie an important paradox in cyberspace research. The vision of the world without distance still remains unmet in many ways. In practice, ubiquitous cloud of communication is underpinned and enabled by a vast, physical (placed) IT infrastructure of cables, data centers, and exchanges. Place still matters, though virtualization in many cities is accelerating.

A *ubiquitous city* (U-city) is a further extension of digital city concept in terms of ubiquitous accessibility and infrastructure [4,5]. It makes the ubiquitous computing available to the urban elements such as people, building, infrastructure and open space [56]. Its aim is to create a built environment where any citizen can get any services anywhere and anytime through any devices. The ubiquitous city is quite different from the well-known virtual city. While the virtual city reproduces urban elements by visualizing them within the virtual space, ubiquitous city is created by the computer chips or sensors inserted to those urban elements.

An *information city* refers to digital environments collecting information from local communities and delivering it to the public via web portals [5,74,76,81]. In that city, many info-habitants are able to live and work on the Internet. An information city is an urban center for commerce, social and civic services, and social interactions among people, businesses and government institutions [74,76].

2.3.2 Human Dimension

Creativity is recognized as a key driver to smart city, and thus people, education, learning and knowledge have central importance to smart city. The expansive notion of smart city includes creating a climate suitable for an emerging creative class [12]. A *creative city* is one of smart city visions. Human infrastructure (i.e., creative occupations and workforce, knowledge networks, voluntary organizations, crime-free environments, after-dark entertainment economy) is a crucial axis for city development [31].

Social infrastructure (intellectual capital and social capital) is indispensable endowment to smart cities. That infrastructure is about people and their relationship. Smart people generate and benefit from social capital. Smart city is about a mix of education/training, culture/arts, and business/commerce [7] and a hybrid mix of social enterprise, cultural enterprise, and economic enterprise.

A smart city is a *humane city* that has multiple opportunities to exploit its human potential and lead a creative life. Focusing on education, Winters [83] analyzed why smart cities are growing, who moves, and who stays. In his view, a smart city is a center of higher education and better-educated individuals. Similarly, a smart city is full of skilled workforces [37]. The knowledge worker and the high tech knowledge-sensitive industries migrate into highly livable communities [28]. The smartness of workforce diverges between cities. Smart places are getting smarter while other places getting less smarter because such places act as a magnet for creative people and workers [60]. Along with the inflow of smart people, new creative culture driven by them is a drive to urban development. Švob-Đokiæ [78] lauded the outcome of creative culture that extends beyond diversity and creativity to economic performance and social tolerance.

A smart city is also a *learning city*, which improves the competitiveness of urban contexts in the global knowledge economy [71]. Learning cities are actively involved in building a skilled information economy workforce [66]. Campbell [16] established a typology of cities that are learning to be smart: individually proactive city, city cluster, one-to-one link between cities, and city network.

A *knowledge city* is analogous to a learning city. It refers to "a city that was purposefully designed to encourage the nurturing of knowledge" [26]. Technopolis and ideapolis, early articulations of a knowledge city, have evolved into digital, intelligent or smart city [85]. The notion of knowledge city is interchangeable to a certain degree with similar evolving concepts such as intelligent city, educating city, or smart city [25,52]. However, a knowledge city is heavily related to knowledge economy, and its distinction is stress on innovation [22]. Knowledge-based urban development has become an important mechanism for the development of knowledge cities. The buzz concept of being clever, smart, skillful, creative, networked, connected, and competitive has become some of the key ingredients of knowledge-based urban development [84-7].

2.3.3 Institutional Dimension

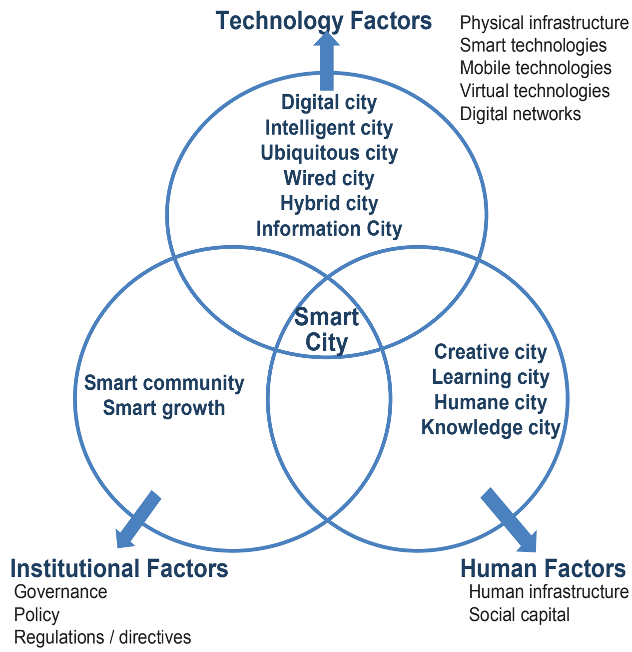
The Smart Communities movement took shape over the 1990s as a strategy to broaden the base of users involved in IT [66]. A smart community should be defined as a community broadly ranging from a small neighborhood to a nation-wide community of common or shared interest, whose members, organizations and governing institutions are working in partnership to use IT to transform their circumstances in significant ways [43]. The concept highlights governance among stakeholders and institutional factors for governance. California Institute for Smart Communities [15] elaborated the concept: "a community in which government, business, and residents understand the potential of information technology, and make a conscious decision to use that technology to transform life and work in their region in significant and positive ways." With a holistic view, a smarter community is composed of not only a more integrated, collaborative, and inclusive "whole" but also of multiple neighborhoods and communities of interest and of kind [49,58]. A smart community makes a conscious and agreed-upon decision to deploy technology as a catalyst to solving its social and business needs [28,29]. Technological propagation is not an end in itself, but only a means to reinventing cities for a new economy and society. Institutional preparation and community governance are essential to the success of smart community initiatives.

Building and planning a smart community seeks for smart growth [66]. Smart growth was the most use of the term smart in the urban context before the concept of smart city emerges [79]. The smart growth movement had prevailed during the 1990s, as a strong government- and community-driven reaction to worsening trends in traffic congestion, school overcrowding, air pollution, loss of open space, effacement of valued historic places, and skyrocketing public facilities cost [34,45,72]. Smart city resembles some functions of smart growth initiatives as an urban problem solver within or beyond the physical jurisdiction of a community. However, the smart growth concept primarily covers urban growth as the alternative or antidote to spatial sprawl [9,67]. The general implication from smart growth is that the ill-planned, ill-coordinated development provoked the smart growth movement [8]. As urban planning based on governance with multiple stakeholders is pivotal to smart growth, smart city initiatives necessitate governance for their success.

3. CORE COMPONENTS OF SMART CITY

This section discusses a set of fundamental factors which make a city smart according to the literature. From the discussion of conceptual variants of smart city in the preceding section, we identify and clarify key conceptual components of smart city, and re-categorize and simplify them into three categories of core factors: technology (infrastructures of hardware and software), people (creativity, diversity, and education), and institution (governance and policy). Given the connection between the factors, a city is smart when investments in human/social capital and IT infrastructure fuel sustainable growth and enhance a quality of life, through participatory governance [17].

Figure 1. Fundamental Components of Smart City



3.1 Technology Factors

Technology is key to being a smart city because of the use of ICT to transform life and work within a city in significant and fundamental ways [41]. A well-functioning infrastructure is absolutely necessary but not enough to become a smart city. IT infrastructure and applications are prerequisites, but without real engagement and willingness to collaborate and cooperate between public institutions, private sector, voluntary organizations, schools and citizens there is no smart city [58].

Most studies on practices of smart city address issues of technological infrastructure and enabling technologies. The focus on infrastructure and technology stresses accessibility and availability of systems [35,36]. Contrasting with human infrastructure, technological infrastructures have other names such as physical infrastructure [12] and technoware [61]. Washburn *et al.* [80] views smart city as a collection of smart computing technologies applied to critical infrastructure components and services. Smart computing refers to "a new generation of integrated hardware, software, and network technologies that provide IT systems with real-time awareness of the real world and advanced analytics to help people make more intelligent decisions about alternatives and actions that will optimize business processes and business balance sheet results" [80]. Al-Hader *et al.* [1,2] specifies technological components with the framework of smart city development pyramid: smart interface (dash board, common operational platform, integrated web services), smart control systems (automatic control network, local operating network), and smart database resources (database, database server).

Mobile, virtual, and ubiquitous technologies gain importance. Those technologies offer benefits to city dwellers in mobile lifestyle. Smart city application evolves from smart places to networked inhabitants [32]. While the wireless infrastructure is a key element of digital city infrastructure, it is only a first step [1,2]. A set of technological requisites for smart city comprises network equipments (fiber optic channels and wi-fi networks), public access points (wireless hotpots, kiosks), and service-oriented information systems [5]. A ubiquitous/pervasive computing infrastructure is a key technological component in the build out of a digital city [88]. A smart city provides interoperable, Internet-based government services that enable ubiquitous connectivity to transform key government processes, both internally across departments and employees and externally to citizens and businesses.

3.2 Human Factors

The availability and quality of the IT infrastructure is not the only definition of smart city [17]. Importantly, other definitions stress the role of human infrastructure, human capital and education in urban development [12]. For urban development, Florida [31] suggested 3T (tolerance, technology, and talent), of which two are germane to people and their relationship. Smart people is an important component of smart city [35,36]. The smart people concept comprises various factors like affinity to life long learning, social and ethnic plurality, flexibility, creativity, cosmopolitanism or open-mindedness, and participation in public life. Problems associated with urban agglomerations can be solved by means of creativity, human capital, cooperation among relevant stakeholders, and their bright scientific ideas: in a nutshell, "smart solutions" [17]. The label smart city therefore points to clever solutions by creative people.

The category of human factors highlights creativity, social learning, and education. Smart city is a center of higher education and smart workforce [37,83]. For smart city, Malek [61] emphasizes the importance of humanware, which represents cognitive/creative capability and human skills. Smart city bolsters a creative environment [86]. The category of human factors also includes social inclusion of various urban residents in public services, soft infrastructure (knowledge networks, voluntary organizations, crime-free environments), urban diversity and cultural mix, social/human/relational capital, and knowledge base such as educational institutions and R&D capacities [41,85].

Education is a critical magnet that makes a city attractive. Businesses, organizations, and individuals of all backgrounds gravitate to dynamic learning environments [10]. IT education enabled the vision of Singapore as an intelligent island [59]. Collective intelligence and social learning make a city smarter [20]. The notion of smart community refers to the locus in which networked intelligence is embedded and continuous learning is nurtured. To explain functioning mechanisms of smart community, the hidden portion of the iceberg is collective intelligence and social learning [20]. A smart city initiative becomes an integrated approach to connecting among entire communities (governments, businesses, schools, non-profits, and individual citizens), creating specific services to address city objectives, and advancing collective skills and capacities.

3.3 Institutional Factors

The support of government and policy for governance is fundamental to the design and implementation of smart city initiatives. This category comprises a variety of institutional factors drawing from the discussion of smart community or smart growth initiatives: not just supportive policies but also the role of government, the relationship between government agencies and non-government parties, and their governance. It is necessary to establish *administrative environment* (initiatives, structure, and engagement) supportive for smart city [86]. To enable smart city initiatives, the category should also include integrated and transparent governance, strategic and promotional activities, networking, and partnerships [68].

IBM [42] presented smart government as one key component for smart city. Smarter government will do more than simply regulate the outputs of economic and societal systems. It interconnects dynamically with citizens, communities, and businesses in real time to spark growth, innovation, and progress. The challenges vary from departmental silos to process delays to the lack of transparency and accountability. Smarter government means collaborating across departments and with communities—to become more transparent and accountable, to manage resources more effectively, and to give citizens access to information about decisions that affect their lives. Leading governments are integrating their service delivery, establishing offices that support multiple services, and placing the most needed transactions on the Web. At the most fundamental level, smarter government means making operations and services truly *citizen-centric*.

The transformation to smart city entails interactions of technological components with political, institutional and transitional components [64]. Political components represent endogenous political elements (directions, city council, city government, city mayor), harmonized by exogenous ones (international pressures, agenda, projects, strategies in prevalence) and verified by best practices. Institutional components are prerequisites as well. Institutional readiness such as removing legal and regulatory barriers is important. Transitional components comprise visions, leadership, and organizational transition in structure.

As a cornerstone of smart city, smart governance means various stakeholders' (especially citizens') engagement in decision-making and public/social services [36,37]. IT-mediated governance, so called e-governance, is key to enabling smart city by bringing citizens to a smart city initiative and keeping the decision and implementation process transparent [70]. The central spirit of governance is a citizen-centric, citizen-driven approach. The consideration of stakeholders (i.e., end-users, groups of end-users, IT experts, policy/service domain experts, and public managers) is fundamental to architecture of smart city [5,57]. Successful initiatives are the result by a coalition of business, education, government and individual citizens [58]. A successful smart city can be built from top down or bottom up approaches, but active involvement from every sector of the community is essential. United efforts create synergy, which allows individual projects to build upon each other for faster progress, resulting in the involved, informed and trained critical mass necessary for transformation of how the entire community carries out its work.

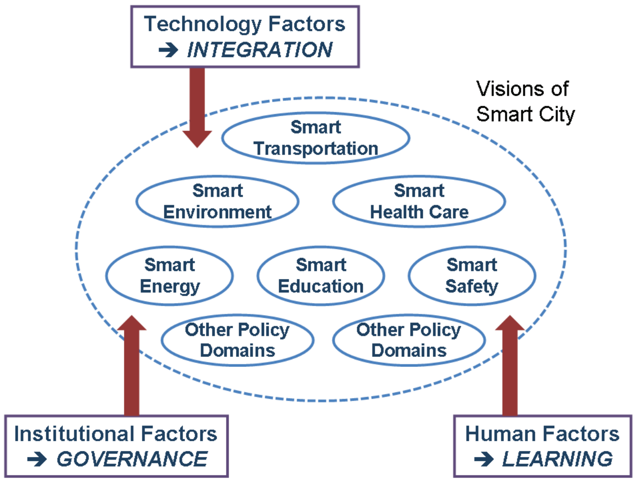
4. STRATEGIC DIRECTIONS IN KEY DIMENSIONS

This section offers strategic principles for making a city smart in order to realize the various visions specified for diverse policy domains, aligning to the three categories of core components identified in the preceding section.

4.1 Integration of Technology Factors

A solution to make a city smarter introduces a new level of complexity [48]. The solution should extend beyond technology, but we should still value the indispensable role of technology. Smart city integrates technologies, systems, infrastructures, services, and capabilities into an organic network that is sufficiently complex for unexpected emergent properties to develop. Integrative service of smart city faces challenges as well as opportunities. The perception of technology in smart city initiatives stresses integration of systems, infrastructures and services mediated through enabling technologies. Technological innovation is a means to smart city, not an ends. IT is just a facilitator for creating a new type of innovative environment, which requires the comprehensive and balanced development of creative skills, innovation-oriented institutions, broadband networks, and virtual collaborative spaces [53].

Figure 2. Strategic Directions of Smart City



4.2 Learning for Human Factors

The emphasis on human infrastructure highlights social learning and education. Towards more progressive smart cities, cities should start with people from the human capital side, rather than blindly believing that IT itself can automatically transform and improve cities [41]. To a substantial extent that is already recognized, the critical factor in any successful city is its people and how they interact. Stronger approaches to awareness, education and leadership offer services that are accessible to all of citizens, get rid of barriers related to language, culture, education, skills development, and disabilities [20]. Social learning soothes the digital divide concern for those who lag behind the prevalent use of the new technologies. Education and training actions should develop IT skills, nurture knowledge workers, facilitate the environment of social learning, and improve IT training in schools, organizations and industries [13].

4.3 Governance of Institutional Factors

Governance encapsulates collaboration, cooperation, partnership, citizen engagement, and participation [20]. Successful cities possess a set of common features [29]. One characteristic is collaboration among different functional sectors and parties (government, business, academics, non-profit and voluntary organizations, and others), and among different jurisdictions within a given geographical region [3,45,58,70]. City government should share concepts (promotional identity and brand), visions, goals, priorities, and even strategic plans of smart city with the public and stakeholders [22,29,68]. Leadership of key leaders and their strong support (championing) of the smart city vision are fundamental to the success of smart city [5,10,14,15,33]. The role of leadership is pivotal both within government and for its relation with citizens.

5. CONCLUDING REMARKS

We expect that the elaborated conceptualization of smart city in this paper will contribute to future studies. As we explored multiple conceptual dimensions of smart city, the concept is an organic connection among technological, human, and institutional components. Nowadays the usage of "smart" captures innovative and transformative changes driven by new technologies. However, social factors other than smart technologies are central to smart cities. In this sense, a socio-technical view on smart city is needed. Leading a smart city initiative requires a comprehensive understanding of the complexities and interconnections among social and technical factors of services and physical environments in a city. For future research based on a socio-technical view, we must explore both "how do smart technologies change a city?" and "how do traditional institutional and human factors in urban dynamics impact a smart city initiative leveraged by new technologies?". This research will also explore the practical implications of the conceptual model suggested. To that end, we will continue studying smart city by focusing on exemplar practices of smart city initiatives, considering the dynamics of various stakeholders in those initiatives, and discussing policy innovation in city governments.

6. REFERENCES

[1] Al-Hader, M., Rodzi, A., Sharif, A. R., & Ahmad, N. (2009a). Smart city components architecture. In *Proceedings of the International Conference on Computational Intelligence, Modelling and Simulation*, (Brno, Czech Republic, Sep 7-9)

[2] Al-Hader, M., Rodzi, A., Sharif, A. R., & Ahmad, N. (2009b). SOA of smart city geospatial management. In *Proceedings of the 3rd UKSim European Symposium on Computer Modeling and Simulation* (Athens, Greece, Nov 25-27). Available at http://doi.ieeecomputersociety.org/10.1109/EMS.2009.112.

[3] Anderson, G., & Tregoning, H. (1998). Smart growth in our future? In Urban Land Institute (Ed.), *ULI on the Future: Smart Growth* (pp. 4-11). Washington, DC: Urban Land Institute.

[4] Anthopoulos, L., & Fitsilis, P. (2010a). From digital to ubiquitous cities: Defining a common architecture for urban development. In *Proceedings of the 6th International Conference on Intelligent Environments* (Kuala Lumpur, Malaysia, Jul 19-21).

[5] Anthopoulos, L., & Fitsilis, P. (2010b). From online to ubiquitous cities: The technical transformation of virtual communities. In A. B. Sideridis & C. Z. Patrikakis (Eds.), *Next Generation Society: Technological and Legal Issues* (*Proceedings of the Third International Conference, e-Democracy 2009*, Athens, Greece, Sep 23-25, 2009) (Vol. 26, pp. 360-372). Berlin, Germany: Springer. Available at http://www.springerlink.com/content/g644776482968k36/fulltext.pdf.

[6] Anthopoulos, L., & Tsoukalas, I. A. (2005). The implementation model of a digital city. *Journal of E-Government*, *2*(2), 91-110.

[7] Bartlett, L. (2005). Smart city: Social entrepreneurship and community engagement in a rural regional city. In *Proceedings of the International Conference on Engaging Communities*, (Brisbane, Australia, Aug 14-17). Available at http://www.engagingcommunities2005.org/abstracts/Bartlet-Leo-final.pdf.

[8] Beatley, T., & Collins, R. (2000). Smart growth and beyond: Transitioning to a sustainable society. *Virginia Environmental Law Journal*, *19*, 287-322.

[9] Benfield, F. K., Terris, J., & Vorsanger, N. (2001). *Solving Sprawl: Models of Smart Growth in Communities across America*. New York: National Resources Defense Council.

[10] Boise Smart City Initiative. (2002). *Boise Smart City Initiative Committee Report*. Boise, ID: Capital City Development Corp. Available at http://www.ccdcboise.com/Documents/SC_committee_report_for_screen.pdf.

[11] Borja, J. (2007). Counterpoint: Intelligent cities and innovative cities. *Universitat Oberta de Catalunya (UOC) Papers: E-Journal on the Knowledge Society*, *5*. Available at http://www.uoc.edu/uocpapers/5/dt/eng/mitchell.pdf.

[12] Boulton, A., Brunn, S. D., & Devriendt, L. (Forthcoming). Cyberinfrastructures and "smart" world cities: Physical, human, and soft infrastructures. In P. Taylor, B. Derudder, M. Hoyler & F. Witlox (Eds.), *International Handbook of Globalization and World Cities*. Cheltenham, U.K.: Edward Elgar. Available at http://www.neogeographies.com/documents/cyberinfrastructure_smart_world_cities.pdf.

[13] Cairney, T., & Speak, G. (2000). *Developing a 'Smart City': Understanding Information Technology Capacity and Establishing an Agenda for Change*. Sydney, Australia: Centre for Regional Research and Innovation, University of Western Sydney. Available at http://trevorcairney.com/file_uploads/cgi-lib.30886.1.IT_Audit.pdf.

[14] California Institute for Smart Communities. (1997). *Smart Communities Implementation Guide: How California's Communities Can Thrive in the Digital Age*. Available at http://www.smartcommunities.org/guide/index.html.

[15] California Institute for Smart Communities. (2001). *Smart Communities Guide Book*. Available at http://www.smartcommunities.org/guidebook.html.

[16] Campbell, T. (2009). Learning cities: Knowledge, capacity and competitiveness. *Habitat International*, *33*(2), 195-201.

[17] Caragliu, A., Del Bo, C., & Nijkamp, P. (2009). Smart cities in Europe. In *Proceedings of the 3rd Central European Conference in Regional Science* (Košice, Slovak Republic, Oct 7-9). Available at http://www.cers.tuke.sk/cers2009/PDF/01_03_Nijkamp.pdf.

[18] Castells, M. (1996). *Rise of the Network Society: The Information Age*. Cambridge, MA: Blackwell.

[19] Center on Governance. (2003). *SmartCapital Evaluation Guidelines Report: Performance Measurement and Assessment of SmartCapital*. Ottawa, Canada: University of Ottawa. Available at http://www.christopherwilson.ca/papers/Guidelines_report_Feb2003.pdf.

[20] Coe, A., Paquet, G., & Roy, J. (2001). E-governance and smart communities: A social learning challenge. *Social Science Computer Review*, *19*(1), 80-93.

[21] Collins, B., Paquet, G., Roy, J., & Wilson, C. (2002). E-governance and smart communities: A social learning challenge. In *Proceedings of the SSHRC Knowledge Based Economy Workshop*, (Newfoundland, Canada, May 10-11). Available at http://www.christopherwilson.ca/papers/Nfld_paper_2002.pdf.

[22] Dirks, S., Gurdgiev, C., & Keeling, M. (2010). *Smarter Cities for Smarter Growth: How Cities Can Optimize Their Systems for the Talent-Based Economy*. Somers, NY: IBM Global Business Services. Available at ftp://public.dhe.ibm.com/common/ssi/ecm/en/gbe03348usen/GBE03348USEN.PDF.

[23] Dirks, S., & Keeling, M. (2009). *A Vision of Smarter Cities: How Cities Can Lead the Way into a Prosperous and Sustainable Future*. Somers, NY: IBM Global Business Services. Available at ftp://public.dhe.ibm.com/common/ssi/ecm/en/gbe03227usen/GBE03227USEN.PDF.

[24] Dutton, W. H. (1987). *Wired Cities: Shaping the Future of Communications*. London: Macmillan.

[25] Dvir, R., & Pasher, E. (2004). Innovation engines for knowledge cities: An innovation ecology perspective. *Journal of Knowledge Management*, *8*(5), 16-27. Available at http://www.emeraldinsight.com/journals.htm?issn=1367-3270&volume=8&issue=5&articleid=1506527&show=pdf&PHPSESSID=nhgj3cfuo0c783rinvovb73am1.

[26] Edvinsson, L. (2006). Aspects on the city as a knowledge tool. *Journal of Knowledge Management*, *10*(5), 6-13. Available at http://www.corporatelongitude.com/download/Aspects_on_city.pdf.

[27] Edvinsson, L., Dvir, R., Roth, N., & Pasher, E. (2004). Innovations: The new unit of analysis in the knowledge era. *Journal of Intellectual Capital*, *5*(1), 40-58.

[28] Eger, J. M. (2000, Feb 13). Cities: Smart growth and the urban future. The San Diego Union Tribune

[29] Eger, J. M. (2009). Smart growth, smart cities, and the crisis at the pump a worldwide phenomenon. *I-Ways*, *32*(1), 47-53.

[30] Eger, J. M., & Maggipinto, A. (2010). Technology as a tool of transformation: e-Cities and the rule of law. In A. D'Atri & D. Saccà (Eds*.), Information Systems: People, Organizations, Institutions, and Technologies* (pp. 23-30). Berlin/Heidelberg, Germany: Physica-Verlag.

[31] Florida, R. (2002). *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday life*. New York: Basic Books. Available at http://www.washingtonmonthly.com/features/2001/0205.florida.html.

[32] Forest, F., Lavoisy, O., Eurich, M., van Gurp, J., & Wilson, D. (2009). Roadmap for real world Internet applications: Socioeconomic scenarios and design recommendations. In G. Tselentis, J. Domingue, A. Galis, A. Gavras, D. Hausheer, S. Krco, V. Lotz & T. Zahariadis (Eds.), *Towards the Future*

Internet: A European Research Perspective (pp. 325-334). Amsterdam, The Netherlands: IOS Press.

[33] Frece, J. W. (2008). *Sprawl & Politics: The Inside Story of Smart Growth in Maryland*. Albany, NY: State University of New York Press.

[34] Freilich, R. H. (1999). *From Sprawl to Smart Growth*. Chicago: American Bar Association.

[35] Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanović, N., & Meijers, E. (2007). *Smart Cities: Ranking of European Medium-Sized Cities*. Vienna, Austria: Centre of Regional Science (SRF), Vienna University of Technology. Available at http://www.smart-cities.eu/download/smart_cities_final_report.pdf.

[36] Giffinger, R., & Gudrun, H. (2010). Smart cities ranking: An effective instrument for the positioning of cities? *ACE: Architecture, City and Environment*, *4*(12), 7-25. Available at http://upcommons.upc.edu/revistes/bitstream/2099/8550/7/ACE_12_SA_10.pdf.

[37] Glaeser, E. L., & Berry, C. R. (2006). Why are smart places getting smarter? *Taubman Center Policy Briefs*, PB-2006-2. Available at http://www.hks.harvard.edu/rappaport/downloads/policybriefs/brief_divergence.pdf.

[38] Hall, P. (2000). Creative cities and economic development. *Urban Studies*, *37*(4), 633-649.

[39] Hall, R. E. (2000). The vision of a smart city. In *Proceedings of the 2nd International Life Extension Technology Workshop* (Paris, France, Sep 28). Available at http://www.osti.gov/bridge/servlets/purl/773961-oyxp82/webviewable/773961.pdf.

[40] Harrison, C., Eckman, B., Hamilton, R., Hartswick, P., Kalagnanam, J., Paraszczak, J., & Williams, P. (2010). Foundations for Smarter Cities. *IBM Journal of Research and Development*, *54*(4). DOI: 10.1147/JRD.2010.2048257.

[41] Hollands, R. G. (2008). Will the real smart city please stand up? *City*, *12*(3), 303-320.

[42] IBM. (2010). *Smarter Thinking for a Smarter Planet*. Available at http://www.ibm.com/smarterplanet/global/files/us__en_us__loud__ibmlbn0041_transtasman_book.pdf.

[43] Industry Canada. (1998). *Report of the Panel on Smart Communities*. Ottawa, Canada: Government of Canada.

[44] Industry Canada. (1999). *Smart Communities: Program Guide*. Ottawa, Canada: Government of Canada.

[45] Ingram, G. K., Carbonell, A., Hong, Y.-H., & Flint, A. (Eds.). (2009). *Smart Growth Policies: An Evaluation of Programs and Outcomes*. Cambridge, MA: Lincoln Institute of Land Policy.

[46] Ishida, T. (2002). Digital city Kyoto. *Communications of the ACM*, *45*(7), 78-81.

[47] Ishida, T., & Isbister, K. (Eds.). (2000). *Digital Cities: Technologies, Experiences, and Future Perspectives* (Vol. 1765). Berlin, Germany: Springer.

[48] Jennings, P. (2010). Managing the risks of Smarter Planet solutions. *IBM Journal of Research and Development*, *54*(4). DOI: 10.1147/JRD.2010.2050540.

[49] Kanter, R. M., & Litow, S. S. (2009). Informed and interconnected: A manifesto for smarter cities. *Harvard Business School General Management Unit Working Paper*, 09-141. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1420236.

[50] Klein, C., & Kaefer, G. (2008). From smart homes to smart cities: Opportunities and challenges from an industrial perspective. In *Proceedings of the 8th International Conference, NEW2AN and 1st Russian Conference on Smart Spaces, ruSMART 2008* (St. Petersburg, Russia, Sep 3-5). Available at http://www.springerlink.com/content/d053p7u7g42u573p/.

[51] Knight, J., & Weedon, A. (1995). Editorial. *Convergence: The Journal of Research into New Media Technologies*, *1*(1), 5-8.

[52] Komninos, N. (2002). *Intelligent Cities*. London: Spon Press.

[53] Komninos, N. (2009). Intelligent cities: Towards interactive and global innovation environments. *International Journal of Innovation and Regional Development*, *1*(4), 337-355.

[54] Komninos, N., & Sefertzi, E. (2009). Intelligent cities: R&D offshoring, Web 2.0 product development and globalization of innovation systems. Paper presented at the Second Knowledge Cities Summit 2009. Available at http://www.urenio.org/wp-content/uploads/2008/11/Intelligent-Cities-Shenzhen-2009-Komninos-Sefertzi.pdf.

[55] Landry, C. (2000). *The Creative City: A Toolkit for Urban Innovation*. London: Earthscan.

[56] Lee, S., Han, J., Leem, Y., & Yigitcanlar, T. (2008). Towards ubiquitous city: Concept, planning, and experiences in the Republic of Korea. In T. Yigitcanlar, K. Velibeyoglu & S. Baum (Eds.), *Knowledge-Based Urban Development : Planning and Applications in the Information Era* (pp. 148-169). Hershey, PA: IGI Global.

[57] Lepouras, G., Vassilakis, C., Halatsis, C., & Georgiadis, P. (2007). Domain expert user development: The SmartGov approach. *Communications of the ACM*, *50*(9), 79-83.

[58] Lindskog, H. (2004). Smart communities initiatives. In *Proceedings of the 3rd ISOneWorld Conference* (Las Vegas, NV, Apr 14-16). Available at http://www.heldag.com/articles/Smart%20communities%20april%202004.pdf.

[59] Mahizhnan, A. (1999). Smart cities: The Singapore case. *Cities*, *16*(1), 13-18.

[60] Malanga, S. (2004). The curse of the creative class. *City*, *14*(1). Available at http://www.city-journal.org/html/14_1_the_curse.html.

[61] Malek, J. A. (2009). Informative global community development index of informative smart city. In *Proceedings of the 8th WSEAS International Conference on Education and Educational Technology* (Genova, Italy, Oct 17-19).

[62] Marsa-Maestre, I., Lopez-Carmona, M. A., Velasco, J. R., & Navarro, A. (2008). Mobile agents for service personalization in smart environments. *Journal of Networks*, *3*(5), 30-41.

[63] Martin, J. L., Varilly, H., Cohn, J., & Wightwick, G. R. (2010). Preface: Technologies for a smarter planet. *IBM Journal of Research and Development*, *54*(4). DOI: 10.1147/JRD.2010.2051498.

[64] Mauher, M., & Smokvina, V. (2006). Digital to intelligent local government transition framework. In *Proceedings of the 29th International Convention of MIPRO* (Opatija, Croatia, May 22-26). Available at http://www.mmc-consulting.hr/Download/2008/03/07/Mauher_M_Digital_to_Intelligent_City_Transition_Framework.pdf.

[65] Mitchell, W. J. (2006). Smart City 2020. *Metropolis*, April. Available at http://www.metropolismag.com/story/20060320/smart-city-2020.

[66] Moser, M. A. (2001). What is smart about the smart communities movement? *EJournal*, *10/11*(1). Available at http://www.ucalgary.ca/ejournal/archive/v10-11/v10-11n1Moser-print.html.

[67] O'Toole, R. (2001). *The Vanishing Automobile and Other Urban Myths: How Smart Growth Will Harm American Cities*. Bandon, OR: The Thoreau Institute.

[68] Odendaal, N. (2003). Information and communication technology and local governance: Understanding the difference between cities in developed and emerging economies. *Computers, Environment and Urban Systems*, *27*(6), 585-607.

[69] Partridge, H. (2004). Developing a human perspective to the digital divide in the smart city. In *Proceedings of the Biennial Conference of Australian Library and information Association* (Queensland, Australia, Sep 21-24). Available at http://eprints.qut.edu.au/1299/1/partridge.h.2.paper.pdf.

[70] Paskaleva, K. A. (2009). Enabling the smart city: The progress of city e-governance in Europe. *International Journal of Innovation and Regional Development*, *1*(4), 405-422.

[71] Plumb, D., Leverman, A., & McGray, R. (2007). The learning city in a 'planet of slums'. *Studies in Continuing Education*, *29*(1), 37-50.

[72] Porter, D. R. (2002). *Making Smart Growth Work*. Washington, DC: Urban Land Institute.

[73] Rios, P. (2008). Creating "the smart city". Available at http://dspace.udmercy.edu:8080/dspace/bitstream/10429/20/1/2008_rios_smart.pdf.

[74] Sairamesh, J., Lee, A., & Anania, L. (2004). Information cities. *Communications of the ACM*, *47*(2), 28-31.

[75] Spangler, W. S., Kreulen, J. T., Chen, Y., Proctor, L., Alba, A., Lelescu, A., & Behal, A. (2010). A smarter process for sensing the information space. *IBM Journal of Research and Development*, *54*(4). DOI: 10.1147/JRD.2010.2050541.

[76] Sproull, L., & Patterson, J. F. (2004). Making information cities livable. *Communications of the ACM*, *47*(2), 33-37.

[77] Streitz, N. (2009). Ambient intelligence landscapes for realizing the cities of the future: Introduction and overview. In *Proceedings of the 3rd European Conference on Ambient Intelligence* (Salzburg, Austria, Nov 18-21). Available at http://www.smart-future.net/14.html.

[78] Švob-Đokiæ, N. (Ed.). (2007). *The Creative City: Crossing Visions and New Realities in the Region*. Zagreb, Croatia: Institute for International Relations, Available at http://www.culturelink.org/publics/joint/cultid08/Svob-Djokic_Creative_City.pdf.

[79] Urban Land Institute. (1998). *ULI on the Future: Smart Growth*. Washington, DC: Urban Land Institute.

[80] Washburn, D., Sindhu, U., Balaouras, S., Dines, R. A., Hayes, N. M., & Nelson, L. E. (2010). *Helping CIOs Understand "Smart City" Initiatives: Defining the Smart City, Its Drivers, and the Role of the CIO*. Cambridge, MA: Forrester Research, Inc. Available at http://public.dhe.ibm.com/partnerworld/pub/smb/smarterplanet/forr_help_cios_und_smart_city_initiatives.pdf.

[81] Widmayer, P. (1999). Building digital metropolis: Chicago's future networks. *IT Professional*, *1*(4), 40-46.

[82] Williams, M. (2010, Nov 18). 2010 digital cities survey winners announced. *Government Technology*. Available at http://www.govtech.com/e-government/2010-Digital-Cities-Survey-Winners-Announced.html.

[83] Winters, J. V. (2010). Why are smart cities growing? Who moves and who stays. *Journal of Regional Science*, *20*(10), 1-18.

[84] Yigitcanlar, T., & McCartney, R. (2010). Strategising knowledge-based urban development: Knowledge city transformations of Brisbane, Australia. In *Proceedings of the 14th International Planning History Society* (*IPHS*) *Conference*, (Istanbul, Turkey, Jul 12-15).

[85] Yigitcanlar, T., O'Connor, K., & Westerman, C. (2008a). The making of knowledge cities: Melbourne's knowledge-based urban development experience. *Cities*, *25*(2), 63-72.

[86] Yigitcanlar, T., & Velibeyoglu, K. (2008). Knowledge-based urban development: The local economic development path of Brisbane, Australia. *Local Economy*, *23*(3), 195-207.

[87] Yigitcanlar, T., Velibeyoglu, K., & Martinez-Fernandez, C. (2008b). Rising knowledge cities: The role of urban knowledge precincts. *Journal of Knowledge Management*, *12*(5), 8-20.

[88] Yovanof, G. S., & Hazapis, G. N. (2009). An architectural framework and enabling wireless technologies for digital cities & intelligent urban environments. *Wireless Personal Communications*, *49*(3), 445-463. Available at http://www.springerlink.com/content/g1v63025217mt8x0/.

[89] Zook, M. A., & Graham, M. (2007). Mapping digiplace: Geocoded Internet data and the representation of place. *Environment and Planning B: Planning and Design*, *34*(3), 466-482. Available at http://www.envplan.com/epb/fulltext/b34/b3311.pdf.

Progress in IS

Renata Paola Dameri
Camille Rosenthal-Sabroux *Editors*

Smart City

How to Create Public and Economic Value with High Technology in Urban Space

Springer

Progress in IS

For further volumes:
http://www.springer.com/series/10440

Renata Paola Dameri · Camille Rosenthal-Sabroux
Editors

Smart City

How to Create Public and Economic Value with High Technology in Urban Space

Springer

Editors
Renata Paola Dameri
Department of Economics
University of Genova
Genova
Italy

Camille Rosenthal-Sabroux
Université Paris IX Dauphine
Paris
France

ISSN 2196-8705 ISSN 2196-8713 (electronic)
ISBN 978-3-319-06159-7 ISBN 978-3-319-06160-3 (eBook)
DOI 10.1007/978-3-319-06160-3
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014942273

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Today, the smart city is a red-hot topic on the urban strategy agendas of governments worldwide. This is especially so in the advanced countries, where fast-paced urban growth has thrown open the door to a mounting number of complex infrastructural and social issues.

Smart cities are being piloted in Europe, the Americas and Asia, from London to Boston to Hong Kong, from Barcelona to Amsterdam to Sau Paulo do Brazil, as citizens across the globe demand their local governments provide urban spaces designed to improve their quality of life. Yet another challenge to the citizens' quality of life is the environmental impact of these ever-larger, more technologically endowed cities, which can only be addressed by reducing pollution levels and through the wise management of natural resources; in other words, by investing in sustainable economic development.

The smart city issue is complex because it straddles several domains, from the city's physical capital to its intellectual and social capital. City planning is not just a question of urban design, but also brings into play social studies, political science, and economics. Further, the concept of smart city is underpinned by its technological core, which in turn is driven by the advances made in the fields of computer science and engineering.

The sharp increase in the number of scientific papers and empirical reports on smart cities forms a loud chorus that underscores the great interest and appeal of this new topic. The book surveys hundreds of scientific contributions on smart cities and affinity concepts, such as digital cities, intelligent cities, and green cities, published since 2010. In addition, the smart city trend has led hundreds of aspiring smart city players to upload their smart city plan to the Internet, making them accessible to all.

The advent of the smart city has sparked great fizz and bang all round, raising public interest to considerable heights, but also sowing confusion. Indeed, an analytical review of the literature reveals several theoretical roadblocks that need to be leaped before we can chart a roadmap that is as smart as the smart city we aspire to live in. Definition, governance, planning, and evaluation are the key steps that need to be addressed on the theoretical and design path that will lead to the best practices, which makes *Smart City—Using High Technology in Urban Spaces to Create Public and Economic Value* edited by Renata Paola Dameri and Camille

Rosenthal-Sabroux a welcome initiative, one that will consolidate our extant knowledge on the complex and multifaceted nature of the smart city.

The book sets out to collate the most important studies written on Europe's smart cities in an attempt to understand whether a smart city truly has the potential to create public value for citizens.

To date, the assumption of all the reviewed smart city studies and implementer reports is that the smart city is a good thing but, strangely, these provide no empirical evidence to support the claims that it helps to improve the quality of life of its citizens. These studies and reports assume that a city is smart exclusively thanks to the technology that is its core component, pointing to it as a winning card, but neglect to study the outcome and impact of the technology on the everyday life of the smart city's people, i.e., the relationship forged by the user with the technology.

As a result, this book dedicates several chapters to the debate on how to measure the impact of smart city initiatives on the creation of public value for the people who live, work, study, and visit a city. To date, studies that explore how to define and measure smart city performance are few and far between, mostly because not only is it difficult to measure a phenomenon that is still embryonic and, hence fuzzy, but also because of the subjective and nuanced view that each citizen has of the quality of life.

Nevertheless, no matter how high the hurdle, it must be leaped if we want solve the crux of how to measure smart city performance and, hence, chart an effective and practical roadmap to achieve the goal of a comprehensive smart city.

The smart cities that exist at present are mainly pilot projects that rely on the use of ICT to transform the traditional city into a better, more liveable place. However, to implement the smart city concept on a global scale takes significant resources, investments, time, and effort, not to mention political commitment. Therefore, if we really want to design and implement projects that create value and generate high returns on investment we need to develop a smart city framework that enables us to gain intelligence and traction on all the gaps in our current knowledge.

We are facing what is called a "grand challenge," meaning that the issue will keep us engrossed for several years to come and, while we are unlikely to arrive at the perfect solution, we still need to explore, investigate, analyze, question, debate, and discuss the smart city to arrive at part-solutions that can put a better and brighter spin on the way we live in our cities.

Rome Marco De Marco

Contents

Smart City and Value Creation

Renata Paola Dameri and Camille Rosenthal-Sabroux

Abstract During the latest five years, the label smart city has been spreading all over the world, impacting on urban strategies in both large and small towns. To face the increasing problems of urban areas, local public government, companies, not-for-profit organizations and the citizens themselves embraced the idea of a smarter city, using more technologies, creating better life conditions and safeguarding the environment. However, today the smart city panorama appears very confused. No acknowledged smart city definition exists till now and several cities defining themselves smart completely lack of a strategic vision about their smart future. This first chapter is the introduction of this book collecting several contributes from different academic studies all over Europe. The aim of this work is to offer a large vision about the smart city phenomenon and to compare researches and considerations regarding how to define a smart city, how to design a smart strategy and how to measure if smart actions really are able to create public value for citizens and a better quality of life in urban spaces. This chapter introduces the most important themes regarding the smart city and further deepened in the ten chapters of the book.

Keywords Smart city • Smart strategy • Smartness • Performance measurement • Public value

1 Searching for a Shared Smart City Idea

During the latest five years, the label smart city has been spreading all over the world, impacting on urban strategies in both large and small towns [1]. To face the increasing problems of urban areas, local public government, companies,

R. P. Dameri (✉)
Department of Economics, University of Genoa, Genoa, Italy
e-mail: dameri@economia.unige.it

C. Rosenthal-Sabroux
Université Paris IX Dauphine, Paris, France
e-mail: camille.rosenthal-sabroux@dauphine.fr

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_1,

not-for-profit organizations and the citizens themselves embraced the idea of a smarter city, using more technologies, creating better life conditions and safeguarding the environment.

However, the smart city idea has more ancient roots [2]. A large literature survey about smart city and digital city scientific papers, realized by Annalisa Cocchia and extensively debated in the next chapter of this book, observes that these themes have been studied from twenty years ago [3]. Therefore the idea of a city able to be smart and digital, that is, to use technology and especially ICT to improve the quality of life in urban space, is quite old [4]. But only during the latest years the attention about this topic has a peek. There are several reasons about this evidence: the larger diffusion of mobile devices and the Internet among citizens, the higher and higher dimensions of cities, the need to safeguard the environment from pollution and energy consumption [5].

Today smart city is in the mood, not only in academic or scientific researches, but especially in public government choices and projects. Looking for smart city web sites, the results are millions. It seems that every city all over the world, across continents and independently from dimension, culture, economic situation, considers important to be smart [6].

For these reasons, the panorama is very confused. A deeper analysis of the literature survey, presented by Cocchia and also by Dameri in their work "Smart and Digital City: A Systematic Literature Review" [3], considering not only the number of papers or their geographical distribution but also their content, shows that a shared and sound definition of smart city still lacks. Even if there are some most cited definitions, their meaning is quite different each other. Moreover, owing to the continuous and fast innovation regarding the smart city enabling technologies, it is difficult to compare definitions written in a time elapse of three/four years [7].

Also the smart city empirical implementation shows the same heterogeneity. Cities have been starting to implement their own smart projects. Both citizens, companies and public governments have very high expectations from the positive impact of smart actions on the quality of life or on the appeal of their city. Sometimes a smart city project is seen like a panacea able to solve all the urban problems, such as pollution, local public transport difficulties, inequalities between people, economic crisis, and so on. But these expectations are often not supported nor by a clear smart vision of the city nor by effective smart programs and initiatives [8].

The smart city implementation generally rises like a bottom up phenomenon, that is, several actors independently each others start to realize a smart initiative, using some public infrastructures or technological solutions. For example, a public hospital realizes an on-line health record access, a company supplies electric cars to its employers and the municipality replaces old buses with new ones, with a lower impact on air pollution. Three smart actions, using technology to improve the quality of life in urban spaces and to reduce pollution and energy consumption, but not included into a comprehensive vision able to define goals, expected results and scheduled time for project realization. Moreover, the lack of a framework to collect all these initiatives prevents to realize important synergies and also to communicate to the citizens the improved smartness of their city [9].

One of the primary defects of this smart city first wave is the excessive stress on the pivotal role of the technology. Indeed, technology is certainly the core aspect of a smart city, but it is not enough to create public value for citizens. The human contribute is necessary, to really embody the smart actions into the daily life of people living, studying, working in the city or also visiting the city for one or a few days for work or tourism. It should be therefore necessary to speak about smart people in smart city and to consider people, technology and strategic vision like indispensable components of a successful smart program [10, 11].

Till now, the lack of a smart strategic vision negatively impacts on the performance obtained by smart projects and initiatives. But, moreover, no city till now has developed and applied a set of key performance indicators and a measurement framework to evaluate the real effectiveness of smart actions. Perhaps it is not severe when smart city is a pioneering project, but it becomes a real obstacle in obtaining success when the smart city project wants to deliver sustainable returns to large public and private investments [12].

The mosaic emerging from the smart city panorama is colorful and rich of suggestions to support both further studies and better implementation plans. It clearly emerges that smart city is a complex challenge, because it involves several dimensions: technology, citizens, public and private bodies, urban vision [13]. Moreover it interests cities all over the world, with very deep differences each other: cultural, economic, social. Each city wants both to apply a shared smart city idea and to pursue its own specific goals.

This complexity requires the development of a governance framework of smart cities, built upon a shared smart city definition, but flexible to be adapted to different and specific needs; it should include all the steps of the governance activity, that is: to define a strategic vision, to design long term strategies, to prioritize and schedule projects and to measure the obtained results for different stakeholders.

In this book, several points of view are collected, to put the theoretical basis for a comprehensive, agile and flexible Smart City Governance Framework.

2 Smart City Definitions and Strategic Vision

We said that a generally accepted definition of smart city still lacks. Why is it so difficult to define a smart city? There are several reasons.

As Cocchia and Dameri show in their chapters in this book, the emerging of smart themes is originally strictly joined with the digital city idea. Indeed, examining the most cited definitions of smart city and digital city listed by Cocchia in the next chapter, several elements are the same in both the topics. But an important reason to explain the difficult to define the smart city should be found in its bottom-up nature.

Rising from the empirical application, the concrete smart city is especially a collection of several projects, initiatives and actions, carried out both by public and by private organizations. Therefore, as these initiatives are the result of spontaneous choices by different actors, depending on their own interests but also on the specificity of a city, the collections are very heterogeneous. To design a definition observing one or several case studies means to write a definition describing a specific smart city, and not a standard [14].

Giffinger, one of the most cited authors in the smart city field of study, examines also the different topics involved in the smart city implementation [15]. Certainly, all these themes are included in smart cities, but not in each smart city and not only these themes are included. Moreover, some of these themes sometimes overlap each other and the clearness of the Giffinger's definition is not satisfying. It says: “A Smart City is a city well performing built on the ‘smart’ combination of endowments and activities of self-decisive, independent and aware citizens”.

This definition is broad enough to include all the good initiatives carried out to improve the city quality, no matters which instruments, outcomes or actors are involved. This definition could be interesting for a theoretical debate about what a smart city is, but it is not very useful to drive its implementation and to measure the obtained results.

Examining also other smart city definitions, as listed in the chapter written by Cocchia, it emerges that there is a large disagreement between the academic view and the empirical view about smart cities. This disagreement regards the main component of a smart city: in the academic debate, it is the intellectual capital, in the empirical vision expressed by large companies such as IBM, Cisco and so on, the main component is the technology.

This different vision impacts on all the further aspects regarding the smart city: strategy definition, implementation, evaluation and performance measurement.

The academic vision considers the intellectual capital the most important resource to increase the smartness of a city. The label intellectual capital is to be interpreted in the broader meaning. It includes the culture of citizens, their educational level, their intellectual capability; but also the culture of companies, that is, trade marks, patents, know how, reputation on the market; and finally the city culture, represented by museums, theatres, cinemas, cultural events and everything could animate the cultural life in the city [16].

Depending on this vision, the smarter city is the one that has the larger cultural capital and is able to use its knowledge to choice the better solutions for the further development of the city quality. Investments in cultural initiatives are therefore welcome, but especially the city should use its awareness to promote sustainable development, equal economic growth and environmental quality in the urban areas.

Also the evaluation system is consequently designed depending on this intangible vision. Indicators regarding the cultural aspect of the city, the citizens and the public and private bodies resident in the city are the main proxy of the city smartness. To increase the cultural level—and by this way the smartness—of the city is the main instrument to further attract the best people and companies: more educated, more innovative, more profitable [17].

The business vision of a smart city is strongly based on the pivotal role of technology, especially the ICT. It derives from both the previous idea of digital city, and from the strong need to solve several concrete problems strongly affecting the life in large metropolis, such as traffic, pollution, energy consumption, waste treatment, water quality. These aspects are also near to the idea of green city and the environmental themes are an important part of the smart city goals.

In this smart city vision, initiatives to improve the city smartness are especially focused on some lines such as:

- energy production from renewable sources, to reduce energy cost, CO_2 emissions and to satisfy the increasing energy demand in urban areas;
- building efficiency, to reduce energy demand and consumption;
- local transport quality and greenness, to reduce pollution deriving from transport in cities;
- and so on.

The evaluation system to be applied to this different smart city vision is more tangible and based on physic indicators such as CO_2 emissions, greenhouse gases, waste tons, megawatts produced by renewable sources and so on [18]. It is important to outline that, even if the ultimate goal is to improve the citizens' quality of life, they are scarcely considered in this smart city vision and smart initiatives are often planned without their involvement. They are seen like the final addresser in the smart city value chain, but this value is not compared with their own expectations about the quality of life in city.

Even if these two smart city visions are quite clear in both academic papers and empirical studies or surveys about smart city, they are scarcely applied when a smart city plan is designed. As Thorne and Griffith explains in their chapter about the London Smart City development, and as it emerges from large literature surveys conducted by several authors in this book, the different smart city souls are merged each other and are not able to distinguish themselves in a smart strategy. Technological, cultural and environmental aspects are the core elements of a smart city, but their role is not the same and it is important to explicitly declare which aspect is the more important, what has the leading role and how this component interacts with the main stakeholders of the smart city strategy, that is, the citizens. To explicitly define the smart city vision and to align it with smart initiatives and desired outcomes is the first step to implement a successful smart city program.

3 Smartness, Public Value and Smart City Performance

What makes a city smart? And how it is possible to define the smartness of a city, and to measure it?

Even if a shared definition of a smart city still lacks, it is possible to describe which are the main characteristics of a smart city, which initiatives could improve its smartness and the most important goals to be reached. To measure created public value and smartness performance, all the goals and processes should be clearly defined and quantified. It requires a city strategic vision (that too often lacks) to sustain all the programs and projects carried out by a city to become smarter [19].

Dameri in chapter "Smart and Digital City: A Systematic Literature Review" tries to put the basis to define the smartness of a city starting from its core components: land, infrastructures, people and government.

Land means the territory, that is, the geographical area upon which the city rises up. Infrastructures are a large element, including all the physical, material components of a city such as buildings, streets, transport facilities, and so on. People includes all the citizens, not only the city inhabitants but also who works, studies or visits the city. Government means the local political bodies which have the power to govern the administrative aspects of the city.

For the first, to become smarter a city should improve the smartness of its core components. What do we mean with smartness? All the authors contributing to this volume agree to consider three main aspects of a smart city: effectiveness, environment consideration, and innovation [20, 21].

- Effectiveness means the capacity of a city to supply effective public and private services to several subjects, such as citizens, companies, not-for-profit organizations; and in detail to different categories of citizens such as students, workers, elder men and women, and so on. It requires to include the subjective role of several stakeholders in the smartness definition. Therefore, a smart city is not smart for itself, but if it creates public value for people.
- Environmental consideration regards the increasing impact that large cities have on the environmental quality of urban areas. One of the main pillars of smarter cities is to prevent a further environmental degradation. The main impacts regard energy consumption, air and water pollution, traffic congestion, land consumption. A smarter city therefore acts to reduce all these aspects to preserve the environmental quality.
- Innovation means that a smart city should use all the new and higher available technologies to improve the quality of its core components, to deliver better services and to reduce its environmental impacts. Technology is therefore a central aspect of smarter city, used at the service of smart initiatives for the quality of life in city.

To improve the smartness of its core components, a city should transform them into more effective, environmental and innovative ones [22].

Therefore, a smarter land means cleaner territory, water and air, a reduced consumption of land for new buildings, environmental reclamation and so on. Smarter infrastructures should be cleaner, more effective in serving the citizens and answering to their needs, using high technology, ICT and mobile devices to spread e-services and information. Smarter people means citizens more informed, more aware about the city goals and the role of technologies in improving the quality of urban land, infrastructures and services, a easier access to the Internet and all the mobile and on-line services and finally a strong decreasing of the digital divide. A smarter government uses ICT and all the new technologies to implement e-government and e-democracy, improving the quality and accessibility of supplied public services and the people satisfaction for the local administration [23, 24].

However, all these activities to improve the smartness of a city are not enough to realize public value to be enjoyed by citizens. Indeed, the creation of public value should be the final goal of a smarter city, but it requires that all the projects and initiatives would be addressed to the citizens [25, 26]. Public value is a complex idea, as it includes several different dimensions [27]:

- it requires to create both economic and social values, that are difficult to merge and sometimes in conflict each other;
- it requires to create value for different stakeholders, that have different expectations not ever compatible each other;
- it requires to create value respect to different dimensions of the life in city, and it further requires to understand which are the real needs and the priorities to carry on.

To create public value in a smart city program means therefore to put together a large set of variables and to compose them into a well-defined general framework, able to collect the needs, the expectations and the perception of citizens respect to the smart city for their daily life [28, 29].

To measure the public value created and supplied thanks to a smart city program is therefore a complex task, but such important as the implementation of the smart initiatives. Indeed, examining some smart city cases all over the world it emerges that often:

- smart city benefits are not defined,
- they are not measured,
- and furthermore they are not communicated.

Even if the smart city program produces improvement in the daily life of the citizens, they are not informed about that, nor involved in the definition of their priorities and not aware about the impact of smart projects in the quality of their own city [30].

To measure the smart city performance, that is, the capacity of a smart city program to really create and spread public value, is the major challenge to be faced, to grant the transparency and the awareness about the smart wave in city, and to prevent that this trend would finish before it starts to create real benefits in urban areas. The importance of this topic emerges also from this volume, as five chapters are explicitly devoted to this aspect of smart city implementation [31].

Negre and Sabroux face the problem of prioritizing smart initiatives. In their chapter, several possible smart actions have been briefly described. Each city has a enormous set of possible actions to be carried out, but… which to choice? Considering that not all the cities have the same characteristics and problems, nor their citizens have the same needs and expectations, and that financial resources are not enough to implement all of the desired smart projects.

Fontana examines how some cities already committed in a smart city plan are defining their own strategies and linking them with the creation of public value in a sustainable way. It requires to include into the smart city strategic vision all the stakeholders, such as, citizens, companies, public authorities, not-for-profit organizations. Each of them rightly wants a part of the created public value, but the expectations of all of them are not easy to harmonize.

The importance of citizens is outlined also by Palumbo and Cossetta. In their chapter regarding social innovations, they introduce the idea that social and open innovations are very important to create novel solutions able to improve the quality of life in cities. Using Living Labs to explore the needs, the expectations and the ideas of citizens abut smart city, it is possible to obtain better performance, more aligned with the co-production of public value.

Zuccardi Merli and Bonollo introduce the crucial topic of performance measurement. Performance is not only the smartness of a city, but a more complex concept: it means to measure the advances of a city towards its capacity to deliver a better quality of life to everybody. Also these authors outline the importance of the citizen involvement, the role of different stakeholders and the need to build a model able to measure smart city performance. They also test their theoretical model on a set of Italian and European smart city cases.

Baccarne, Mechant and Schuurman analyse created value in a smart city case, Ghent Smart City in Belgium; they face an important aspect, that is, the sustainability of smart programs over the time. Indeed, all the smart city projects implemented till now are pioneer implementations, especially aiming at testing new solutions to find best practices in smart city realization. However, it is time now to overcome this phase and to transform demonstrators towards real sustainable value.

4 Specific Smart Projects

It emerges from both all the chapters in this book and the international literature that the final aim of a smart city program is to improve the quality of the city and in the meantime the quality of life in city. These two aspects—city quality and life quality in city—are not the same thing, but they are strictly linked. What especially links these two different way to understand the smart city benefits are the specific smart projects [32].

The smartness of a city is indeed composed by several dimensions. Giffinger, one of the most cited authors, indentifies six different dimensions of smart city, as showed in Fig. 1: Smart mobility, smart environment, smart people, smart living, smart governance, smart economy [15].

However, it is difficult to use this schema to classify the specific smart projects, and furthermore to use this classification to build an evaluation framework, because some of these aspects are linked each other or are overlapped in some aspects. For example, a new public local transport system, based on low carbon emissions, impacts on both smart mobility and smart environment. It is moreover difficult to find projects not impacting on smart living, as this dimension seams to summarize all the benefits deriving from smart initiatives [33].

Therefore, it is perhaps better to use a descriptive framework, based on the core components of a smart city—land, infrastructure, people, government—composed by a project portfolio and aiming at a better quality of life and/or of city, measured by a set of key performance indicators representing the different benefits created

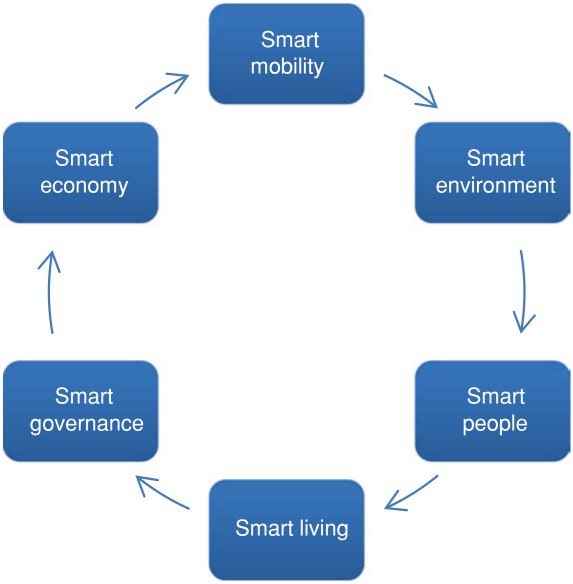


Fig. 1 The Smart city dimensions (*Source* [15])

by each project. It is very difficult to separate the benefits produced by a project in different streams: it is better to describe and to measure the numerous benefits generated by a sole project.

We can imagine that:

- the core components should become more and more smart;
- the smart initiatives are actions planned to both improve the smartness of the core components of a city and impact on the quality of life;
- the quality of life depends on both the smartness of the core components and the capacity of single smart projects or set of smart projects to impact on one or more dimensions of the daily life in city [34].

Therefore, single projects are the instrument to realize the smart city. These projects should have some characteristics, such as use advanced technological solutions, harmonize environment and economy, and address the needs and expectations of citizens. Till now, smart projects are mainly focused on some themes such as buildings energy efficiency, greenhouse gases reduction, broadband diffusion, e-services delivery, mobile government and so on. It would be more and more important to offer technical solutions to city problems, but also to include each project into a comprehensive smart framework [35].

In this book, two chapters regard crucial topics for smart city realization. Di Bella et al. analyze the so-called smart security systems, that is, applications using ICT to improve the safety and security in urban areas, especially the more degraded. This is indeed a good example of a smart project, aiming at improving both the citizens' quality of life and the city quality tout court: the safety of neighborhoods is a real benefit for citizens and in the meantime the reputation of a city increases.

Carrabs, Cerulli and Sciomachen apply the smart city framework to a logistic problem, regarding goods distribution inside the city boundaries. The aim of their study is to suggest a mathematical programming model to face and solve the inconveniences deriving from business to consumer goods deliver in city centre. Optimizing the vehicle routing, it is possible to reduce traffic and pollution, just two main goals of a smart city strategy.

5 Conclusions and Further Works

Smart city is one of the most interesting research themes in the latest few years. One of the main reasons is that Smart city is a multidisciplinary topic, impacting on human, social, economic and technical research fields. The need to face with the harder and harder problems deriving from increasing dimension cities, along with the desire to gain the higher benefits from the urban life, are formidable engines that sustain the research about smart city.

Till now, this topic has been a pioneering field, both in theoretical research and in empirical applications. Academic researchers are still trying to understand what exactly a smart city is, and local governments are trying to realize prototypes of smart city or, at least, of smart projects. But to realize the expected returns from smart city projects, it is necessary to overcome the first stage of smart city study and realization and to increase the maturity level of this promising urban strategy.

This book is the result of a series of writings from all over Europe; researchers give their contribution about this topic, searching to clarify the concepts still dark and confused. They agree about the most important themes to be deepened and interesting also for further works:

1. The definition of a smart city, to be shared and useful to clarify which initiatives are included into a smart city strategy;
2. The smart city goals and the measurements needed to evaluate its success or failure;
3. The collection of best practices, the repeatability of prototypes and the financial sustainability of smart initiatives.

The definition of a smart city is indispensable to trace its perimeter and to understand which initiatives can be considered smart and which can not. Moreover, a standard definition is also the first step for each city to specify its own vision of a smart city strategy and to build a comprehensive smart city framework able to link together all projects and initiatives.

The definition and the comprehensive smart city framework are the necessary basis on which to build the smart city goals system. The multisciplinarity of a smart city program requires to define a set of objectives to be reached. To support the monitoring of projects and initiatives, all the goals should be measurable and key performance indicators are the instrument to evaluate the progress of a smart strategy. Citizens should even be involved, both in the plan phase and in the smart city implementation steps; communication is at the centre of a shared participation in defining smart city goals and in spreading awareness about the smart city role and benefits for people.

Finally, smart cities are now leaving their youngness, but they need to reach their maturity, to extend best practices collected in smart city pioneering implementation all over the world and increase the return on investments—financial, but also political, social, human—of smart projects. Local governments, together with businesses, universities, not-for-profit organizations and the citizens themselves should share their work to grant the maximum of benefits delivery to everybody, so that a smart city could also be considered an inclusive city.

All these topics are examined in this book, establishing sound basis for further studies about all of them.

References

1. Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart cities in Europe. *Journal of Urban Technology, 18*(2), 65–82. Routledge.
2. Toh, M. H., & Low, L. (1993). The intelligent city: Singapore achieving the next lap: practitoners forum. *Technology Analysis & Strategic Management, 5*(2), 187–202.
3. Dameri, R. P., & Cocchia, A. (2013). Smart city and digital city: twenty years of terminology evolution. *Proceeding of ITAIS conference*, Milano, December.
4. Tokmakoff, A., & Billington, J. (1994). Consumer services in smart city Adelaide. In Bjerg and Borreby (Eds.), *Home-oriented informatics, telematics and automation.*
5. Oberti, I., & Pavesi, A. S. (2013). The triumph of the smart city. *TECHNE: Journal of Technology for Architecture & Environment 5*.
6. Mainka, A. et al. (2014). Government and social media: a case study of 31 informational world cities. arXiv preprint arXiv:1401.4533.
7. Dameri, R. P. (2013). Searching for smart city definition: a comprehensive proposal. *International Journal of Computers & Technology, 11*(5), 2544–2551.
8. Mulligan, C. E. A., & Olsson, M. (2013). Architectural implications of smart city business models: an evolutionary perspective. *IEEE Communications Magazine, 51*, 6.
9. Paroutis, S., Bennett, M., & Heracleous, L. (2013). A strategic view on smart city technology: the case of IBM smarter cities during a recession. *Technological Forecasting and Social Change*.
10. Suppa, A., & Zardini, A. (2011). The implementation of a performance management system in the Italian army. *Communications in Computer and Information Science, 210*(3), 139–146.
11. Vakali, A., Angelis, L., & Giatsoglou, M. (2013). Sensors talk and humans sense towards a reciprocal collective awareness smart city framework. *IEEE International Conference on Communications Workshops* (ICC).
12. Kourtit, K. et al. (2013). 11 An advanced triple helix network framework for smart cities performance. *Smart Cities: Governing, Modelling and Analysing the Transition 196*.
13. Pardo, T., Taewoo, N. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. *Proceedings of the 12th Annual International Conference on Digital Government Research* (pp. 282–291). ACM, New York.
14. Hollands, R. G. (2008). Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? *City, 12*(3), 303–320.
15. Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovi, N., & Meijers, E. (2007). Smart cities: ranking of European medium-sized cities. Centre of Regional Science (SRF), Vienna University of Technology, Vienna, Austria, from http://www.smart-cities.eu/download/smart_cities_final_report.pdf.

16. Leydesdorff, L., & Deakin, M. (2011). The triple-helix model of smart cities: a neo-evolutionary perspective. *Journal of Urban Technology, 18*(2), 53–63.
17. Paskaleva, K. A. (2009). Enabling the smart city: the progress of city e-governance in Europe. *International Journal of Innovation and Regional Development, 1*(4), 405–422.
18. Al-Hader, M., & Rodzi, A. (2009). The smart city infrastructure development and monitoring. *Theoretical & Empirical Researches in Urban Management, 2*, 11.
19. Zygiaris, S. (2013). Smart city reference model: assisting planners to conceptualize the building of smart city innovation ecosystems. *Journal of the Knowledge Economy, 4*(2), 217–231.
20. Casalino, N., Buonocuore, F., Rossignoli, C., Ricciardi, F. (2013). Transparency, openness and knowledge sharing for rebuilding and strengthening government institutions. *IASTED Multiconferences—Proceedings of the IASTED International Conference on Web-Based Education*, WBE, pp. 866–871.
21. Dameri, R. P. (2012) Defining an evaluation framework for digital cities implementation. *IEEE International Conference on Information Society* (*i-Society*).
22. Chourabi, H. et al. (2012). Understanding smart cities: an integrative framework. *IEEE 45th Hawaii International Conference on System Science* (*HICSS*).
23. Alawadhi, S., et al. (2012). *Building understanding of smart city initiatives "Electronic Government"* (pp. 40–53). Berlin: Springer.
24. Nam, T., & Pardo, T.A. (2011). Smart city as urban innovation: focusing on management, policy, and context. *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance*. ACM.
25. Moore, M. H. (1995). *Creating public value: Strategic management in government*. Cambridge: Harvard University Press.
26. Sorrentino, M., & Niehaves, B. (2010). Intermediaries in E-inclusion: a literature review. Paper presented at the *43rd Hawaii International Conference on System Sciences* (HICSS-43, 2010), Kauai, Hawaii, USA, IEEE.
27. Benington, J., & Moore, M. H. (Eds.) (2010). Public value: Theory and practice. Basingstoke: Palgrave Macmillan.
28. O'Flynn, J. (2007). From new public management to public value: paradigmatic change and managerial implications. *Australian journal of public administration, 66*(3), 353–366.
29. Walravens, N., & Ballon, P. (2013). Platform business models for smart cities: from control and value to governance and public value. *IEEE Communications Magazine, 51*, 6.
30. Anttiroiko, A. V., Valkama, P., & Bailey, S. J. (2013). Smart cities in the new service economy: building platforms for smart services. *AI & Society* (2013), 1–12.
31. Bowerman, B. et al. (2000). The vision of a smart city. *2nd International Life Extension Technology Workshop*, Paris.
32. Al-Hader, M. et al. (2009). Smart city components architicture. *IEEE International Conference on Computational Intelligence, Modelling and Simulation*, CSSim'09.
33. Glaeser, E. L., & Berry, C. R. (2006). Why are smart places getting smarter. Rappaport Institute/Taubman Center Policy Brief 2.
34. Shapiro, J. M. (2006). Smart cities: quality of life, productivity, and the growth effects of human capital. *The Review of Economics and Statistics, 88*(2), 324–335.
35. Lee, J. H., Phaal, R., & Lee, S Ho. (2013). An integrated service-device-technology roadmap for smart city development. *Technological Forecasting and Social Change, 80*(2), 286–306.

Smart and Digital City: A Systematic Literature Review

Annalisa Cocchia

Abstract The concept of Smart City embraces several definitions depending on the meanings of the word "smart": intelligent city, knowledge city, ubiquitous city, sustainable city, digital city, etc. Many definitions of Smart City exist, but no one has been universally acknowledged yet. From literature analysis it emerges that Smart City and Digital City are the most used terminologies in literature to indicate the smartness of a city. This Chapter explores the literature about Smart City and Digital City from 1993 to the end of 2012 in order to investigate how these two concepts were born, how they have developed, which are the shared features and differences between them. To accomplish with these goals, three steps were followed: (1) to set up a search strategy for systematic literature review to collect a representative subset of papers about Smart City and Digital City using Google Scholar; (2) to store the selected subset in an ad-doc database to synthesize the literature review; (3) to organize the literature review subset to extract quantitative and qualitative data and information about Smart City and Digital City evolution. The author proposes a literature review taxonomy through five specific analysis: (1) time analysis, to explore the causes of the trend of Smart City and Digital City literature in the latest twenty years; (2) terminology analysis, to examine how and where these two ideas were born and what have been the main events influenced their development; (3) definitions analysis, to select and compare the most cited and validated definitions of Smart City and Digital City trying to identify similarities, differences or overlaps between these two concepts; (4) typology analysis, to investigate if Smart City and Digital City are included into a specific urban strategy pursued by government or if they face specific urban problems without a comprehensive framework; (5) geographic analysis, to understand where are the largest concentrations of Smart Cities and Digital Cities in the world and which are their main characteristics and best practices.

A. Cocchia (✉)
Department of Economics, University of Genoa, Genoa, Italy
e-mail: cocchia@economia.unige.it

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_2,

Keywords Smart city • Digital city • Literature review • ICT • Urban strategies

1 Introduction

During the latest years of XX century, two important phenomena have been emerging: urbanization and information and communication technologies (e.g. ICT). Eighties' and nineties' technological advancement and economic growth contributed to increase well-being, mainly in the greater urban centers. This fostered the urbanization leading to a progressive abandonment of rural areas towards greater cities and metropolis, which can offer many opportunities in terms of work, education, social life and so on. People strong inclination to concentrate in cities generated both positive and negative effects at global level [1]. On one hand it causes the increasing of cultural level, the creation of new job opportunities and an improvement of economic conditions. On the other hand, concentration in cities increased traffic jam, carbon dioxide, greenhouse gases emissions and waste disposal with consequences on health conditions. City dimension drives energy and natural resources demand, the need of territory redevelopment and adequate infrastructures availability. In this scenario, to save the earth and people health, the idea of smart cities emerges, that is, cities able to solve urban issues paying attention to the environment. For this reason, in the nineties, the concept of smart growth has begun to spread: it implies a community-driven reaction to solve traffic congestion, school overcrowding, air pollution, loss of open space and skyrocketing public facilities cost [2].

In the international context, in order to achieve the objectives established in the Kyoto Protocol, the Smart City concept was born and has been adopted by many institutions (e.g. European Commission, Setis-EU, OECD, etc.) which labeled as "smart" initiatives and projects relevant to cities sustainability. Indeed, if we "google" the term "smart city", we can find more and more results about: city sustainable initiatives, institution road-maps to enhance green growth and quality of life, the usefulness of ICT infrastructures, the involvement of citizens in public life, the need to reduce digital divide, and so on. But giving a comprehensive definition of Smart City is far to be done. A major hurdle in identifying such definition is the ambiguity of meanings attributed to the word "smart" and to the label "Smart City". Some examples of these several meanings are: Digital City, Wired City, Knowledge City and Green City, "which often link together technological informational transformations with economic, political and socio-cultural change" [3]. All those meanings are somehow part of the fuzzy concept Smart City and they cannot be regarded as mere correlated themes of it. Especially Smart City and Digital City are often used without specifying their similarities and differences. A brief analysis of scientific literature and paper title shows that smart city and digital city are most recurrent terms, but their meaning is rarely clarified.

This paper aims to light the shared features and dissimilarities between Smart City and Digital City concepts looking at a sound definition of both. To lead this study, the author inquires about the beginnings of Smart City and Digital City phenomena, considering the time frame 1993–2012. In this context, the author organized the research in the following steps:

- paragraph 2 defines the research method, aiming to identify and organize a subset of papers relevant for the literature review analysis;
- paragraph 3 examines the collected data trying to answer to the specific questions through five analysis:

  1. *time analysis* answers to the Research Question 1: "How and when Smart City and Digital City concepts were born?". This analysis aims to show the trend of papers about Smart City and Digital City in the time frame 1993–2012 and to investigate about the drivers of this trend;
  2. *terminology analysis* answers to the Research Question 2: "Which events mainly influenced the development of Smart City and Digital City ideas?". This analysis aims to show the time distribution of papers regarding Smart City and Digital City during the latest twenty years, highlighting the most important causes which influenced the widespread of these two concepts; definitions analysis answers to the research Question 3: "Which are the main shared features, differences and overlaps between Smart City and Digital City contents?". This analysis aims to compare each other the most recurrent and acknowledged definitions about Smart City and Digital City, identifying their similarities and differences;
  3. *typology analysis* answers to the Research Question 4: "Are Smart City and Digital City included into a comprehensive urban strategy? Or are they the sum of standalone projects? In this context do they follow a top-down or a bottom-up approach?". This analysis uses papers labeled as "empirical study" or "theoretical study" regarding smart/digital cities/initiatives during the time frame 1993–2012 to understand if their origin is more theoretical and top-down, or more empirical and bottom-up;
  4. *geographic analysis* answers to the Research Question 5: "How these two types of city strategies are widespread in the world?". This analysis aims to show the location of empirical studies regarding Smart City or Digital City implementation, highlighting the concentration of these concepts per continents;

- paragraph 4 addresses concluding remarks.

2 Methods: Search Strategy for Systematic Literature Review

The literature review about Smart City and Digital City has been carried out using the methodological model proposed by Vom Brocke et al. [4] in their study about the importance of rigor in documenting the literature search process. This methodological model is based on a five phases framework for the literature search process. These phases are: (1) definition of the review scope, (2) conceptualization of topic, (3) literature search, (4) literature analysis and synthesis, (5) research agenda.

In the following paragraphs these phases are introduced referring to the literature review about Smart City and Digital City.

2.1 Definition of Review Scope

In order to clearly define the scope of this literature review, the author refers to an established taxonomy presented by Cooper [5] that includes six characteristics for literature review: (a) focus; (b) goal; (c) organization; (d) perspective; (e) audience; (f) coverage.

a. *Focus* is the central area of interest to the reviewer. This area could concern: research outcomes, research methods, theories, practices or applications; this literature search focus regards all types of papers, from theoretical to application-centered ones;
b. *Goal* regards what the author hopes the review will fulfill. The aim of literature review could regard: integration (such as, to clarify contradictory ideas or to bridge the gap between theories and practices), criticism (such as, to critically examine the literature to demonstrate the unwarranted previous theories), central issue (such as, what has been studied in the past, what researchers will study in the future, what has hindered the development of some topics, and so on); the aim of this study is to synthesize past literature and to identify the central issue of the literature review about Smart City and Digital City, that is to investigate how these two concepts were born, how they have developed, what are their similarities, differences and overlaps;
c. *Organization* concerns how the reviewer organizes his search study. The literature review could be organized by: chronological order, conceptual order (that is, to group the same ideas), methodological order (that is, to group the same methods of work); this literature is sorted by chronological order first and by conceptual order after;
d. *Perspective* is the point of view of reviewer in discussing the literature. The reviewer could lead the study with: a neutral position (he plays the impartial role as an honest "judge") or an espousal of position (he plays the role of an "advocate"); the author considers worthwhile to adopt an essentially neutral literature search perspective, because there is no interest to foster a specific position or policy about the topic;
e. *Audience* concerns groups of people (such as specialized researchers, general researchers, practitioners, policy makers, general public and so on) whom the review is addressed; the audience of the literature review are specialized scholars and industry makers.

Table 1 The Cooper's taxonomy applied to the smart city and digital city literature review

	Characteristic	Cooper's options	Author choice
a	Focus	Type of papers involved (methodological, theoretical, practices, applications, outcomes)	All types of paper
b	Goal	Integration, criticism, central issue	Central issue
c	Organization	Chronological, conceptual, methodological	Chronological first, conceptual after
d	Perspective	Neutral, espousal of a position	Neutral
e	Audience	Groups of people whom the review is addressed	Specialized scholars and industry makers
f	Coverage	Exhaustive, with selective citation, representative, central, pivotal	Representative

f. *Coverage*, it regards how the reviewer searches the literature and how he makes decisions about the suitability and quality of documents. The coverage could be: exhaustive, exhaustive with selective citation, representative, central or pivotal; the author decided to choose a reasonably representative coverage.

Table 1 summarizes the choices made by the author, regarding the Cooper's taxonomy about the review scope.

2.2 *Conceptualization of Topic*

Vom Brocke et al. [4] suggest that "a review must begin with a broad conception of what is known about the topic and potential areas where knowledge may be needed". Therefore, in order to choose the key concepts on which to base the literature review, the author began the study on Smart City by looking:

- several papers about the meaning of the word "smart" (in particular the paper of Hollands [3] and the IBM report [6]), because Smart City is a broad concept including many aspects of urban life, such as urban planning, sustainable development, environment, energy grid, economic development, technologies, social participation, and so on; therefore, also the word smart assumes a large range of meanings, linked with its different field of application;
- several papers about the different terminologies which identify a "smart" city, because it is not clear if these different terms want to say the same thing or if they define different cities, strategies and technologies (especially the paper of Pardo and Nam [2], Dameri [7], Su et al. [8]);
- several papers about the Smart City definitions (especially the paper of Pardo and Nam [2, 9], Chourabi et al. [10]).

However, exploring these papers, we cannot find a comprehensive definition accepted by academics, businesses and institutions about what Smart City is and which are its key elements and boundaries are. The difficulty to define Smart City regards mainly two aspects:

1. the adjective “smart”, because it depends on the meaning we attribute to this word. In literature, several typologies of city refer to Smart City concept, such as Intelligent City, Knowledge City, Wired City, Digital City, and so on [4];
2. the label “smart city”, because it is a fuzzy concept and it is used in ways that are not always in accordance each other. There are many cities that define themselves as Smart City when they identify some own characteristics as “smart”, but without referring to a standard meaning.

For these reasons, nowadays, a unique definition of Smart City does not exist yet [2]. Moreover, from this first literature review it emerges that there are many terminologies of Smart City, but to understand if each of them could be considered as synonymous of Smart City, it is necessary to clearly define if they have some shared features, overlaps or differences. To achieve this goal, the author analyzed and compared different definitions of city linked to the label “smart city”. Table 2 shows this comparison, listing the definition and the reference. For each definition, there are in evidence some words to highlight the meaning of these concepts: the bold character is used to outline the human component of different city concepts; while the italic character is used to outline the applied technologies.

From the literature analysis, it emerges that all these concepts are not in contradiction with each others and they disclose some shared features and are partially overlapped [7].

Pardo and Nam [2] organizes these definitions in dimensions depending on some recurrent shared characteristics, in order to define the most used terms. These dimensions are:

1. *Technology dimension*; it is based on the use of infrastructures (especially ICT) to improve and transform life and work within a city in relevant way. This dimension includes the concepts about Digital City; Virtual City, Information City, Wired City, Ubiquitous City and Intelligent City;
2. *Human dimension*; it is based on people, education, learning and knowledge because they are key drivers for the smart city. This dimension includes the concepts about Learning City and Knowledge City;
3. *Institutional dimension*; it is based on governance and policy, because the cooperation between stakeholders and institutional governments is very important to design and implement smart city initiatives. This dimension may include the concepts about Smart Community, Sustainable City and Green City.

Finally, we can see that Digital City also embraces several meanings of “smart”, such as virtual city, cyber city, wired city, ubiquitous city and so on [11, 12–14]. Moreover, Digital City is sometimes considered as a Smart City based on ICT infrastructures, because one of the most important technologies used to support Smart City strategies is ICT [7]. Therefore, from this analysis, it appears that Digital City is the most recurrent terminology linked to the meaning of Smart City.

Table 2 The different meanings of smart city

Concept	Definition	Reference
Wired city	"Wired cities refer literally to the laying down of *cable and connectivity* not itself necessary smart"	Hollands [3]
Virtual city	"Virtual City concentrates on *digital representations* and manifestations of cities"	Schuler [11]
Ubiquitous city	"Ubiquitous city (U-City) is a further extension of digital city concept. This definition evolved to the ubiquitous city: a city or region with *ubiquitous information technology*"	Anthopoulos et al. [39]
Intelligent city	"Intelligent cities are territories with high capability for learning and innovation, which is built-in the creativity of their population, their institutions of knowledge creation, and their *digital infrastructure* for communication and knowledge management"	Komninos [40]
Information city	"Digital environments collecting official and unofficial information from local communities and delivering it to the public via *web portals* are called information cities"	Anthopoulos et al. [39]
Digital city	"The digital city is as a comprehensive, *web-based representation*, or reproduction, of several aspects or functions of a specific real city, open to non-experts. The digital city has several dimensions: social, cultural, political, ideological, and also theoretical"	Couclelis [41]
Smart community	"A geographical area ranging in size from neighborhood to a multi-county region whose residents, organizations, and governing institutions are using *information technology* to transform their region in significant ways. Co-operation among government, industry, educators, and the citizenry, instead of individual groups acting in isolation, is preferred"	California Institute [42]
Knowledge city	"A Knowledge City is a city that aims at a knowledge-based development, by encouraging the continuous creation, sharing, evaluation, renewal and update of knowledge. This can be achieved through the continuous interaction between its citizens themselves and at the same time between them and other cities' citizens. The citizens' knowledge-sharing culture as well as the city's appropriate design, *IT networks and infrastructures* support these interactions"	Ergazakis [43]
Learning city	"The term 'learning' in 'learning cities' covers both individual and institutional learning. Individual learning refers to the acquisition of knowledge, skills and understanding by individual people, whether formally or informally. It often refers to lifelong learning, not just initial schooling and training. By learning, individuals gain through improved wages and employment opportunities, while society benefits by having a more flexible and technological up-to-date workforce"	OECD [44]

(continued)

Table 2 (continued)

Concept	Definition	Reference
Sustainable city	"Sustainable city uses *technology* to reduce CO_2 emissions, to produce efficient energy, to improve the buildings efficiency. Its main aim is to become a green city"	Batagan [45]
Green city	"Green City follows the Green Growth which is a new paradigm that promotes economic development while reducing greenhouse gas emissions and pollution, minimizing waste and inefficient use of natural resources and maintaining biodiversity"	OECD [46]

For these reasons, the author decided to focus on Smart City and Digital City; these two topics will be analyzed together because they are the most used and representative terminologies in the literature to indicate the smartness of the city; but they are quite different to require a distinct analysis and further comparison.

2.3 Literature Search

This phase "involves database, keyword, backward and forward search, as well as an ongoing evaluation of sources" [4]. To conduct this literature search process, the author evaluated the following search strategy steps: (a) to choose the database source; (b) to choose keywords and search criteria; (c) to choose if to apply backward and forward search; (d) to evaluate the literature subset suitability.

a. First, it is needed to choose the *database source* among the available ones. The selected on-line database has been Google Scholar, because it includes a broad field of publications (especially such as papers and journal articles) which focus on the chosen topic.
b. Second, it is needed to choose the most suitable *keywords* and *search criteria* in order to extract a representative subset from the selected database. In the present case, the search was conducted between February and May 2013. The system was request to search the words "*Smart City*" *OR* "*Digital City*" *OR* "*Smart Cities*" *OR* "*Digital Cities*" only in the title of paper and excluding all citations and patents. In this way, the search results included 987 papers. Then, the Google Scholar database was request to sort all the results by year of publication within 1993–2012 range. The author chooses this twenty years range in order to have a reasonable representative subset which does not include the work in progress (such as it could happen including 2013). After the filtering, the search results reduced to 843 papers which have been stored in an ad-hoc database. This database has been called "Literature Review Storage DataBase"

(LRS-DB) where each record corresponds to a paper. It is characterized by the following attributes:

– publication year;
– authors' name;
– title of work;
– source which identifies the typology of scientific publication;
– affiliation of authors, country included;
– abstract;
– keywords used by authors to index their work;
– citations only when superior to twenty;
– tag "smart" or "digital" in order to classify papers into Smart/Digital City label on the base of adjective used in the title;
– type of study identifies if the contribution is a theoretical study, a case study or a report;
– if case study, the object of it (that is, a smart or digital project or a smart or digital city);
– the name of the city in case of empirical study;
– the continent of city in case of empirical study;
– abstract available to delete the papers which are not abstract available.

c. Third, it is needed to choose to apply *backward and forward search*. In the present case, the amount of papers was considered an appropriate pool to investigate how and when Smart City and Digital City concepts were born, how they have been developing and if there are some shared features between these terminologies. Thus, the author decided not to apply any backward search nor forward search.
d. Fourth, *evaluation* in "all phases means limiting the amount of literature identified by keyword search to only those articles relevant to the topic at hand" [4]. In this phase, the LRS-DB was used as a source input platform and some criteria were applied to it to restrict the search. Indeed, the author removed: duplicates, thesis, power point presentations, white papers, book's introductions, competition announcements, all works which are not in English language and/or have not the full abstract available. The application of these criteria resulted in the exclusion of 115 papers, leading to a total 705 ones relevant to the present study.

Figure 1 shows in a sketch the steps of these described systematic review process.

The LRS-DB is an important tool for data-mining aiming to fulfill the following objectives around the Smart City and Digital City concepts:

- to filter the literature;
- to identify how Smart City and Digital City have been evolving during time;
- to identify research trends during the last two decades;
- to pinpoint the most studied research themes;
- to pinpoint the less studied research themes which can perhaps be expanded in future.

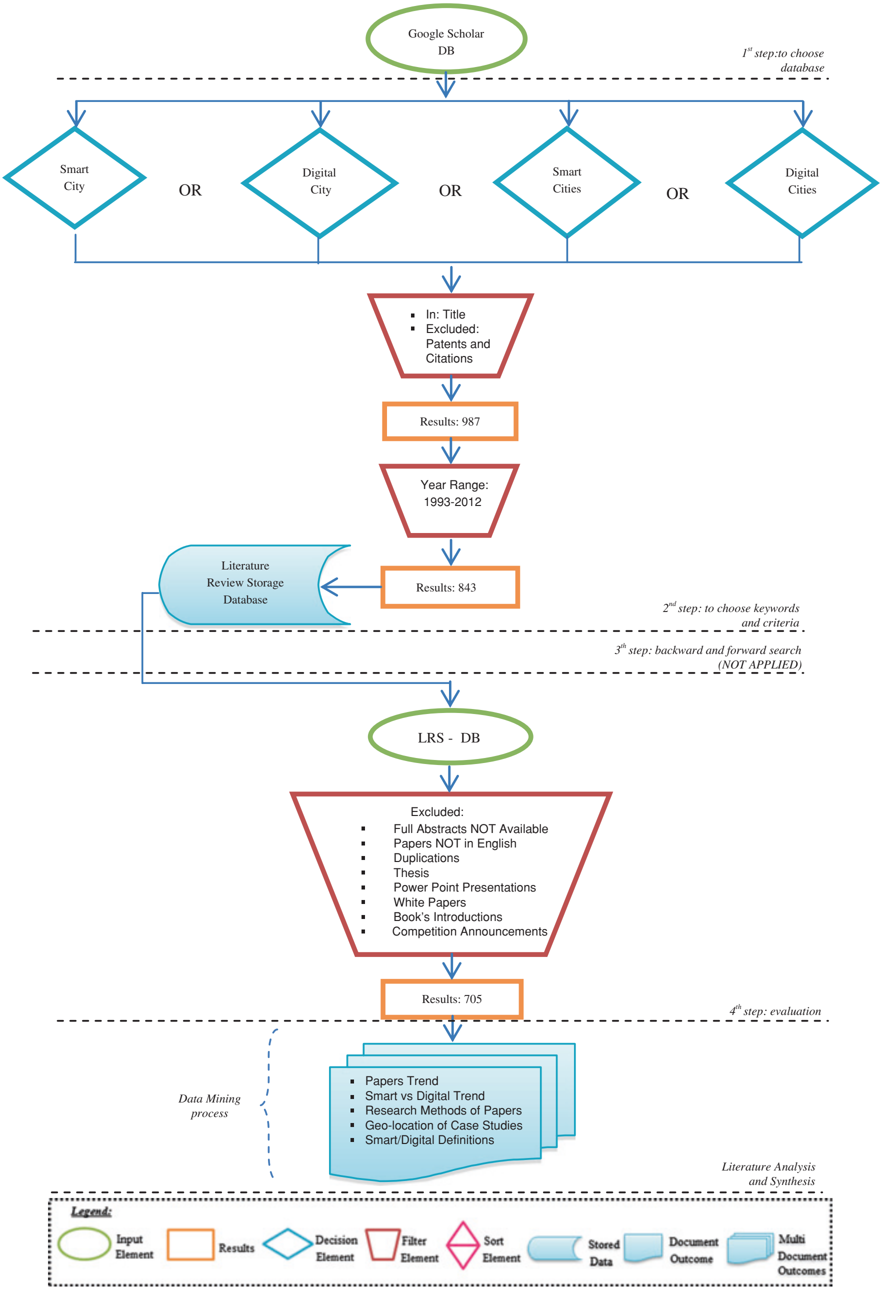


Fig. 1 Search strategy for systematic review

A part of these goals will be reached in this work,[1] while others are going to be accomplished in the future research studies.

2.4 Literature Analysis and Synthesis

"After collecting sufficient literature on a topic it has to be analyzed and synthesized" [4]. Therefore, the aim of this phase is to organize the papers stored in the LRS-DB to analyze systematically the collected literature. To accomplish with this goal, the 705 papers were organized to investigate about:

a. *time analysis*, it explores the evolution of researches about Smart City and Digital City during the latest twenty years. To achieve this purpose, the stored papers were organized by years of publication to count them per each year. The output of this analysis is to show in a graphic the trend of papers about Smart City and Digital City idea in the time frame 1993–2012;
b. *terminology analysis*, it explores how and when Smart City and Digital City concepts have being conceived and the relationships between the two topics. To achieve this goal, the stored papers were organized by years of publication and by "smart" or "digital" label, according to the adjective used in the title of paper. The output of this analysis shows in a graphic the time distribution of papers regarding Smart City or Digital City during the latest twenty years;
c. *definitions analysis*, it explores the most cited definitions of Smart City and Digital City, to compare their meanings and contents. To accomplish with this goal, the stored papers were organized by the most recurrent and validated definitions, according to the citation number and to the paper focus;
d. *typology analysis*, it explores if Smart City and Digital City are two initiatives which follow a specific urban strategy defined by governments (they have in this case a top-down approach) or if they solve specific and stand-alone urban issues without a comprehensive strategic vision (therefore they have a bottom-up approach). To achieve this aim, the stored papers were organized by years of publication and by "empirical study" or "theoretical study" label, according to the research method highlighted in the abstract of paper. The output of this analysis is to show in a graphic the distribution of papers labeled as "empirical study" or "theoretical study" regarding smart/digital cities/initiatives during

[1] Thanks to the use of LRS-DB, in this chapter will be investigated about the evolution of Smart City and Digital City concepts during the latest twenty years, in order to understand: how and where these two concepts were born, what causes mainly influenced their evolution, if Smart City and Digital City follow a bottom-up approach, where are the most large concentrations of Smart cities and Digital cities in the world, which are the empirical cases more frequently studied by researchers, how much these two topic are overlapping strategies and how much they are different. Therefore, this study is the first step to deep explore other research dimensions aiming to clearly design the contents and the boundaries of Smart City and Digital City idea.

the time frame 1993–2012; it helps to understand if the empirical studies come first, or after, the theoretical study of these topics, that is, if the bottom-up approach prevails on the top-down one, or vice versa;

e. *geographic analysis*, it explores where Smart cities and Digital cities are more concentrated in the world. To achieve this purpose, the stored papers were organized by "empirical study" label and by the "city of empirical study", according to the paper abstract. The output of this analysis shows in a graphic the distribution of empirical studies regarding Smart City or Digital City implementation all over the world.

2.5 Research Agenda

"The literature search process never comes to a definitive end" [15]. The final purpose of this literature review is not only to clarify the similarities and differences between Smart City and Digital City, or to find a good definition to identify both of them; but also to result in a new research agenda, which should be more insightful than the research question posed at the beginning. This new research agenda will aim to deep the characteristics of Smart City and Digital City, investigating about the contents of the papers included into the LRS-DB and also collecting several empirical case studies, to verify if the theoretical definitions designed in the present works are suitable to embrace the real implementation of Smart City or Digital City experiences all over the world.

3 Results

This section describes and explains the results obtained by the analysis of the LRS-DB, answering to the Research Questions #1, 2, 3, 4, 5 exposed in the previous paragraph.

3.1 Time Analysis

The purpose of time analysis is twofold; it aims both to analyze the time trend and distribution of researches regarding Smart City and Digital City and to understand which are the main determinants of this time trend.

To accomplish these objectives, the 705 papers stored in the LRS-DB were organized by chronological order, classified depending on publishing year to count them.

Figure 2 shows the number of papers about Smart City and Digital City during the latest twenty years. As trend line highlights, the first study concerning this topic is dated 1994. Between this start point and 1997, no more publications were found. After that, the total number has been gradually increasing until 2005. From

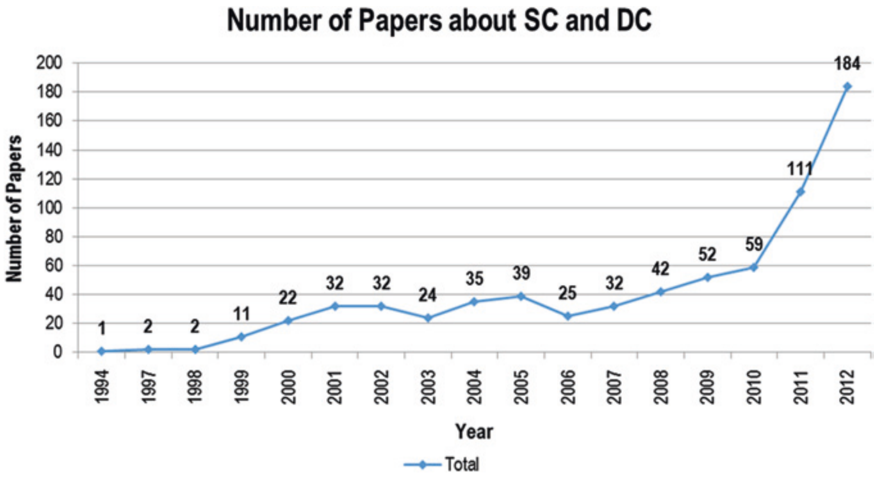


Fig. 2 Time analysis: number of papers about smart city and digital city

2006 to 2009, the trend line shows a steady increase (plus 10 units per year), while from 2010 its growth was doubled year by year up to 184 units at the end of 2012. Therefore, the interest about Smart City and Digital City is quite stable from 1993 to 2010 and it increases exponentially from 2010 to now.

Examining time analysis results, five dates have been identified as possible causes which could have influenced the development of Smart City and Digital City concept. These dates are: 1997, 2000, 2005, 2008, 2010.

1. *1997*. This year was characterized by *Kyoto Protocol*. Its main purpose is to limit CO_2 emissions and consequently to safeguard the environment all over the world. The Kyoto Protocol was signed by 192 Parties, including European Union and 191 States (such as, all United Nations members with the exception of the United States, Andorra, Canada, South Sudan). Nevertheless, it was entered in force in 2005 after Russia ratified it in October 2004. All parties were required to prepare policies and measures to decrease CO_2 emissions in their respective countries. There are two commitments periods in which developed countries have to achieve binding limitations or reductions emissions of greenhouse gases: the time frame between 2008 and 2012 and the time frame between 2013 and 2020 [16]. The Kyoto Protocol has certainly influenced the way to think the city, especially modern and industrialized cities characterized by a strong urbanization. In this context, during the latest twenty years, all State Parties have begun to foster several initiatives about CO_2 emissions reduction to apply them within their own boundaries with consequent studies increase about these scenarios. This role of Kyoto Protocol in driving countries and cities to design and apply environmental policies is also one of the main drivers of interest about the Smart City topic; less relationships are between Kyoto Protocol and Digital City.
2. *2000*. Two thousands were characterized by widespread of *Internet* all over the world, not only in business or academic context but especially in everyday life

[17]. In these years the *ICT* infrastructure, such as broadband infrastructures, wireless sensor networks, Internet-based networked applications, open platforms, were spread more and more, in order "to work together to form the backbone of a large intelligent infrastructure" [18]. Thanks to the use of Internet-based infrastructures, the e-services supply regarding healthcare, energy, education, environmental management, transportation, mobility and public safety, has begun to spread among citizens. At the same time, mobile phones have become more accessible for everybody (not only for businessmen but also for each citizen) evolving in technologically sophisticated products able to use the Internet access-point and to supply intelligent services to the users. The accessibility to the Internet in urban life has become easier and more popular [17, 19]. The newness it is that the city increases its cooperation with the surrounding territory in physic and virtual terms, in order "to build an arena where people in communities can share knowledge, experience and mutual interests" [20]. The author observes that this scenario supports more and more the concept of Digital City as a wired-city based on Internet, in which it is possible to provide public and private services to create socio-economic value for customers, citizens and the civil society [21]. Therefore, the Internet diffusion is one of the main driver of interest regarding the Digital City concept; less relationships are between the Internet and Smart City;

3. *2005*. During this year the *Kyoto Protocol entered in force* on the 16th February. After this moment, the international initiatives about the safeguard of environment have spread to achieve the Kyoto Protocol aims. Therefore, this scenario has fostered the development of smart strategies all over the world, focused on the environment safeguard.
4. *2008*. In 2008 two important events could have been influenced smart/digital researches: *the IBM Smart Planet concept* and *the Covenant of Mayors*.

   – IBM is the first company paying attention to the concept of "*Smart Planet*". For IBM, Smart Planet is as an instrumented, interconnected and intelligent planet in which leaders in business, government and civil society around the world could use Big Data to "transform enterprises and institutions through analytics, mobile technology, social business and the cloud" [6]. For IBM, this is the way to compete in the "smart" era, to have a good quality of life and to improve the city. Therefore, IBM has started a new business in this sector supplying to governments smart solutions focused on communications, energy and utilities, healthcare, insurance, retail, transportation, and so on. After that, many companies worldwide (such as Cisco, ABB, HP, Siemens, Ericsson, etc.) followed the IBM idea studying new smart projects for urban city issues. Therefore, putting together the event of the entered in force of Kyoto Protocol and the diffusion of Smart Planet concept, the author observes that the adjective "smart" gather with the word "city" has begun to widespread in every research field. However, the Smart City referred by IBM is not only smart, but also digital, because the role of ICT in pursuing the Smart City goals is crucial This is one of the most important example of confusion between smart and digital.
   – *Covenant of Mayors* is a self-started initiatives of European Cities. This initiative is finalized to spread the Smart City concept and to reduce CO_2

emissions by more than 20 % by 2020 through increased energy efficiency and fostering renewable energies. The agreement is fostered by European Commission in the frame of fulfilling the objectives of Strategy 2020 [22]. Its actions mainly focuses on: clean mobility, private and public buildings redevelopment, citizen awareness on the energy consumption theme. The signatory cities agreed to issue their own Action Plan for Sustainable Energy (PAES), which consists in a roadmap for fulfilling the agreement objectives. The Covenant of Mayors initiative is not isolated, it interacts with a number of projects, policies and initiatives. The main synergy in place is with the Strategic Plan for Energy Technologies (SET Plan). Indeed, SET Plan outlines the logic frame where the actions to fulfill Agenda 2020 objectives have to be developed [23]. In this context, the Smart City concept develops more than Digital City and it mainly regards the sustainability in terms of pollution reduction and environmental quality improvement.

5. *2010*. In 2010 the EU launched the *Europe 2020 Strategy* about delivering [22]: smart growth investing in education, research and innovation areas; sustainable growth investing in technologies and resources low-carbon economy; inclusive growth giving a strong emphasis on job creation and poverty reduction. The Europe 2020 strategy is focused on five goals in different areas, which should be achieved by European Union within the end of 2020 year. These goals concern: employment, research and developments, climate change and energy sustainability, education, poverty and social inclusion. Aiming to achieve these objectives, each European country commits to carry out smart or digital initiatives in its own major cities. In this way, the Europe 2020 Strategy has increased the widespread of both Smart City and Digital City concepts and, of consequence, the research studies about them.

Linking the papers trend with the most important events, the author identifies the links between technological and political situations and smart/digital researches. It emerges from the literature analysis that the time increasing of papers has been influenced by the Internet development, that justifies investments in Digital City initiatives, and environmental global policies, such as Kyoto protocol and EU 2020 Strategy, explicitly focused on sustainable growth and CO_2 emissions reduction, more influencing the Smart City investments. Therefore, both Smart City and Digital City strategies, and consequently also researches about these topics, are the effect of technology advancements and environmental sensibility. These two causes explain the most of papers published about Smart and Digital City and the exponential increase of papers after 2009.

3.2 Terminology Analysis: Smart Versus Digital

The terminology analysis aims to separate papers regarding Smart City from papers regarding Digital City and to order them chronologically, to distinguish the different time trends characterizing these two research topics. To accomplish with

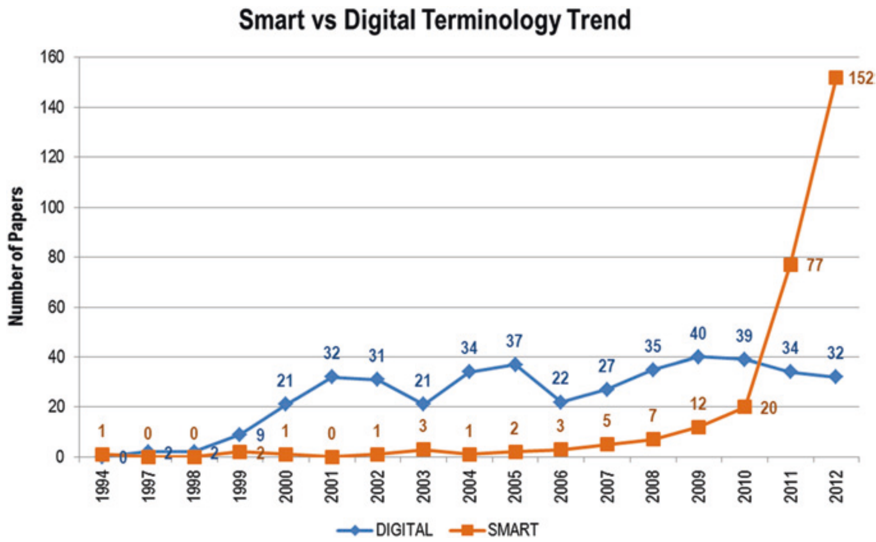


Fig. 3 Terminology analysis: smart versus digital terminology trend

these objectives, the 705 papers stored in the LRS-DB were organized and counted by chronological order of publication and by the label “smart” or “digital” according to the title of paper. The papers terminology trend is represented in Fig. 3.

Figure 3 shows the time distribution of Smart City papers and Digital City papers year by year. The graph underlines that these two topics have a very different time trend and it clearly appears from the figure that the Digital City concept was born before the Smart City one. Indeed, Digital City was conceived and developed in the nineties, in the context of Internet adoption in everyday urban life [20]. Smart City was born in 1994, but papers regarding this topic are few or zero until 2010, when the European Union started to use “smart” to qualify sustainability projects and actions in the urban space [24].

About papers labeled “Digital City”, they have been rapidly increasing from 1997 until 2009. In this time frame, “Digital City” is always more used respect to “Smart City”. Instead, after 2009, papers labeled Smart City have begun to exceed respect to Digital City papers. In this context, the author identifies two main events influencing the high interest in Digital City topic, in year 1994 and 2000.

1. *1994*, in this year Digital City Amsterdam was born. Amsterdam is the first Digital City in the Netherlands and in Europe. Nowadays, “it is usually taken as example of a successful project in that field” [25]. A major part of its success depends on the use of a virtual *metaphor of city*, because “the use of appropriate navigation metaphors can help to make the structure of modern information systems easier to understand and therefore easier to use” [26]. From this success, other cities tried to repeat the Amsterdam experiment contributing to spread the metaphor of “Digital City”.

2. *2000*, this year was characterized by a large widespread of the *Internet* in everyday life and by *ICT* diffusion among citizens [17, 18]. In this context, the author believes that the newness of topic, the accessibility and affordability of the Internet and digital devices fostered the increase of papers labeled "Digital City". In these years, there were several studies about Digital City definition: some researchers affirmed that a Digital City is a wired-city [3]; while others affirmed that it is a virtual reconstruction of city [13]; but the most famous Digital City definition is by Ishida in his study about Kyoto Digital City, in which he defines Digital City as "an arena in which people can interact and share knowledge, experiences, and mutual interests" [20]. Ishida's studies have been very important for the development of Digital City topic, because they have paid attention to how a city could summarize different aspects (such as data and information, e-services, etc.) of urban everyday life on the Internet in order to facilitate people in their decisions-making process. Moreover, the Digital City idea is strictly linked with the use of ICT in public administrations and with the e-Government practices, regarding both central governments and also local governments such as municipalities and city councils and administrations. Adopting an e-Government policy a city starts a transformation path towards a digital city.

About papers labeled "Smart City", they had a flat growth until 2010. The author identifies three dates which characterized the most important increase of Smart City papers: 2005, 2007, 2010.

1. *2005*, after 2005 Smart City papers increased little by little year by year. The author identifies the reason of this growth starting point in the *entered in force of Kyoto Protocol*, which is already explained in the previous paragraph.
2. *2007*, Apple Ltd. launched the i-phone, the first *smart phone* and, from this moment, the use of smart devices has been spreading more and more in everyday life. The adjective "smart" identifies devices that combine telephony and computing: smart phone have high-speed data access by Wi-Fi and mobile broadband in order to supply in real-time digital services to their users and, at the same time, to improve their quality of life [27]. So, the success of word "smart" in mobile context could have influenced the adjective "smart" in urban city context; in this case, the smart label identifies a digital device and more generally the building of a digital urban arena, rather than the definition of smart strategies, and it contributes to overlap smart and digital meanings of innovative urban policies;
3. *2010*, papers labeled "Smart City" shows a huge hike till 2012. This strong excess of Smart City papers respect to Digital City ones could be caused by the *Europe 2020 Strategy* approved by European Commission (this issue was analyzed in the previous paragraph). Indeed, Europe 2020 Strategy have widespread the Smart City label in terms of urban space sustainability to the detriment of Digital City label, because the Europe 2020 Strategy focuses its attention on environment safeguard, sustainability and social issues [22].

Therefore, this terminology analysis regarding Smart City papers and Digital City ones highlights the evolution of these two topics:

- both Smart City and Digital City research fields start to develop in the Nineties, however Digital City has been gaining a steady interest for twenty years, while Smart City has a very low number of papers till 2009 and an exponential increasing from 2010 till now;
- the time trends of researches about both these topics are strongly influenced by external drivers, such as technological drivers like the Internet diffusion or the smart devices use, or political drivers such as the Kyoto protocol and the EU 2020 strategy;
- the different time trends regarding these two topics and the different drivers they are influenced by, shows to us that even if Smart City and Digital City are often used like synonymous, they are quite different. Their main differences regards:
  - their contents: the Digital City regards the use of ICT in urban areas, the Smart City regards the attention to be paid to the environmental quality in cities;
  - their nature and relationship with the government: Digital City is a free trend emerging from the daily use of smart and digital devices by citizens, and it incites the local governments to supply e-services, that is, to gradually transform the city into a Digital City; Smart City is a political trend, driven by international institutions, to implement adequate initiatives to improve the environmental quality in cities.

3.3 Definition Analysis

The purpose of the definition analysis is to compare the most cited definitions of Smart City and Digital City, to understand which are the main similarities and the differences between these two concepts, often overlapped or confused.

To accomplish with this goal, the 705 papers stored in the LRS-DB have been organized by the citation number and by the paper focus. Afterwards, the author analyzes and compares the most recurrent and validated definitions of Smart City and Digital City respectively in Tables 3 and 4. Each table discloses the definition and the reference, putting in evidence some words to extract the meaning of these concepts: the bold character is used to outline the human component of Smart/ Digital City; while the italic character is used to outline the applied technologies.

The comparison of these definitions helps us to create a sound relationship between these two topics and to understand if and which are the links between these two different urban strategies.

If we consider the human aspect, both the topics refer to people or citizens. 7 out of 9 Smart City definitions regards citizens or people; several definitions explicitly refer to their quality of life in city. Some definitions recall the role of public

Table 3 Most cited definitions of smart city

Definition	Reference
"A Smart City is a city well performing built on the 'smart' combination of endowments and activities of self-decisive, independent and aware citizens"	Giffinger [37]
"A smart community is a community that has made a conscious effort to use *information technology* to transform life and work within its region in significant and fundamental rather than incremental ways"	California Institute [42]
"A city to be smart when investments in human and social capital and traditional (transport) and modern (*ICT*) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory governance"	Caragliu et al. [1]
"Smart city is defined by IBM as the use of *information and communication technology* to sense, analyze and integrate the key information of core systems in running cities"	IBM [6]
"Smart City is the product of *Digital City* combined with the *Internet of Things*"	Su et al. [8]
"Concept of a Smart City where citizens, objects, utilities, etc., connect in a seamless manner using *ubiquitous technologies*, so as to significantly enhance the living experience in 21st century urban environments"	Northstream [47]
"A city that monitors and integrates conditions of all of its critical infrastructures, including roads, bridges, tunnels, rails, subways, airports, seaports, *communications*, water, power, even major buildings, can better optimize its resources, plan its preventive maintenance activities, and monitor security aspects while maximizing services to its citizens"	Hall [36]
"Smart City is a city in which it can combine *technologies* as diverse as water recycling, advanced energy grids and mobile communications in order to reduce environmental impact and to offer its citizens better lives"	Setis-Eu [48]
"A smart city is a well-defined geographical area, in which high technologies such as *ICT*, logistic, energy production, and so on, cooperate to create benefits for citizens in terms of well-being, inclusion and participation, environmental quality, intelligent development; it is governed by a well-defined pool of subjects, able to state the rules and policy for the city government and development"	Dameri [7]

and private services in improving the quality of life; only 4 out of 9 definitions explicitly recall the importance of the environmental impact of urban activities, the concept of better natural resources management or sustainable economic growth. Therefore, in the academic milieu, the idea of Smart City is not so focused on environmental goals, like it happens applying the EU definition; however, in empirical implementations, cities are ever more applying the EU definition, driven by the aim to obtain public funding to realize their own projects for better cities. Instead, in the academic definitions the quality of life is linked with the quality of people and community, depending on the cultural level, the data information and knowledge sharing, but also some other aspects of community life, for example awareness or consciousness, human capital, communication between people and so on.

Table 4 Most cited definitions of digital city

Definition	Reference
"A digital city is substantively an open, complex and adaptive system based on computer network and urban information resources, which forms a *virtual digital space* for a city. It creates an information service marketplace and information resource deployment center"	Qi et al. [49]
"A Digital City has at least two plausible meanings: (1) a city that is being transformed or re-oriented through *digital technology* and (2) a *digital representation* or reflection of some aspects of an actual or imagined city"	Schuler [11]
"The concept of Digital City is to build an arena in which people in regional communities can interact and share knowledge, experiences, and mutual interests. Digital City integrates urban information (both achievable and real time) and create public spaces in the *Internet* for people living/visiting the city"	Ishida [20]
"Digital city denotes an area that combines broadband communication infrastructure with flexible, service-oriented computing systems. These new *digital infrastructures* seek to ensure better services for citizens, consumers and business in a specific area"	Komninos [50]
"The term Digital City (a.k.a., digital community, information city and e-city) refers to: a connected community that combines broadband *communications infrastructure*; a flexible, service-oriented *computing infrastructure* based on open industry standards; and, *innovative services* to meet the needs of governments and their employees, citizens and businesses. The goal of a Digital City is to create an environment for information sharing, collaboration, interoperability & seamless experience for all its inhabitants anywhere in the city"	Yovanof et al. [51]
"Digital City does not refer to a specific urban entity or formal communications mechanism, but it refers to a functional approach which describes four interdependent action types: Digital City supports data and information related to a city in digital format; Digital City supports a *communication infrastructure* (physical or virtual means for enabling information flows); Digital City delivers value added information and innovative services (these are likely to synthesize data from a range of sources, be location based and may include analytical interfaces); Digital City uses *virtual environments* in planning, decision-making and analysis (when data collected by citizens are used in the process of modeling or digitally recorded citizen behavior is influenced by formal planning an analysis a feedback loop is completed)"	Schiewe et al. [13], Dykes [52]

In Digital City definitions, people or citizens are cited, but their role is less proactive. Also the idea of improving the citizens' quality of life is not explicitly enounced in Digital City definitions. Instead, it appears crucial the virtualization process, that is, the transformation of a material city into a virtual city, able to create a new intangible urban dimension where people, relationships and services are virtually joined and shared to build a smarter community.

This different approach probably derives from the different role of the ICT in these two city ideas. The Digital City is obviously based on ICT: the ICT is the core component of a Digital City and all the other aspects—citizens, services, communities, relationships, communications, information and knowledge, human and social capital—are joined through the technology. Also the Smart City has the ICT like an important element: 7 out of 9 definitions explicitly or not recall the ICT, or the Internet, or similar concepts. Reading the Smart City definitions, it emerges that the ICT is ever an important element characterizing the Smart City, but not the only one, instead together with other aspects.

Examining the time distribution of the Smart City and Digital City definitions, it emerges also that the Smart City definitions are more recent respect to the Digital City ones: 6 out of 9 are after 2010, whereas only 3 out of 6 Digital City definitions are after 2010. We can argue that the Smart City concept somewhat includes the Digital City idea, that is, the present concept of Smart City actually is a merge of both the environmental requirements of a smart city with the digital requirements and attitudes.

This latest evidence contradicts the idea of two different contents in Smart City and Digital City; indeed, the Smart City has born from three different sources: the EU source, focusing on the environmental requirements; the digital source, based on the previous experiences of Digital Cities; and the cultural source, that is, the human and social capital able to build the smart community.

For these reasons, the Smart City definition analysis discloses a wide range of meanings associated with a smart city, including environmental, social and digital components.

3.4 Typology Analysis: Theoretical Versus Empirical

The purpose of the typology analysis is to separate theoretical papers respect to empirical papers and to count them during the time frame regarding the latest twenty years (1993–2012). The reason of this type of examination is to understand if Smart City and Digital City are mainly academic ideas, built on theoretical basis, or if they emerges from concrete implementations in cities, and which is the relationship between theoretical studies and empirical implementations. This analysis, moreover, aims also to investigate if Smart City and Digital City are top-down phenomena or bottom-up ones. For example, analyzing the Amsterdam Digital City case, one of the most successful in Europe and recognized like a pilot case, it emerges that it has been a bottom-up phenomenon, grown from the free use of the Internet by citizens to share their opinion before the local elections. On the contrary, the further experience of Amsterdam Smart City appears like a top-down project, where the Municipality of Amsterdam assumes the leading role to implement several smart initiatives in the urban area. To investigate about the distribution in time of both theoretical and empirical papers helps us to better understand the Smart/Digital City phenomenon and its origin.

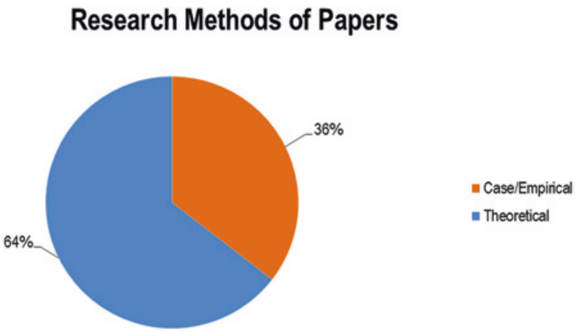


Fig. 4 Typology analysis: research methods of papers

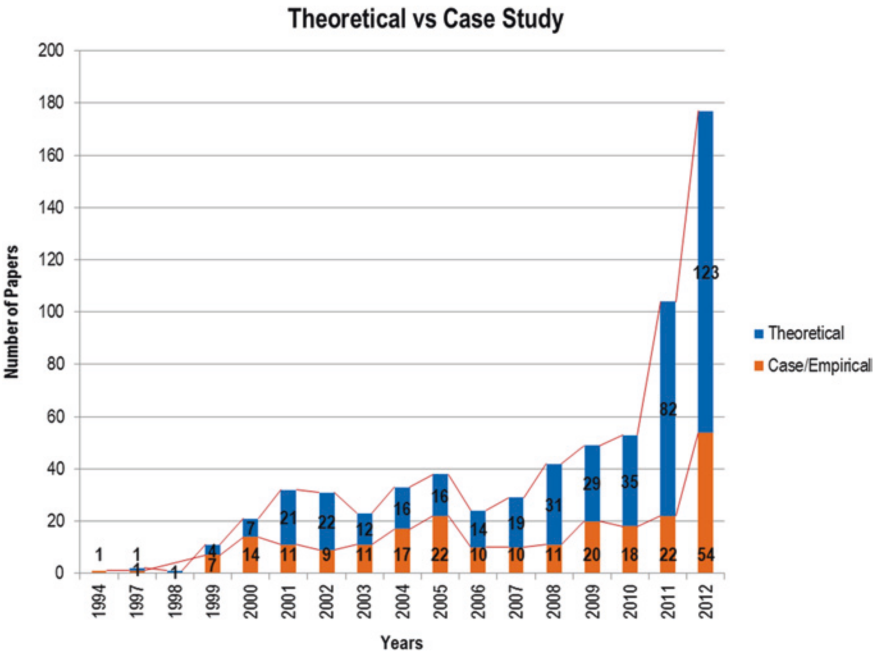


Fig. 5 Typology analysis: theoretical versus case study papers

To accomplish with these objectives, the 705 papers stored in the LRS-DB have been organized and counted by chronological order of publication and by the label "theoretical study" or "empirical study" according to the research method highlighted in the abstract. The share of each research method used in stored papers is shown in Fig. 4; while the theoretical/empirical study time trend is represented in Fig. 5.

As the pie chart shows:

1. *theoretical studies* are 64 % of the reviewed papers. As defined by Wacker [28] in his study about theory research, this category includes papers regarding the following research methods: "conceptual definitions, domain limitations, relationship-building, and predictions". This kind of study is therefore the most adopted among researchers to conduct the study about Smart City and Digital City;
2. *case studies* (also known as empirical studies) are 36 % of the reviewed papers. This type of papers includes researches where "the methodology must use data from external organizations or businesses to test if relationships hold in the external world" [28]. This kind of study includes all papers which investigate on the Smart/Digital City phenomena within its real-life context, in order to verify the empirical applications of theoretical concepts [29]. In this literature review, the object of empirical studies are especially case studies regarding cities implementations and projects defined as smart or digital.

Figure 5 shows the use of different research methods during the time frame 1993–2012: in the first decade, empirical studies are almost always more than theoretical studies. Depending on this evidence, it is possible to argue that the first steps regarding Smart City and Digital City are empirical and therefore the diffusion process probably has been mainly bottom-up; cities or other agents started to implement smart or digital initiatives, without a comprehensive strategy or a leader driving the implementation of a common and shared strategy [30].

Analyzing only the empirical papers, we can note that the smart or digital label is chosen depending on some smart or digital characteristics of one or a group of empirical implementations or projects, without referring to a smart or digital strategic vision regarding the whole city. If we consider a comprehensive smart or a digital vision of the city applied to the empirical papers, we can observe three different contexts:

1. Smart City, when the city follows sustainable strategies through the innovative and sustainable use of its own natural resources;
2. Digital City, when the city follows digital policies aiming to supply e-services to the citizens through the use of technologies such as Web 2.0, Cloud Computing, Internet of Things, and so on [31];
3. Smart City based on Digital City, when the city follows sustainable strategies using technologies applied in Digital City. In this case, the Digital City represents the ICT component on which the Smart City strategy is based.

From the empirical case studies survey it emerges also the central role of technological, innovative solutions for the Smart/Digital City implementations. We can say that the Smart or Digital City development is largely based on the application of innovative technologies to urban projects. It can also explain why at the beginning a shared definition of Smart City has not been conceived: the smart or digital projects have been influenced from the technological innovation and its application to urban areas and themes. It means that the idea of a Smart or a Digital

City has been mainly technology driven, instead of policy driven. However, after several different technological applications have been implemented in cities, and each of them has been qualified as smart, to express a unique, universal Smart City definition has become very difficult. The origin of smart implementations explains therefore why a shared definition of Smart City still lacks.

3.5 Geographic Analysis

The aims of geographic analysis is to discover where all over the world Smart City and Digital City strategies and projects have been more implemented, evidencing geographical areas more interested in smart themes and cities that are considered pilot cases worldwide.

To accomplish with these objectives, each of the 705 papers stored in the LRS-DB has been labeled by the city it refers to, according to the abstract of paper, and each city has been assigned to a continent. In this way, the subset of 705 papers has been analyzed only considering the empirical studies of Smart/Digital City, for a total of 162 case studies. Table 5 shows the relationship between cities and continents. Then the papers have been organized and counted by "empirical study" label and by the "continent". The geographical distribution of Smart/Digital Cities all over the world is shown in Figs. 6 and 7: the first graphic shows a pie chart counting the number of Smart/Digital Cities for each continent; the second one shows a bubble chart about the Smart/Digital Cities geo-location in the world.

Moreover, to better understand Fig. 7, we can see Table 5 depicting the list of Smart/Digital Cities per Continent and per Country.

The geographic analysis of Fig. 6 highlights that:

1. *Asia* is the continent where there is the highest number of Smart/Digital Cities with 49 % of reviewed papers;
2. *Europe,* after Asia, is the continent where there is a relevant number of Smart/Digital Cities with 36 % of reviewed papers;
3. *North America*, is the third continent in terms of Smart/Digital Cities with 9 % of reviewed papers;
4. *Oceania*, *Africa* and *Middle/South America,* are the continents in which there are the lowest number of Smart/Digital Cities, with respectively 3, 2 and 1 % of reviewed papers.

Figure 7 shows the Smart/Digital Cities position in the world, in order to detect if and where they form some clusters. From the exam of the above bubble chart, different aspects emerge:

- there are macro-clusters of Smart/Digital Cities both in *Asia* and in *Europe*. If we compare this two clusters depending on their extension, we can say that Asia presents a greater cities dispersion than Europe; while Europe highlights a greater cities concentration. In Asia we observe the greatest Smart or Digital

Table 5 List of smart/digital cities geo-location in the world on the basis of 162 case studies analyzed

Continent	Cities
Asia	Amman (Giordania), Bangladesh Region (Bangladesh), Beijing (China), Caofeidian island area (China), Changzhu (China), Chengdu (China), Cheongna (Korea), Dongying (China), Dubai (Emirates), Guangdong (China), Guangzhou (China), Guiyang (China), Hangzhou (China), Harbin (China), Heilongjiang (China), Hong Kong (China), Huizhou China), Jiangan (China), Kochi (India), Kuwait Region (Kuwait), Kyoto (Japan), Lianyuangang (China), Lijiang City (China), Linyi (China), Macao (China), Masdar City (Emirates), Panzhihua (China), Pudong (China), Qianjiang City (China), Seoul (Korea), Shanghai (China), Shenzhen (China), Singapore (China), Subang Jaya (Malaysia), Taipei (Taiwan), Tang Chang'An City (China), Tokyio (Japan), Wuhan (China), Xiamen City (China), Zhengzhou (China)
Europa	Amsterdam (Netherlands), Anpwerp (Belgium), Aveiro (Portugal), Barcelona (Spain), Beaufort (France), Berlin (Germany), Bilbao (Spain), Bologna (Italy), Bolzano (Italy), Bragança (Portugal), Bristol (UK), Como (Italy), Copenhagen (Denmark), Dubiln (Ireland), Fredrikstad (Norway), Genova (Italy), Ghent (Belgium), Helsinki (Finland), London (UK), Luxembourg (Luxembourg), Malta (UK), Manchester (UK), Marseilles (France), Milano (Italy), Nottingham (UK), Oulu (Finland), Parthenay (France), Trikala (Greece), Turin (Italy), Venice (Italy), Vienna (Austria), Vilnius (Lithuania)
North America	Charlotte (USA), Cleveland (USA), Edmonton (Canada), Iowa (USA), Montreal (Canada), New York (USA), Portland (USA), Quebec City (Canada), Saskatoon (Canada), Seattle (USA), Sudbury (Canada)
Oceania	Adelaide (Australia), Brisbane (Australia), Melbourne (Australia), Parramatta (Australia)
Middle/South America	Curitiba, Paranà (Brazil), Juarez (Mexico), Mexico City (Mexico), Nassau (Bahamas), Rio de Jainero (Brazil)

Cities concentration only in correspondence to the Chinese east coast; while European Smart/Digital Cities appear to be more concentrated in the North Sea Region (that is the Netherlands, Belgium, United Kingdom, Scandinavia) and in the Mediterranean Region (that is Spain, France, Italy);

- there is a little Smart/Digital Cities cluster also near the Great Lakes Region between United States and Canada (*North America*);
- in *Oceania* and in *Africa*, there are the smallest clusters observed and in both cases they are located along the most populated and developed areas: Australian east coast and South African coast.

The same results have been reached by the Ericsson Report about Networked Society City Index. This report shows that "cities located in Northern Europe, North America and parts of East Asia have a longer tradition of producing and using ICT equipment, and have therefore been able to benefit from their

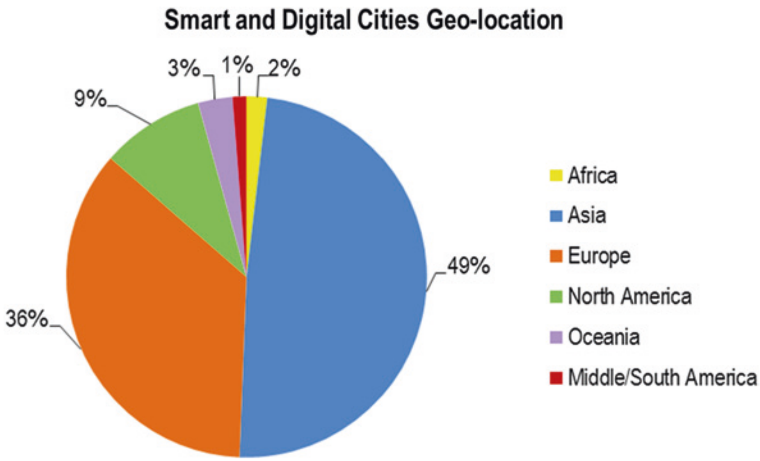


Fig. 6 Geography analysis: smart/digital cities geo-location (per continent) on the basis of 162 case studies analyzed

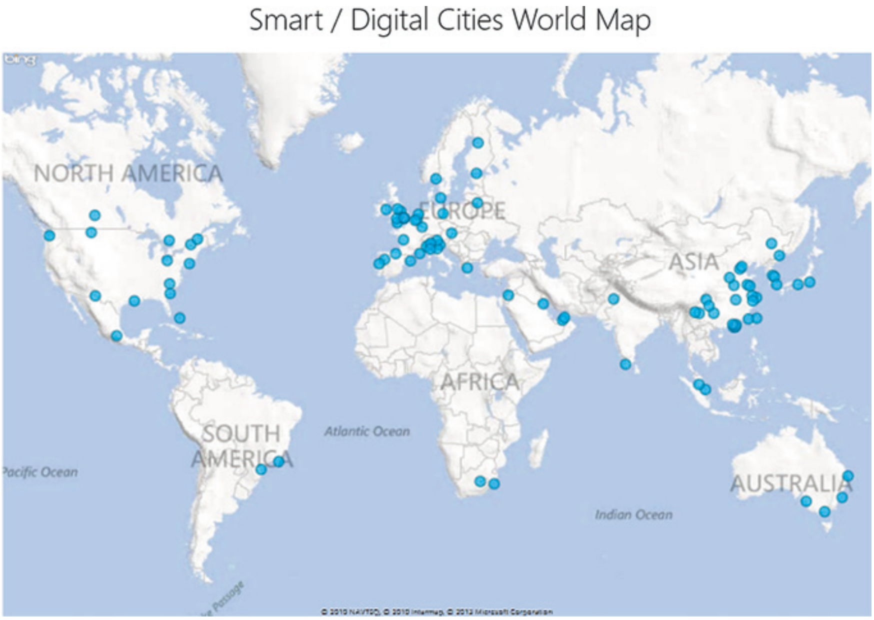


Fig. 7 Geography analysis: smart/digital cities geo-location in the world on the basis of 162 case studies analyzed

investments over longer periods of time" [32]. Indeed, from the literature review about city case studies, we can observe that the spread of Smart/Digital Cities in Asia, Europe and North America have some shared features:

1. the widespread and development of ICT infrastructures, considered like:
   - Internet diffusion among citizens in everyday life [17];
   - Internet more accessible and affordable for many people to reduce digital divide [33];
   - data sharing and open data;
   - increase the adoption of Community Network to supply e-government services (this aspect is more relevant especially among Smart/Digital Cities in Europe);
   - focuses on the use of ICTs for public administration;
   - to provide better public services and e-services also using Web 2.0 technology;
   - increase urban wealth [1];
   - increase innovation and entrepreneurship [32];
   - increase social cohesion;
2. the adoption of green policies for a smart growth, in order to:
   - reduce issues about urban crowding in terms of pollution reduction, improvement of urban planning, safety and sanitary conditions, power demand sustainable, and so on (these aspects is more relevant especially among Chinese Smart/Digital Cities);
   - reduce CO_2 emissions and greenhouse gases;
   - improve mobility services to reduce traffic congestion and then pollution;
   - achieve sustainable urban development and a better urban landscape [1].

Instead, Smart/Digital Cities in Middle/South America and Africa have in common the widespread and development of ICT infrastructures, but for other reasons in respect to Asia, Europe and North America. For example:

- to attract foreign investment promoting local advantage and to improve cultural, economic and social development [34];
- to enable service delivery and economic development;
- to enable the transition to a knowledge economy;
- to focus on ICT access in rural and periphery urban areas [35].

4 Conclusions

The large literature survey described in this work aims to clarify several aspects regarding the new, still immature strategies of Smart City and Digital City. Several goals have been reached thanks to this deep survey. We can summarize them respect to three large themes:

- Smart/Digital City definition;
- birth of Smart/Digital City ideas;
- diffusion of Smart/Digital City implementations.

Regarding the definition of both Smart City and Digital City, we can observe that a shared and acknowledged definition of both Smart City and Digital City still lacks. However, there are several most cited definitions and they are establishing themselves like standards (see for example Hall [36], Caragliu [1], Giffinger [37]). Digital City definitions show a higher uniformity, because all of them are focused on the key role of ICT in improving the quality of services and information supplied to citizens. Smart City definitions are more different each others, mainly because the purpose of a Smart City is often too large, that is, to improve the quality of urban life; depending on this goal, everything could be considered smart! However, deepening our analysis, we discover some shared features characterizing Smart Cities, that is, the role of innovation and technology, the environmental requirements, the economic and social development. Sometimes, especially during the latest years, also the use of ICT has been included into the Smart City perimeter; it means that the Digital City is becoming a subset of the Smart City.

Regarding the birth of Smart City and Digital City, both of them date back to twenty years ago environ. However, the Digital City has a development synchronized with the Internet diffusion, especially in everyday life and in e-government. This development has been quite stable during the latest twenty years, with some peaks around 2000. Smart City, on the contrary, had a very slow development till 2010, when the UE assumed the Smart City like one of its key development paths. From this year, papers and researches about this topic have a strong outburst.

It is interesting to note that the birth of both Smart City and Digital City has been mainly empirical, and only after sometime a theoretical research activity about these topics started to increase. Moreover, this empirical birth is also bottom-up, that is, it derives from the independent, free application of ICT or other innovative technologies to smart and digital aims, to improve the quality of life in cities. For this reason, Smart City and Digital City are often the result of a sum or collection of single initiatives, instead of the outcome of a well conceived strategy. Only the latest Smart City implementations show a new trend, towards a top-down path, where municipalities are assuming a leading role in defining and driving a comprehensive vision about the Smart or Digital City programs. Both Smart City and Digital City empirical implementations are strongly driven by the technology. ICT or engineering technologies are the real engine of the Smart/Digital projects, even if different are their application fields: information sharing, communication and citizens involvement for Digital City projects, environmental safeguard, pollution reduction and infrastructure quality for Smart City projects.

Also the diffusion of SmartCity and Digital City is largely driver from the technological progress. Indeed, the presence of Smart Cities or Digital Cities among the continents is higher where higher is the economic and scientific development of a country. Obviously, a strong driver for a Smart/Digital City implementation is the city dimension: indeed, the larger is the city, the worse is its environmental impact, to be reduced thanks to Smart City programs; the larger is the city, the better are the benefits deriving from data and knowledge sharing and e-services supply, to be taken thanks to Digital City programs [38].

Finally, we can say that the Smart City and Digital City phenomena are strongly spreading both in theoretical researches and in empirical implementations. Sometimes it is the result of a support from national or international governments, institutions or political bodies, such as the EU, that also finances Smart City projects in Europe; but more frequently it is the result of a new, innovative idea about city and urban life: more pleasant, more inclusive, greener and cleaner. The Smart City is nowadays seen like a key strategy to improve the quality of life of billions of people living in cities all over the world.

References

1. Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82.
2. Pardo, T., & Taewoo, N. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. *Proceedings of the 12th Annual International Conference on Digital Government Research* (pp. 282–291). New York: ACM.
3. Hollands, R. G. (2008). Will the real smart city please stand up? *City: Analysis of Urban Trend, Culture, Theory, Policy, Action, 12*(3), 303–320.
4. vom Brocke, J., Simons, A., Niehaves, B., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: on the importance of rigour in documenting the literature search process. *ECIS 17th European Conference on Information Systems* (pp. 2–13).
5. Cooper, H. M. (1988). Organizing knowledge syntheses: a taxonomy of literature review. *Knowledge Society*, *1*, 104–126.
6. IBM (2010). *Smarter thinking for a smarter planet*.
7. Dameri, R. P. (2013). Searching for smart city definition: a comprehensive proposal. *International Journal of Computers & Technology*, *11*(5), 2544–2551(Council for Innovative Research).
8. Su, K., Li, J., & Fu, H. (2011). Smart city and the applications. *IEEE International Conference on Electronics, Communications and Control* (*ICECC*), pp. 1028–1031(IEEE Xplore).
9. Pardo, T. A., & Nam, T. (2011). Smart city as urban innovation: focusing on management, policy and context. *Proceeding of the 5th International Conference on theory and Practice of Electronic Governance* (pp. 185–194). New York: ACM.
10. Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., Pardo, T. A., & Scholl, H. J. (2012). Understanding smart cities: an integrative framework. *45th Hawaii International Conference on System Sciences* (pp. 2289–2297). IEEE Xplore.
11. Schuler, D. (2002). Digital cities and digital citizens. In: M. Tanabe, P. van den Besselaar, T. Ishida (Eds.), *Digital cities II: computational and sociological approaches*. LNCS, vol. 2362, (pp. 71–85). Berlin: Springer.
12. Schuurman, D., Baccarne, B., De Marez, L., & Mechant, P. (2012). Smart ideas for smart cities: investigating crowdsourcing for generating and selecting ideas for ICT innovation in a city context. *Journal of Theoretical and Applied Electronic Commerce Research, 7*(3), 49–62. (Universidad de Talca, Chile).
13. Schiewe, J., Krek, A., Peters, I., Sternberg, H., & Traub, K. P. (2008) HCU research group “Digital City”: developing and evaluating tools for urban research. In: Ehlers et al. (Eds.) *Digital earth summit on geoinformatics*.
14. Moutinho, J. L. (2008). Building the information society in Portugal: lessons from the digital cities programm 1998–2000. In van Geenhuizen (Eds.), *Value-added partnering and innovation in a changing world*.
15. Ricciardi, F. (2010). ICTs in an ageing society: an overview of emerging research streams. In A. D’Atri, M. De Marco, A. M. Braccini, F. Cabiddu (Eds.), *Management of the interconnected world*. ITAIS, vol. 1, (pp. 37–44). Berlin: Springer.

16. Cosgrove, S. (2009). The United Nations framework convention on climate change. *15th Conference of the Parties—The Copenhagen Protocol*. Background Paper, AMUNC.
17. Ishida, T. (2000). Understanding digital cities. In: T. Ishida, K. Isbister (Eds.), *Digital cities*. LNCS, vol. 1765, (pp. 7–17). Berlin: Springer.
18. Schaffers, H., Ratti, C., & Komninos N. (2012). Special issue on smart applications for smart cities—new approaches to innovation: guest editors' introduction. *Journal of Theoretical and Applied Electronic Commerce Research*, *7*(3). (Universidad de Talca, Chile).
19. Ishida, T., & Hiramotsu K. (2001). An augmented web space for digital cities. *Proceedings of Symposium on Applications and the Internet* (pp. 105–112).
20. Sorrentino, M., & Simonetta, M. (2013). Incentivising inter-municipal collaboration: the Lombard experience. *Journal of Management and Governance, 17*(4), 887–906.
21. Dameri, R. P. (2012). Defining an evaluation framework for digital cities implementation. *IEEE International Conference on Information Society* (i-Society), (pp. 466-470). IEEE Xplore.
22. Europe 2020 Strategy. from ec.europa.eu/regional_policy/what/europe2020/index_en.cfm.
23. ABB Report (2012). Smart cities in Italia: un'opportunità nello spirito del Rinascimento per una nuova qualità della vita. Technical Report, ABB.
24. Al-Hader, M., & Rodzi A. (2009). The smart city infrastructure development and monitoring. *Theoretical and Empirical Researchers in Urban Management*, *2*(11).
25. van Bastelaer, B., & Lobet-Maris C. (1998). Social learning regarding multimedia developments at a local level. The case of digital cities. In B. van Bastelaer, C. Lobet-Maris (Eds.), *SLIM research—DG XII—TSER program. Final integrated study—Social learning in the public sector*, University of Namur.
26. Dieberger, A., & Frank, A. U. (1998). A city metaphor to support navigation in complex information spaces. *Journal of Visual Languages & Computing, 9*(6), 597–622. (Elsevier).
27. Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic, N., & Meijers, E. (2007). *Smart cities. Ranking of European medium-sized cities*. Vienna: University of Technology.
28. Wacker, J. G. (1998). A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of Operations Management, 16*, 361–385. (Elsevier).
29. Yin, R. K. (2003). *Case study research. Design and methods*. California: SAGE Publications.
30. Casalino, N., Buonocuore, F., Rossignoli, C., & Ricciardi F. (2013). Transparency, openness and knowledge sharing for rebuilding and strengthening government institutions. *IASTED Multiconferences—Proceedings of the IASTED International Conference on Web-Based Education*, (pp. 866–871). WBE.
31. Anthopoulos, L., & Tougountzoglou T. E. (2012). A viability model for digital cities: economic and acceptability factors. In: C. Reddick, S. Aikins (Eds.), *Web 2.0 technologies and democratic governance*. PAIT, vol. 1, (pp. 79–96). Berlin: Springer.
32. Ericcson Report (2012). Networked society city index—Part III: triple bottom line benefits for city business. Technical report, Ericsson.
33. Partridge, H., Hall, H., McAllister, L., & Hallam G. (2005). Effecting social change in the "Smart City": the west end connect community project. *Conference on the Social Change in the 21st Century. Centre for Social Change Research*, Queensland University of Technology.
34. Anci Cittalia (2011). Smart Cities nel Mondo. Technical report, Fondazione Anci Ricerche.
35. Odendaal, N. (2007). Towards the digital city in South Africa: issues and constraints. *Journal of Urban Technology*, *13*(3), 29–48. (Routledge).
36. Hall, P. (2000). Creative cities and economic development. *Urban Studies, 37*(4), 633–649.
37. OECD (2011). M-Government. Mobile technologies for responsive government and connected societies. Technical report, OECD Publishing.
38. ZArdini, A., Mola, L., Vom Brocke, J., & Rossignoli, C. (2010). The role of ECM and its contribution in decision-making processes. *Journal of Decision Systems*, *19*(4), 389–406.
39. Anthopoulos, L., & Fitsilis, P. (2010). From digital to ubiquitous cities: defining a common architecture for urban development. *IEEE 6th International conference on Intelligent Environments*, (pp. 301–306). IEEE Xplore.

40. Komninos, N. (2006). The architecture of intelligent cities: integrating human, collective and artificial intelligence to enhance knowledge and innovation. *IEEE 2nd IET International Conference on Intelligent Environments* (pp. 13–20). IEEE Xplore.
41. Couclelis, H. (2004). The construction of the digital city. *Planning and Design, 31*(1), 5–19 (Environment and Planning).
42. California Institute (2001), http://smartcommunities.org/concept.php.
43. Ergazakis, M., Metaxiotis, M., & Psarras, J. (2004). Towards knowledge cities: conceptual analysis and success stories. *Journal of Knowledge Management*, *8*(5), 5–15 (Emerald Group Publishing Limited).
44. OECD Observed (1999). Learning cities: the new recipe in regional development.
45. Batagan, L. (2011). Smart cities and sustainability models. *Revista de Informatica Economica, 15*(3), 80–87.
46. OECD: Green Cities Programme (2010). http://www.oecd.org/gov/regional-policy/49318965.pdf.
47. Northstream (2010). White paper on revenue opportunities, from http://northstream.se/white-paper/archive.
48. Setis-Eu (2012). setis.ec.europa.eu/implementation/technology-roadmap/European-initiative-on-smart-cities.
49. Qi, L., & Shaofu, L. (2001). Research on digital city framework architecture. *IEEE International Conferences on Info-Tech and Info-Net*, vol. 1, (pp. 30–36). Proceedings ICII.
50. Komninos, N. (2008). *Intelligent cities and globalization of innovation networks*. Lomdon: Routledge.
51. Yovanof, G. S., & Hazapis G. N. (2009). An architectural framework and enabling wireless technologies for digital cities and intelligent urban environments. *Wireless Personal Communications*, *49*(3), 445–463. (Springer).
52. Dykes, J. (2010). GeoVisualization and the digital city. *Computers, Environment and Urban Systems*, *34*, 443–451. (Elsevier).

Comparing Smart and Digital City: Initiatives and Strategies in Amsterdam and Genoa. Are They Digital and/or Smart?

Renata Paola Dameri

Abstract The objective of this research is to investigate the relations between Smart city and Digital city concepts and strategies. The author examines the international literature about these topics, comparing smart city and digital city definitions, components and goals. This survey shows that a clear definition of both smart city and digital city still lacks and that these two topics are often overlapped or confused. The same thing happens in empirical implementation of smart and/or digital strategies in cities. The research methodology includes the study and comparison of two important empirical implementations of Smart/Digital strategies in Europe: Amsterdam and Genoa. The results show that smart city and digital city are not the same, even if they are strictly linked each other and sometimes merged in common initiatives. Moreover, this empirical research highlights the key role of players, programs and governance in realizing smart/digital cities really effective for a best quality of life in the urban space.

Keywords Smart city • Digital city • Digital agenda • Case study

1 Introduction

The concepts of smart city and digital city are in the mood, however they are not clearly defined till now and several aspects of these two concepts are overlapping each other [1–3].

At present, several cities all around the world define themselves like smart city, but this definition is far from to be well stated. Indeed, these cities use the word smart to name a wide strategy, including a large spectrum of heterogeneous

R. P. Dameri (✉)
Department of Economics, University of Genoa, Genoa, Italy
e-mail: dameri@economia.unige.it

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_3,

initiatives involving several different technologies. A comprehensive vision of the smart city strategy lacks, so as a roadmap to implement it or a set of performance indicators to evaluate the success or failure of smart initiatives.

In the meantime, cities are also committed to create an ICT infrastructure to support big data collection and processing, communications between citizens and institutions, digital private and public services, and so on. EU, several national governments and cities themselves have their own digital agenda to be implemented and are becoming therefore digital cities [4, 5].

This panorama is quite confused and it impacts on the quality and effectiveness of public and private programs, to reach measurable and useful results by smart and digital city initiatives. To better drive city strategies, choices, investments, a clear definition of smart city and digital city is necessary, able both to understand these important phenomena and to support strategic decisions [6, 7].

This chapter aims to reach a definition of both smart city and digital city by comparing two smart city case studies: Amsterdam and Genoa. Amsterdam is the first European city launching a smart program. Genoa is the city leader in winning funding at the latest EU call for smart initiatives proposal. What are they doing? Which are they goals? Who are the main actors and stakeholders involved in the smart and/or digital city programs? The empirical study of the projects portfolio in Amsterdam and in Genoa is the instrument to understand their strategy and the meaning of the words smart and digital. Contents, goals and actors are examined, compared and evaluated, towards a theoretical definition of smart city and digital city supported by the empirical evidence of two case studies leader in Europe.

2 Smart City and Digital City

2.1 Why?

During the latest years, population in cities has been growing faster and faster. At present, 53 % of world population lives in cities, that occupy environ 2 % of the global space on the earth [8]. By 2050, 70 % of population will live in cities. This phenomenon is continuously increasing, and it is spread all over the five continents, even if some countries in particular could have a dramatic augment in urban population during the next 20 years: China, South Korea, Mexico, Brazil and several African countries [9]. Today's urban population is 3.3 billion and by UN calculations it is expected to double by 2050. It means that two out of every three people will live in city in 2050. In Fig. 1, we can see the world situation of urban density and population. There are 24 cities over 10 million inhabitants; several countries—have more that 75 % of population in cities, and many others are predominantly urban too, with more that 50 % of people living in the metropolitan areas.

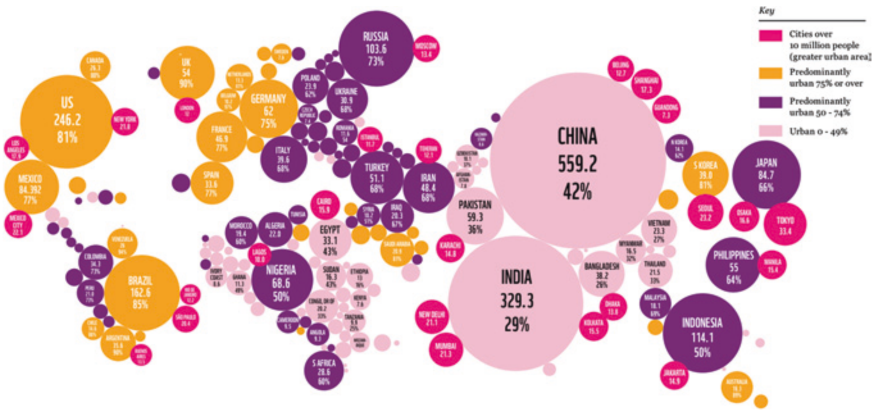


Fig. 1 The number of people living in cities in each country of the world in 2010, together with the percentage of the population in countries with large urban populations

This phenomenon regards not only large cities, but also the medium ones and all the countries in the world. Seeing as megacities such as New York, London, Beijing, Mumbai, and Mexico City can only grow so much, most of the urban growth will take place in smaller cities. The WWF predicts that the highest growth rate of 4.19 % will occur in cities with fewer than one million residents. Cities with more than one million residents will grow at rates less than 2 %.

Whenever in the history, the city has been the crucial space where economic and cultural development has happened, and nowadays the post-industrial development is more and more concentrated in the urban space [10]. However, the larger and larger dimension of cities all over the world poses the dramatic problem of their management. Pollution, overpopulation, scarcity of natural resources and food, difficulties in supplying public and private services are only some of the urgent challenges to face.

These two aspects—the good and the bad one—of the urban life are at present the most interesting drivers for the development of a smart city strategy [11]. Indeed, we could consider that if the metropolitan dimension of a city is a problem to solve, on the other hand it is also able to produce attractiveness and humus for a better development. A smart city strategy can aim in the meantime both to face the negative aspects—the threats—and to empower the positive ones—the opportunities of the greater dimensions of a city (see Table 1).

Opportunities in city are given by universities where to study, companies and public bodies and organizations where to work, theatres, cinemas, libraries, concert halls and all the public spaces where to catch cultural opportunities and spend his own free time in leisure and sport; the city is obviously a place where to live and a formidable milieu where to meet people. In city ideas born, circulate, create initiatives and business; in city things happen… How many initiatives could be exploited and supported thanks to smart and/or digital strategies?

Table 1 City threats and opportunities

Opportunities	Threats
Living	Traffic
Studying	Pollution
Working	Poverty
Cultural opportunities	Energy consumption
Leasure	Resources scarcity
Meeting people	Social tension
…	…

However, the presence of people, organizations, business in the urban space is a threat for the daily life: inefficient local public transports, traffic, the high cost of houses and loans are some of the more diffused negative impacts of the population high density in cities. Pollution is a main characteristic of several large and medium cities, with low rate of green spaces; the buildings use a high quantity of energy, the environmental impact is strongly negative and the social differences, especially in large cities, create poverty and social tensions [12]. Could a smart/digital strategy make something against these problems?

It is time to try to develop an answer to face the situation, before it becomes too difficult to face. Somehow, somewhere, the idea to use the technology to support a better way of life in cities, especially in large metropolitan areas, begins to emerge [13]. The use of high technology to shape the urban skyline is driven by the potential of the technology to enforce new strategies, initiatives, projects and infrastructures aiming at improving the quality of life in urban space along different axes: a smart development trend, able to create much economic value thanks to the use of better informed and linked people and business [14]; a sustainable development trend, using technology to implement low carbon economy, resource efficiency, sustainable transport [15]; and an inclusive development trend, using especially information and communication technology (ICT) to create social inclusion, civil participation at the political debate, higher education and information quality [16].

The smart city idea therefore follows a bottom-up path, growing from single initiatives of business, non profit organizations, public bodies, local governments, universities, … aiming at using the technology to struggle against the menaces in large cities like pollution, energy shortage, water and air bad quality, poverty and social exclusion and to create opportunities for sustainable growth, green cities, shared information, social communication and a higher quality of life in the urban space [17].

The smart city concept often overlaps with the digital city idea [18, 19]. These two urban strategies are not the same, but in the meantime they are not so different each other. Both of them use the technology—especially the ICT—to improve the life quality in city, to create economic development, to save the environment. But they are different both in their history and in their present implementation, in goals and aims to be reached, in strategies and projects to be implemented [20]. However, no one of key subjects—governments, businesses, universities and the citizens their own—is aware about the real differences between smart and digital; a clear and sound definition of smart and digital city also lacks in the academic debate. But a well-conceived definition is necessary to drive choices and to increase the probabilities of success in a so difficult context. Therefore, on the next paragraph the author will introduce some different aspects characterizing smart and digital city concepts. Further in the chapter, the differences in smart and digital urban strategies will be searched in two success case histories, Amsterdam and Genoa, two of the smarter cities in Europe.

Finally, both the theoretical and the empirical investigation will support the conclusions, lessons learned and further work aims.

2.2 What Smart City and Digital City Are?

The smart city idea was born from the application of hi-tech solutions to urban problems, but especially from the use of ICT in connecting people, political institutions and business. This use of ICT is also at the basis of the digital city idea. For this reason, these two concepts are quite confused. Moreover, each city implementing a smart or a digital strategy defines itself like smart or digital, using this word in relation with its own initiatives and projects, without referring to a shared and recognized standard.

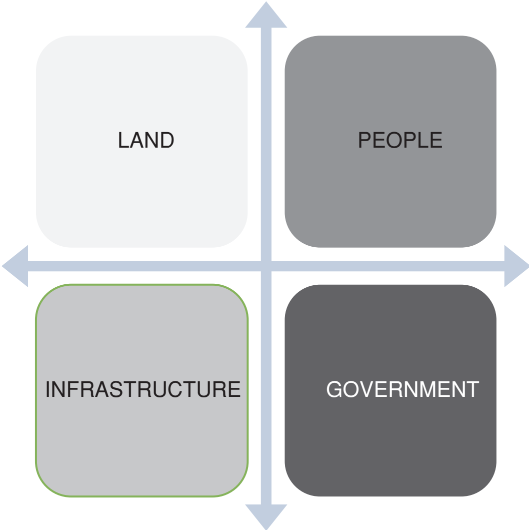
The literature survey shows that the topic is not so recent, because researchers started to study the ICT application to urban life several years ago, twenty environ. However, especially the Internet wave and the Web 2.0 technology have been the main drivers for the development of the digital city research topic.

One of the most interesting aspects regarding smart city and digital city is the use of heterogeneous terminology to define them. People often uses the same word to define different things, and in the same way different words are used to define the same thing.

Analyzing the international literature about this aspect, the following keyword definitions have been found, to be compared with the smart city definition explained above [21].

- Intelligent city. It is a city that has several competences, able to produce knowledge and to translate it into unique and distinctive abilities; it is also able to produce synergies from knowledge and competences mixed in an original way, difficult to imitate; this city is smart because it is able to create intellectual capital and to ground development and well-being on this intellectual capital [22].
- Digital city. It is a wired, digitalized city, using ICT both for data processing and for information sharing, but also to support communication and Web 2.0 democracy [23, 24].
- Sustainable city. It is a city that uses the technology to reduce CO_2 emissions, to produce clean energy, to improve the buildings efficiency; it aims to become a green city [25].

Fig. 2 The basic components of a city



- Technocity. It is a city that uses the technology to improve the efficiency and effectiveness of its infrastructures and services: it focuses its smart projects on urban space quality, mobility, public transports, logistic [26].
- Well-being city. It aims to produce the best quality of life for citizens, but also to create regional attractiveness both for people and for business. The technology is only a part of the instruments used to obtain these goals, but also culture, climate, history and monuments are considered important success factors [27].

Obviously, all these concepts are not in contradiction each other, as they share some aspects and are partially overlapping. But to consider all these aspects enlarges at maximum the concept of smart city, and it is misleading both to understand this concept and to compare it with the digital city concept. Too many definitions mean a lack of focus on the really important factors.

To face this complexity, it could be useful to start from the analysis of the concept of a city, especially identifying such components functional to support a smart or digital city implementation. As showed in Fig. 2, we will consider four basic elements to compose a city:

- **land**, that is, the territory on which the city is built, the geographical area on which the city has its own boundaries;
- **infrastructures**, that is, all the material or technological facilities supporting the urban life, such as public and private buildings, streets, transports, production sites, and so on;
- **people**, that is, the citizens living in the city, but also who works or studies in the city, or comes to visit the city or to enjoy there some cultural or leisure facilities;
- **government**, that is, the public powers to govern the city and the public administrative agencies to manage and supply public services.

Also in the city tout court, all these components are not so well-defined.

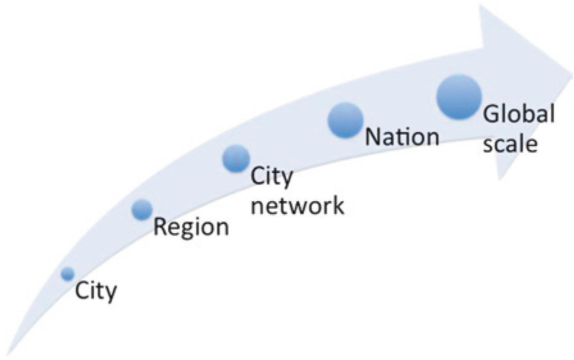


Fig. 3 The territorial dimension of city and urban policies [28]

Regarding the **land**, the territorial dimension not ever corresponds to the administrative boundaries of a city. Sometimes, a city extends its role of economic and social attractiveness well beyond its administrative boundaries. OECD is developing a new way to define metropolitan areas, using a methodology based on the economic function of the city, rather than its administrative boundaries [29]. Also the political aspect is important; in Italy a deep reform of administrative metropolitan areas is underway, extending the administrative boundaries of large cities to the metropolitan area interested by common public services and characterized by high population density and working fluxes from the neighborhoods to the city centre [30]. Sometimes cities link together to create city networks, to share best practices and face together deep urban problems; and not ever these cities are contiguous, but perhaps they are similar in their own characteristics.

Infrastructures are one of the most important aspects of the quality in urban space. Private and public buildings, and their quality, create the urban skyline and define the city character. Streets, traffic and public transports heavily impact on the quality of urban life, but infrastructures also have an important role in the quality of urban environment. Buildings and transports consume energy and produce pollution; they play a double role, both positive and negative, on the quality of their city.

Regarding **people**, it is too simple to include in the city perimeter only who resides in the city. Cities are daily interested by fluxes of workers and students living in the neighborhoods and reaching their own work place or school or university. Moreover, cities are visited by travelers for work or tourism.

About **government**, urban policies are defined not only at urban level, but also at the regional, or national or global level; therefore the urban area and its form are fuzzy and they change depending on the topic, the action, the project, … In Fig. 3 the different levels of urban policies and government are showed. They go from the local dimension, to regional, network, national and finally the global dimension.

Therefore, all the basic components of a city could be seen both from the positive and from the negative side, considering their impact on the urban quality of life; the city dimensions could be the main driver of both a city success and its problems. How smart city and digital city strategies could help to face and solve these problems, but also to highlight the good resources of a city? In the following

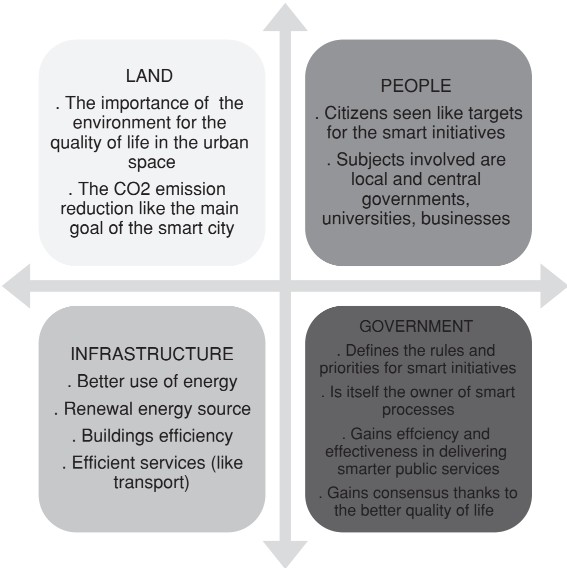


Fig. 4 The basic components of the smart city

paragraphs the basic components of a city are examined considering smart city and digital city strategies, to explore the achievable goals to improve the quality of life in the urban space.

2.3 Smart City

The smart city idea has born in the nineties, but only recently it has become a current topic. Two are the main reasons: the use of the word smart to indicate the so-called smart devices like smart phones, tablets, and so on; and the impulse of the EU to implement smart cities, conceived like low emissions cities, with the main aim to reduce CO_2 emissions.

Therefore, the idea of smart city is mainly focused on the use of high technologies to improve the quality of urban infrastructures and to reduce their environmental impact in the metropolitan area. Indeed, the EU impulse is so strong to overcome all the previous academic visions, based more on knowledge and human capital in city, than on the environmental aspects. Depending on the EU vision, the basic components of the smart city are introduced in Fig. 4.

The **land** component is mainly considered looking at its environmental dimension. Pollution, traffic, waste and energy consumption are important aspects of the daily urban life, they have a high cost for both the citizens and the public administration, they are able to differentiate nice, clean, livable cities from dirty and unlivable ones. The CO_2 emissions in the urban areas are under the main attention of governments at the global level (see the Kyoto protocol) and the EU strategy for

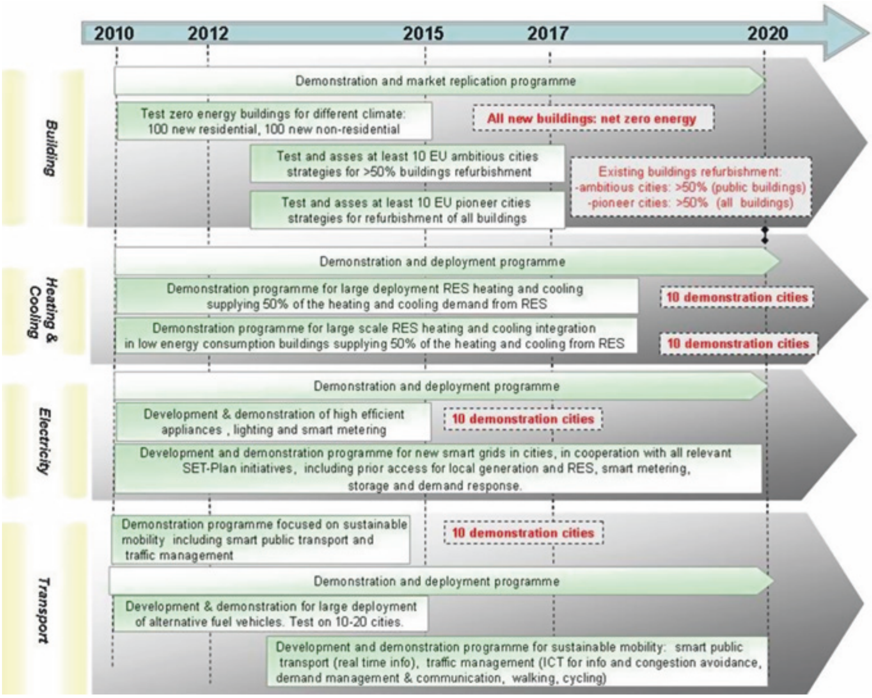


Fig. 5 The EU-SETIS plan implementation

better quality of life in metropolitan areas especially focuses on this goal, easy to define and to measure. Therefore, the land dimension in the smart city is to be considered in the material, environmental sense.

Also the **infrastructure** dimension has a material meaning; streets, buildings, public transport facilities are the instruments to both supply services to the citizens, and to reduce the CO_2 emissions by their quality improvement. A summary of smart city EU strategies could be found in the EU-SETIS program; it is focused on four pillars: clean energy production, low energy consumption, buildings efficiency, sustainable transport (Fig. 5). Each of these pillars has several subsets of goals, strategies, actions and projects [30]. Several targets regard the technological research, the development of prototypes, the implementation of new solutions to support energy, transport and buildings characterized by the near-zero impact goal and the concurrent better quality of public services [31].

Regarding **people** in the EU smart city vision, the role of citizens is not very proactive, as they are mainly seen like the address of this strategy. The subjects involved in the smart city strategy are the triple helix subjects, that is: public administration, universities and research centers, businesses. They play their own role (to govern, to discover, to produce), but are involved to cooperate to design better answers to reach several different goals in the same time: to improve the quality of the technical solutions, thanks to the research outcomes; to deliver better public services and public value to the citizens, thanks to the capacity of local and central government to drive the technological solutions towards the real needs of the people; to create economic value and development, thank to the capacity of companies to produce the desired products and services.

Regarding **government**, local government is generally the main actor involved in supporting smart city projects. Municipalities have been everywhere the first mover in implementing city wide programs regarding smart and/or digital plans for city. Central government plays a key role especially in supporting the city choice to implement a smart city program. However, an important role for European cities is played by the EU; indeed, the scarcity of financial resources available for the municipalities drives the local government to try to obtain funding from the EU programs about smart city. For this reason, smart city strategies are mainly driven by the SETIS program and the EU addresses to implement low carbon programs and projects in urban areas. Only recently, defining the Horizon 2020 goals, the EU changed a little its vision about the smart city idea, conceived now like a larger plan focused not only on the energy pillar, but on three main aspects: economic development, sustainability and inclusion. This new trend in smart city strategy for European cities enlarges its perimeter from the material aspect to the socio-economic aspects, putting inclusion and the social impact of better city at the top of the 2020 agenda.

However, till now the smart city is still considered like a set of strategies and programs regarding the reduction of CO_2; therefore, the smart city perimeter is defined through its goal, and not by the technologies used to reach it. For this reason, smart city is a heterogeneous idea, using several technologies, applied to several and different topics, with the common aim to reduce the environmental impact of city life. In this sense, it is quite easy to define the smart city boundaries, its goals and also to measure the reached results. But both in the academic literature and in the empirical applications, smart city is defined differently and with a larger scope and it is the main reason for its open-endedness.

2.4 Digital City

The idea of a digital city also has born in the nineties, and it has become to spread especially in the so-called Internet era, at the beginning of the millennium. The use of the web both in the private and in the public sector, the social networks and other communication means, the e-services delivery and the availability of rich and up-to-dated online information are the main drivers to implement a digital city, able to exploit all the ICT instruments and devices to create a virtual urban space.

Respect to the smart city, based on several technologies, the digital city is based on the ICT. It means that the main digital city aspects regard:

- the diffusion of rich and updated information online;
- the use of social media or other communication media, to both connect people each others and to create a dialogue between the citizens and the public administration;

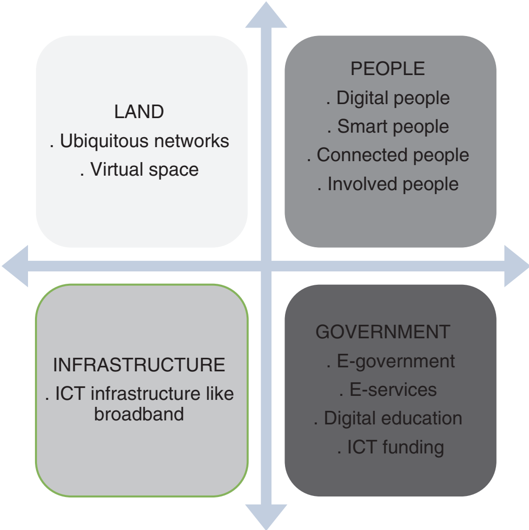


Fig. 6 The basic components of the digital city

- the e-service delivery, by both public agencies and private entities and companies;
- the ubiquity of information, communication and services, thanks to the mobile technology.

Also the digital city could be described adapting the basic components of the city to its characteristics, as shown in Fig. 6.

The **land** dimension is not very important, actually the ICT is used to overcome the material boundaries of cities to create relationships between citizens, among citizens and the public administration, between citizens in the same city or in different geographical areas, all over the world. Indeed the digital city, when fully implemented, is able to support ubiquitous networks and to create a virtual space.

The main component in a digital city is the **ICT infrastructure**, especially the Internet connection based on the broadband. This is the main driver of the digital city implementation. However, also other aspects are important, for example:

- the diffusion and use of smart devices among citizens;
- the high speed connection;
- cloud computing;
- open data;
- system security and resilience;
- and so on.

Some of these components are realized thanks to the cooperation between public administration and enterprises; a wired city needs a comprehensive project, able to support long-term strategy and investments.

Another crucial digital city component is **people**. Indeed, the ICT infrastructure itself is not enough to support the digital city implementation: also the involvement of citizens is necessary. Indeed, otherwise respect to the smart city, the role of people in implementing a digital city should be highly proactive, because they should take part in communication, data processing, information use and e-service enjoyment. For this reason, a digital city should include digital, smart, connected, involved people, able to enjoy the benefits deriving from this urban strategy. One of the main obstacles in digital city implementation is not the broadband diffusion or the lack of high-speed connection, but the overcoming of the digital divide and the increase of people access to digital knowledge and services.

The **government** component is summarized by the word e-Government; the digital city is the main instrument to deliver e-services from the public administration to citizens, aiming both at reducing the service cost and at improving the service quality and effectiveness. However, the complete implementation of the e-government strategy meets several strong obstacles like the digital divide, the lacks of public funding, the lack of digital culture in the public administration. The government should increase the digital readiness of public bodies and workers to success in the digital city strategy [32].

The previous dimensions of the digital city are at first glance deeply different respect to the smart city dimensions. However, smart city and digital city often are two faces of the same coin, as explained in the further paragraph.

2.5 Smart city and Digital City: Two Faces of the Same Coin

The analysis of the literature and of the empirical implementation of some smart or digital city prototypes shows us that smart city and digital city are different in their components, enabling technologies and goals. However, they are often linked together in the urban strategies for better quality of life. They have been also confused in several academic papers and public policies. There are two main reasons.

The first depends on the use of some words like smart, digital, green, to define innovative urban policies, without a clear reference to a sound definition or standard. This is therefore a terminological confusion, but it has few impacts on the concrete implementation of smart or digital city programs.

The latter derives from the interlaced role of technologies and goals, that needs of both smart and digital projects and actions to realize a better city for people.

In Fig. 7 the main components and actions of smart city and digital city are summarized. It is evident that infrastructures and solutions regarding the digital city are useful—or necessary—also for the smart city. Several plants and devices used for smart transport or energy efficiency are based on ICT and wired houses, buildings and cities. Data availability and processing are crucial to support planning and delivering smart products and services and e-government, e-commerce, e-business are the instruments to exploit the smart initiatives.

Digital city	***Digital infrastructure*** - new ICT infrastructure - high speed broadband - fibre optic cables - wireless technology - networked information systems	***Data*** - Data collection, storage and analysis at city level, potentially through the cloud, which can enhance a city's ability to predict and plan for the future ***Information*** - Processing of information to service programmes ***Service development*** - Development of service application	
Smart city	***Smart transport and Mobility*** - Bike schemes - Real time bus timetable information - Electric Vehicle car pools - Congestion charging	***Rebewable energy & energy efficiency*** - Combined Heat and Power - Renewables - Electric Vehicle Charging Poimnts - Sensor to monitor traffic, pollution, emissions - Street lighting - Waste collection systems - Smart grids	***Smart and sustainable buildings*** - Smart meters - Energy efficiency measures: insulation, low energy lighting, efficient boilers - Building Integrated Renewables - Electric Vehicle Charging Point - Smart appliances - Motion detectors - Automatic weather forecasting

Fig. 7 Comparing smart city and digital city

The comparison between smart city and digital city and the analytical individuation of their components aim not to separate, but to create a sound basis for the strategy definition of quality of life in urban areas. The role of technologies, environmental quality, energy safety, information and communication access should work together, but with the awareness of their differences, and not putting all of them on the same footing.

The separation between smart city and digital city could be functional to better investigate about what and how to plan smart and digital strategies, and especially how much results and returns are awaited and finally reached. However, it is important not to be wrong, considering smart city and digital city two different, separate urban strategies. They should be linked together and harmonized to individuate priorities and better investments to create the maximum outcome and public value for citizens.

To support this vision of smart city and digital city, defining them like two different but integrated innovation paths for urban areas, the empirical analysis is necessary. Two important cases have been examined: Amsterdam, The Nederland and Genova, Italy. The analysis is deep and it permits to understand how the smart or digital city idea has born, how it has been developing during time and which are the main aspects of urban strategy in these two cities. For a complete empirical research, two aspects are examined: the key partners involved in the project and the initiative portfolio. Actors and partners define how the project is thought, that is, if top down—drove by the local government, or bottom up—gathering the private and public initiatives by enterprises, associations, citizens and so on. They define also the work method, centralized or federal or some other topological choice. Initiative portfolio permits to understand which are the contents prioritized by the city, and to outline if they have a smart or a digital profile, or both. It is helpful to define the characteristics of a smart/digital city strategy.

The aim of this investigation is to understand if these two cities, defining themselves smart city, are pursuing a smart city strategy, a digital city strategy or a blend of them.

3 Case Study: Amsterdam

3.1 Introduction

Amsterdam Smart City is universally recognized like the first smart city not only in Europe, but in the world. However, the development of Amsterdam Smart City has crossed several different phases, starting just from the digital city strategy. The literature analysis helps us to discover how the Amsterdam case has been becoming the most important in the smart city panorama.

Looking at papers regarding Amsterdam Smart City, Google Scholar shows that the first writings regarding this topic date 2009, whereas papers regarding Amsterdam Digital City date from 1995. Indeed, the Digital City concept has born just in Amsterdam in 1994, when ICT was used to create an online connection and community to enforce Amsterdam citizens in facing political election. Amsterdam Digital City is therefore in its first phase a political and social instrument, arranged by people to communicate and exchange political opinions. Environ 170 papers focus on the Amsterdam pioneer case in digitalizing a city from 1995 till today.

The high success obtained by this project—140.000 subscribers in few months in 1994, well before the Internet boom—was the motor to transform an occasional initiative in a permanent instrument to connect people in the city. However, as the Digital City platform was not a public initiative, but a private project, public funding were not enough to support the infrastructure and its daily functioning, therefore the Amsterdam Digital City became a company and started to test some new business models to use e-commerce for financing the social side of this initiative.

Unfortunately, these economic returns were not enough to support Amsterdam Digital City and this project had a certain decline, especially at the beginning of the new millennium. At the end, we should say that this interesting and pioneering experiment failed to become a sustainable local information and communication infrastructure, but opening new paths of urban development.

In the meantime, the awareness of the city environmental footprint begun to grow; Amsterdam was one of the first cities to think about a strategy to face pollution and energy consumption in urban areas. In 2009, three subjects: Liander, a grid energy operator, Accenture, and Amsterdamse Innovatie Motor, a public agency founded to support innovation in the city of Amsterdam, joined their forces to create the Amsterdam Smart City program; its aim was mainly to create collaborative pilot projects to support a better use of energy and a reduction of pollution and CO_2 emissions in Amsterdam. From this date, we have found 84 papers speaking about Amsterdam Smart City.

Amsterdam assumes the following definition regarding smart city: “A city is smart when investments in capital and communication infrastructure fuel sustainable economic growth and a high quality of life, in combination with an efficient use of natural resources”. Applying this definition, the Amsterdam Smart City partnership defines its own strategies to build a smart city in Amsterdam urban area.

However, Amsterdam was no more the first mover, as it was for the digital city experience: [33] show that a lot of large and medium cities in Europe begun to go through the smart city path, even if most of them are not aware of this strategy or are pursuing smart goals without using a clear or explicit smart city framework.

Nevertheless, Amsterdam is an interesting case study, mainly because during its digital phase it developed a virtual community and the people involvement is at the basis of both its digital and its smart strategy. Moreover, the capability of Amsterdam Municipality to involve also private actors and to design a comprehensive smart city plan, able to include near every aspect of the urban life, qualifies this experience at the top level in Europe. In the further pages, we will examine the Amsterdam case to understand the relationships between digital and smart aspects and if these two paths are alternative or complementary.

3.2 Key Players

As already explained, the Amsterdam Smart City initially started like a Digital City initiative in 1994. Only in 2009 the municipality, with some key partners, moved towards a clear smart city project. One of the most interesting aspects of the Dutch experience is the involvement of several players, belonging to different but complementary categories.

To examine the key players both in Amsterdam Digital City and in Amsterdam Smart City, we should organize our analysis in two streams:

1. who are the “shareholders”, that is, who decides about the planning of a smart or a digital urban strategy;
2. who are the “stakeholders”, that is, who benefits from the smart/digital urban strategy implementation.

We clear also which are the subjects participating to the different implementation processes.

In the Amsterdam Digital City experience, first mover were the citizens, organized in associations; precisely, the political-cultural center The Balie and the computer activists group Hacktic launched the DDS (in Dutch: De Digitale Stad, abbreviated as DDS) as a ten weeks experiment to provide an electronic democratic forum to the citizens of Amsterdam. The pilot project had a great success and it continued well over ten weeks, till 2001.

DDS has been conceived like an information platform, designed like a virtual city, hosting several private and public institutions sites, and also citizens' ones, to deliver data and information to the registered users. Institutions cover several categories, such as health, education, ICT, leisure, media, politics, and business bodies. These subjects are the shareholders, that is, the key actors aiming at using DDS to diffuse their own information and to publicize their activities among citizens, both for commercial and for social aims. DDS has born and ever remains a flat initiative, with nor governance nor formal leadership. Perhaps this lack is also one of the reasons of the failure—or better the extinction—of this pilot project; nobody had enough interest to invest important sums of money in the maintenance and innovation of this platform, and it was just for financial reasons that the initiative expired. Moreover, DDS was not able to renewal its offer and to face competition from followers in the use of the Internet to provide information to the citizens.

DDS was a first experiment of social platform to share information about the life in the urban area. For this reason, its stakeholders were for the first the citizens, even if DDS attracted visitors and users from elsewhere, more interested to the innovative communication medium than to the contents. However, during its life, DDS involved more and more business players, offering free information but with the aim to publicize their products and services and to attract customers. Therefore, DDS lost its social profile to acquire a public–private nature; stakeholders are therefore also the business system and the economic players in the city of Amsterdam.

A few important role has been played by the public institutions; some of them—schools, hospitals, and so on—participated to the initiative with their own web site, but they hadn't a leader role in the DDS. Therefore, we can conclude that the DDS was a bottom-up, flat program to share information among citizens in the Amsterdam urban area, without a formal organization or governance structure.

Very different is the experience of Amsterdam Smart City initiative. In 2009, the Municipality of Amsterdam begun to think about some instruments and projects to face the problem of pollution, energy consumption and environmental quality in city. The "Am*smart*erdam city" project has been founded on this basis. The first mover is therefore a public body and the initiative is top down, as it is driven by a pool of four founding partners, involving in the following several other actors. They are the shareholders of the initiative.

To implement the Amsmarterdam program, the founders settled an association to gather all the players working for the smart goals. Therefore, the governance platform is a closed one, including all the associated partners, and a hierarchical body, because the main actors are the founding subjects, that is: Amsterdam Economic Board, Gemmente Amsterdam, KPN and Liander. If in the DDS the first mover had been a private subject, in the Amsmarterdam initiative it is a public subject and the shareholders are both public and private. Finally, in the DDS project each partner was working alone and there were no interactions between all the

Table 2 Key characteristics in Amsterdam digital city and in Amsterdam smart city

	Amsterdam digital city	Amsterdam smart city
Starting process	Bottom-up	Top-down
Participation	Open	Closed
Structure	Flat	Hierarchical
First mover	Private body	Public body
Actors	Mainly private ones	Public–private partnership
Governance	No interactions between the actors (self organizing platform)	Formal organization (Quadruple helix model)

participants to the DDS platform. In Amsmarterdam, there is a strong connection and cooperation between all the shareholders of the initiative; the aim of the platform is to keep together different categories of players, such as public bodies, universities and research centers, companies and social bodies, to build a quadruple helix able to create also a regional knowledge network to enforce the smart city development in the future.

The Amsmarterdam initiative involves these actors, including also social bodies and therefore the citizens, also if their active role is few represented. Indeed, the citizens are the final stakeholders of the Amsmarterdam project, but they obtain benefits in a mediate manner, that is, thanks to the improvement of environment and life quality in city. Therefore, even if the citizens are the final stakeholders of this urban strategy, they are often not really aware of this.

This analysis shows that DDS and Amsmarterdam—digital and smart strategies in Amsterdam—are very different respect to the role of key players. Their differences are summarized in Table 2.

The quadruple helix model describing the governance and cooperation model in the Amsterdam Smart City initiative is not a declared choice, but the result of the urban vision pursued by the key players and the cooperation model they are trying to implement. The most known triple helix model is a theoretical framework explained by [34]. This model refers to a spiral involving different categories of actors playing in different stages of knowledge capitalization: public sector, industry, and academy. Thanks to their relationships, they are able to support a faster, deeper and higher value innovation process.

However, in the triple helix model the civil society is not included and it is not considered like a key actor in innovation process. The smart city idea is quite different, because it considers the involvement of citizens like a winning weapon to build successful smart strategies. This idea includes citizens and the civil society not only like stakeholders of the smart initiatives, but like active actors, playing a crucial role in supporting innovation in the culture, knowledge and mentality of people, changing their behavior towards a smart awareness.

Amsmarterdam is an initiatives explicitly involving not only the civil society through the active participation of citizens and social bodies in defining the smart priorities and projects, but also declaring that active behavior of citizens and knowledge sharing permits the successful smart implementation in urban spaces.

The quadruple helix in Amsmarterdam is not an explicit strategy, but the result of strategic choices regarding the active role of the citizens' intellect, awareness and commitment. Citizens are therefore both the main shareholders and stakeholders of Amsterdam Smart City.

3.3 Initiatives

To realize its own goals the Amsterdam Smart City partnership defines an implementation strategy including a initiative portfolio; each initiative in some way contributes to create a smart city in Amsterdam.

At present, this portfolio is made by 43 projects, organized by 5 themes and regarding three geographical areas inside Amsterdam urban boundaries. These projects are very heterogeneous, on all points of view: involved actors, applied technologies, role of citizens, and so on. However, all of them are mainly focused on energy transition and open connectivity. These streams recall both smart city aims (energy transition) and digital city ones (open connectivity). Also the definition of a smart city used by Amsterdam to drive its activity recalls investments in communication infrastructure and the aim to pursue sustainability. It seems therefore that the present Amsterdam City strategy includes both smart and digital initiatives.

To better verify these hypothesis all the 43 projects have been deeply analyzed, examining both their content, their aim and the involved actors. To understand if a project is smart or digital, or both, and to classify it depending on its nature, goals and technological contents, a schema has been defined, explained in the following Fig. 8.

The main classification in smart or digital initiatives derives from the smart city definition suggested by Amsterdam Smart City and related in Sect. 3.1. In this definition, a smart city should both invest in ICT and obtain sustainability, that is, environmental footprint reduction and a better use of natural resources. In Fig. 8, we define smart initiatives the ones aiming at sustainability, and digital initiatives the ones based on ICT, web communication and data sharing. Moreover, we consider some other factors in classifying the smart/digital projects.

In smart projects, we consider also the use of ICT like functional technology (that is, ICT is not the aim of the project but the instrument to realize smart goals) and the high or low involvement of citizens: indeed, smart projects could be essentially technological, applied to buildings, transport facilities and other infrastructures without involving the proactive behavior of citizens, or on the contrary they could base their success on the concrete participation of people.

In digital projects, we consider also the eventual impact of digital initiative on smart goals: for example, an ICT system aiming at monitoring energy consumption in private houses has also a smart impact, as it drives people's behavior towards a better use of energy through their higher awareness about consumption.

The project portfolio analysis is showed in Table 3. In column 1 a progressive number is reported, column 2 contains the name of the project, column 3 a brief description; column 4 has the label SC for smart project, DC for digital project,

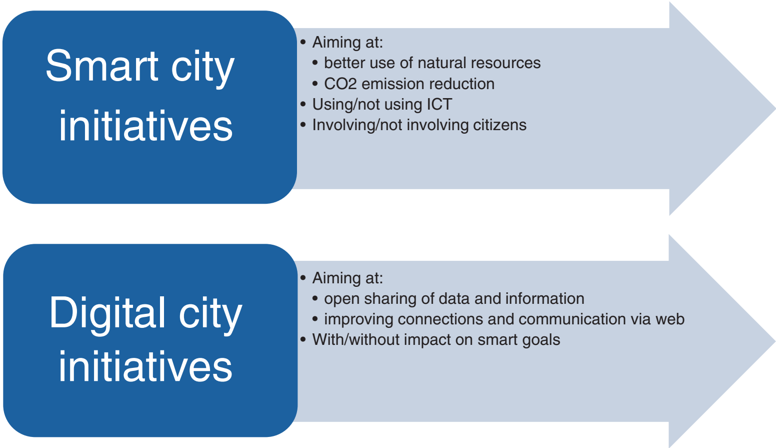
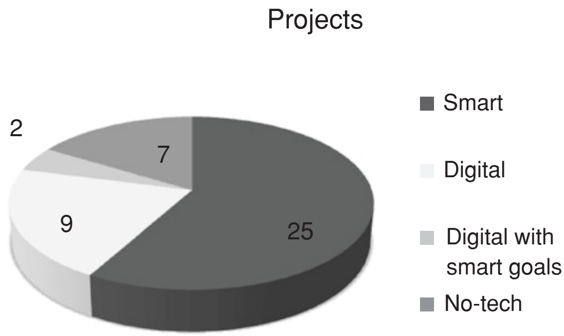


Fig. 8 Smart and/or digital projects classification

Fig. 9 Smart and/or digital project range in Amsterdam Smart City



NOTECH for project not based on technology, DC → SC when a digital initiative produces outputs also on smart goals; column 5 reports other specifications, as described in Fig. 8: EFF for smart projects aiming at energy saving and environmental impact, +ICT for smart projects with a strong ICT base, PEOPLE for smart projects involving active citizens participation, DATA for digital projects aiming at open data and information sharing, COMM for digital projects aiming at a better communication with citizens.

This analysis shows that 25 projects out of 43 are smart projects; 9 are digital projects; 2 are digital projects with a strong impact on smart goals; 7 projects are no-tech projects, that is, initiatives aiming at smart goals, but without using technology. For example, these projects regards to human behavior, legal instruments, and so on, to improve city sustainability. The range of projects is displayed in Fig. 9.

Table 3 Project portfolio in Amsterdam Smart City

	Project	Description	Type	Spec
1	Almere smart society	Almere Smart Society, is a vision of living and working in Almere, in all its facets supported by ICT and technology	DC	DATA COMM
2	Amsterdam free Wifi	Offering Free Wifi on Amsterdam IJburg harbour connected on KPN consumer fiber	DC	COMM
3	AmsterdamOpent.nl	AmsterdamOpent.nl is the platform where civil servants of the city of Amsterdam can propose questions and ask the Amsterdam people to share their ideas	DC	COMM
4	Apps for Amsterdam	Apps for Amsterdam 2 is the second open data contest of the municipality of Amsterdam in which developers are challenged to build apps based on municipality's data	DC	DATA
5	Climate street	Together with entrepreneurs, a typical Amsterdam street, the Utrechtsestraat, is transformed into a sustainable shopping street where innovative technologies are tested	SC	EFF PEOPLE
6	E-harbours —Innovative energy contract Zaanstad	The Municipality of Zaanstad has negotiated a new energy contract, that saves energy, stimulates the local production of renewables, and reduces energy costs substantially.	NOTECH	EFF
7	E-Harbours—ReloadIT	The core of Zaanstad's showcase 'REloadIT' is innovative technology for clean mobility	SC	EFF
8	Energy management Haarlem	250 customers in the Haarlem region tested an energy management system free of charge for four months	SC	EFF PEOPLE + ICT
9	Flexible street lighting	No description found	SC	EFF
10	Fuel cell technology	Using innovative local energy generation technology will enable the "Groene Bocht" building to provide in its own electricity and will reduce CO_2 emissions by 50 %	SC	EFF
11	Geuzenveld—Sustainable Neighborhood	More than 500 homes were provided with smart meters and some with an energy feedback display that can make residents more aware of their energy consumption	SC	EFF PEOPLE + ICT
12	Health-Lab	Health-Lab is a collaboration between companies, government, care and research institutes to stimulate ICT and Care developments	DC	DATA

(continued)

Table 3 (continued)

	Project	Description	Type	Spec
13	IJburg—Fiber-to-the-Home	A new fiber network has been unrolled by Reggefiber in cooperation with KPN to facilitate the inhabitants of Amsterdam with 3 play services	DC	COMM
14	IJburg—Smart Work@IJburg	Amsterdam Smart City offers IJburgers alternatives for the traffic jam: work at home or at a Smart Work center	DC	COMM
15	IJburg—Wijk TV	A local private TV channel via fast fiber internet	DC	COMM
16	IJburg: YOU decide!	In the EUDI (End User Driven Innovation) project, IJburgers are asked to describe their issues and ideas on energy and mobility in their neighborhood	NOTECH	
17	IRIS—Research into the legal frameworks of energy provisions	The goal of project IRIS is to establish legal frameworks that offer the best opportunities to develop local sustainable energy provisions	NOTECH	
18	ITO	By applying Smart Building technology, even a modern building like the ITO Tower can greatly reduce its energy use	SC	EFF + ICT
19	Moet je Watt—charging system	The Moet Je Watt (MJW) is a smart electrical battery charging system for electrical cars that communicates with a smart meter in the meter box to prevent power wastage and overcharging. The purpose of this project is to test the combination	SC	EFF
20	Monumental buildings	The purpose of the project was to find out which technologies and methodologies are practical when it comes to rendering monumental buildings. The shared office building, De Groene Bocht, was a part of this pilot	SC	EFF
21	Municipal buildings	Measuring energy consumption in municipal buildings via an online portal enhances awareness and shows that energy-saving measures do yield real results	SC	EFF + ICT PEOPLE
22	Nieuw-West—City-Zen	A total of €30 million will be invested in innovative projects in urban areas of the city in the years ahead, primarily in the District of Nieuw-West	SC	EFF
23	Nieuw-West—Energy storage for households	Technology development for energy in households which are linked to the smart grid	SC	EFF + ICT PEOPLE

(continued)

Table 3 (continued)

	Project	Description	Type	Spec
24	Nieuw-West—Serious Gaming	A serious Game designed to playfully enable and encourage bottom-up participation of residents in creating a Smarter City	NOTECH	
25	Nieuw-West—Sloten Windmill: smart meeting spot	The Sloten Windmill is a meeting place in the district of Nieuw-West. Together with local residents and our partners, we are working to develop and introduce smart initiatives throughout the district	NOTECH	
26	Nieuw-West—Smart Grid	In the Amsterdam New-West area the first intelligent self healing grid has been implemented with which the city of Amsterdam can realize its sustainability objectives	SC	EFF + ICT
27	PLAY DECIDE	A discussion card for young and old that aims to raise awareness of participants on the topic of smart cities and the city of Amsterdam, intrigue them to see the theme of smart cities in a critical and rounded way and enhance their debating skills	NOTECH	
28	Ring–Ring	Bicyclists are worth everything. By choosing to ride a bike over other transportation, the environment, public space and our own health benefits from this directly. That should be rewarded	NOTECH	
29	Ship to grid	Almost 200 shore power stations are installed allowing ships to connect to green energy instead of relying on polluting on-board diesel generators for their power supply	SC	EFF
30	Smart challenge	Eleven companies compete in the Smart Challenge where the Wattcher gives employees insight in energy consumption. The winner of the contest will be the company whose employees save the biggest amount of energy	SC	EFF + ICT PEOPLE
31	Smart schools contest	In the Smart Schools project 6 primary schools in 7 locations compete on energy efficiency program results, comparing performance through an online portal	SC	EFF + ICT PEOPLE
32	Smart sports parks	In this project sports associations, the local city council and several entrepreneurs work together to build and maintain sustainable and strong sports grounds. Main focus is on Energy Efficiency, Smart Lighting, Bio Diversity and Shared Resources	SC	EFF

(continued)

Table 3 (continued)

	Project	Description	Type	Spec
33	Smart traffic management	Amsterdam has it's own 'virtual traffic manager', a technical tour which enables traffic to be managed almost automatically. Unique in The Netherlands	DC → SC	COMM DATA
34	Swimming pools	Swimming pools are public buildings that consume a great deal of energy. Amsterdam Smart City wishes to work with partners in an effort to find sustainable, energy-efficient solutions to secure maximum cost-efficiency in swimming pool management	SC	EFF
35	The Green Canals of Amsterdam	"De Groene Grachten", an initiative of Wubbo Ockels, has as purpose to make the canal ring of Amsterdam sustainable	SC	EFF
36	The smart home	Benext-iHome demohouse, in which more than 60 domotics products control the house and minimize the energy consumption	SC	EFF + ICT
37	TPEX—Smart Airmiles	To operate TelePresence Conference Centers (meeting rooms, board-rooms or classrooms) in Amsterdam and environs, connected to a worldwide network of international conference centers	DC	COMM
38	Watt for Watt	Watt for Watt uses a neighbourhood-level approach to improve the energy efficiency of houses. The campaign is also dedicated to making residents aware of how they use energy with the aim of keeping energy costs at an affordable level	SC	EFF PEOPLE
39	WEGO Car sharing	WeGo is a new sustainable platform that allows neighbours and friends to safely rent their cars to each other	DC → SC	COMM
40	West Orange	400 households in Amsterdam tested a new energy management system. This system can make residents more aware of their energy consumption and will help them to save energy	SC	EFF + ICT PEOPLE
41	Zuid Oost—Laws and regulations	Free zone for sustainable energy	SC	EFF
42	Zuidoost—Energetic Zuidoost	Energiek Zuidoost wants to reduce the ecological footprint of roughly the area between Amsterdam Arena and the hospital AMC	SC	EFF
43	Zuidoost—Stakeholders in the drivers seat	Value case development in a area where several stakeholders work together to develop the area in a sustainable and integrated way	SC	EFF PEOPLE

Further considering Smart projects, 16 out of 25 are based on a strong participation of citizens in implementing home technologies to improve sustainability in private spaces, or in modifying their behavior to reduce Amsterdam environmental footprint. It means that Amsterdam Smart City is a strong-human-based strategy, were technologies and behaviors should work together to reach the expected results. It is confirmed also by the high rate of no-tech projects, 7 out of 43, demonstrating that a smart city is not only based on technologies, but also on best practices and awareness. Moreover, 10 out of 25 smart projects requires a strong role of ICT in implementing digital platforms, control systems, sensors or other digital devices integrated in other plants or buildings, transport facilities and so on. It means that digital and smart are two attributes difficult to separate in smart contexts.

Further examining digital projects, 7 out of 11 regards an improvement of web-based communication between citizens, or between citizens and public administration bodies. 2 out of 11 are based on open data and 2 out of 11 mixes data sharing and communication. Also these evidences show that the involvement of citizens is at the core of smart city strategy.

3.4 Analysis

The history, the actors and the projects portfolio of Amsterdam Smart City are at the basis of our empirical analysis to understand what a smart city is, if it is similar or different respect to a digital city, where and how much they overlap and mix each other and so on.

Originally, Amsterdam knew an important, pioneering experience of digital city, in 1994. This experience was born from the citizen, it was a bottom-up initiative and it was able to involve thousands of citizens, using the Internet and creating the first digital community in the world. However, the Amsterdam Digital City project failed, especially because it was not able to create the conditions for its economic survival.

In 2009, the Municipality of Amsterdam started a new experience, labeled smart this time. No surprise that, considering the reasons of the failure of the digital experience, nowadays the first actor is the Amsterdam Economic Board, a public body representing governmental agencies, research institutes and the business world. Therefore, the economic dimension plays a key role in implementing the smart plan.

The main goals of Amsterdam Smart City are two: economic development and quality of life. Quality of life is the instrument to attract young and educated people to live in Amsterdam, producing therefore the economic development. The quality of life is obtained mainly through three different paths: environmental quality, digitalization of public and private communication and services, and a more general supply of public services and facilities. These paths are the drivers for determining the goals of smart city initiatives, that is: a better use of natural resources; a strong attention towards energy consumption, clean energy production and reduced environmental blueprint, especially conceived as CO_2 emission reduction; a pivotal role of ICT, web communication and data sharing, continuing the tradition of Amsterdam Digital City, but with a top-down process this time; a special focus on people, their behavior, their inclusion, their democratic participation to the city planning.

All these aspects—environmental attention, digital maturity and high democratic sentiment—traditionally define the cultural profile of The Nederlands. Therefore an idea of smart city based on these drivers is easy to share with Dutch citizens but also to transmit to who wants to reside in Amsterdam. It outlines also the need to define smart strategies well rooted into the culture and the specific history and profile of each city; no standard smart strategy exists, but standard themes specified in each specific city.

Examining the project portfolio, we could also answer to the question, if smart city and digital city are the same thing or if they are different, and if Amsterdam is a smart city, a digital city, both of them or smart/digital at the same time, without distinction of these two urban strategies.

Our survey permits to say that smart city and digital city are indeed two different things. A close and delimited definition of smart city says that a smart city is a strategy aiming at improving the environment quality in the urban area. A close and delimited definition of digital city says that a digital city is a strategy aiming at wiring and digitalizing data, information and public and private services in the urban area. These close definitions permit to trace well-conceived boundaries between smart and digital. It could be very useful to both classify cities, strategies, projects, and to prioritize investments, assess policies, evaluate expected and obtained returns.

However, the reality is not so simple. As we have seen, in Amsterdam a lot of projects classified like smart use ICT, even if a smart project generally uses ICT to process data and not to share information or to connect people; but not ever. In Amsterdam Smart City it is the specific city vision that puts these two urban innovations out the same hat, called smart city program. Amsmarterdam applies a more comprehensive definition of smart city, including both ICT investments and sustainable development. It is therefore a specific, political choice of Amsterdam to join smart and digital initiatives in a unique, large program to improve the quality of life, to sustain economic and social development, to digitalize information and services.

But Amsmarterdam should take into consideration that smart and digital initiatives require different policies. For example, digital initiatives are strongly based on the digital literacy of almost all citizens, to prevent digital divide and to grant the larger participation. It is based on the daily use of web and mobile devices to enjoy digital information and services. It requires therefore a digital maturity of both infrastructure and people. Smart city on the contrary especially requires strong investments in facilities and plants and it is based on active participation of private companies in funding smart investments. Therefore, an effective economic plan should support the smart city implementation, to prevent it fails owing to the lack of financial resources.

Despite that, Amsmarterdam shows all the success drivers to succeed in implementing its smart plan, joining both smart and digital measures.

3.5 Conclusions

The analysis of the Amsmarterdam case has been very useful to better understand the contents of smart city and digital city strategies, to compare these two urban development paths and to verify if the empirical implementation of smart city programs reflects the theoretical definitions.

The Amsmarterdam projects portfolio reveals that a smart city is indeed a mix of smart and digital projects, but also of no-technological based activities. What links together smart, digital and no-tech projects is simply the aim to improve the quality of life in urban space. However, this perimeter would be too large and potentially includes all urban initiatives. We can find two common aspects in all the examined projects, composing the Amsmarterdam projects portfolio: the information and services digitalization and the environmental footprint reduction.

Starting from these empirical evidences, we could rewrite a comprehensive smart city definition able both to include all the smart activities, but also to exclude initiatives out of scope. The definition is: "Smart city is a wired urban space aiming at implementing digital data, services and communication and clean infrastructures, to improve the quality of life in the city through a large web connection and a reduced environmental footprint".

Assuming this definition, a digital city is indeed a subset of a smart city, but a required part, because a city without wired connections and web communications is not conceivable like a smart city. Moreover, the role of ICT in supporting several smart infrastructures in reducing their environmental impact creates a strict relationship between digital and smart technologies.

Finally, the role of citizens has been often neglected in the past implementations of smart city initiatives, giving more importance to the technological aspects. However, Amsmarterdam is a good case to outline best practices in involving citizens in smart and digital projects, aiming at changing their behavior towards more digital relationships each other and with the public administration, and a more careful respect of the urban environment. A smart city becomes therefore also an instrument to increase the democratic participation of people in city government and therefore to create higher consensus and a better quality of life in a social sense. This aspect is not less important, but a core element in the smart city definition and implementation.

4 Case Study: Genova

4.1 Introduction

If Amsterdam is recognized like the first digital city in the world, Genova is the leader city in winning European calls for smart cities. Genova submitted three projects to all the three calls for smart cities launched in 2011, obtaining a funding of 5.5 ml/€, in the amount of 8 % of the total EU funding for these calls.

Moreover, Genova presents a best practice in smart city governance, as it has been the first city creating from the beginning a governance authority to drive smart public policies and smart private initiatives towards a unique goal.

Genova could be defined like a “big bang case” in the smart city strategy; indeed, the idea to participate at the EU calls for funding smart projects has been the first step to start the smart action in Genova. No other initiatives had been implemented before.

The Genova success derives especially from the strength of the team defining the projects and a comprehensive strategic vision for Genova Smart City. This team was initially composed by three big players, that is: the Municipality of Genova, the real mover of the strategy; a couple of large companies in the energy and building industries; and the University of Genova, especially the Polytechnic Faculty. This team includes from the beginning all the main actors able to activate the triple helix and to create a positive synergy in research, innovation and technological transfer from the smart projects to businesses, public bodies and citizens. In the following, Genova settled an association, Genova Smart City Association (GSCA) to drive all the further initiatives, projects and strategies in developing a smart urban area.

The main goal of GSCA was especially to innovate the obsolete public infrastructures, especially in transport, building and energy production, pursuing in the same time the goal to create a more sustainable city. Indeed, the GSCA definition of a smart city recalls the main goals of sustainable cities: “Genova Smart City aims to improve the quality of life through the sustainable development, based on research, innovation and technology, driven by local leadership and applying integrated strategic planning”.

To concretely implement Genova smart city, a large portfolio of actions has been developed, based on 9 big projects and 51 smart initiatives. Each of them is focused on one or more smart goals, but ever aiming at contributing to the shared goal included into the Genova smart city definition. To pursue a comprehensive result, the governance structure and processes are crucial; for this reason, Genova could be considered a best practice case, as it implemented a governance body and specific processes able to effectively drive the multi-purpose, multi-subject smart initiatives towards a unique objective.

4.2 Key Players

To drive and govern the smart strategy implementation, Genova settled a governance body, Genova Smart City Association. It was initially composed by the three main partners participating to the EU calls for smart projects funding, that is: Genova Municipality, Enel Spa (the Italian main electricity producer) and the University of Genova. The first aim of GSCA was to involve the smart city main stakeholders in joining the Association and participating to the smart strategy implementation, but also to the dissemination among companies and citizens of the smart culture.

GSCA is an open association, that is, each public or private body interested in smart actions and projects could join the association, paying a fee and participating to a democratic governance board; each member indeed has voting right to modify the statute, to elect the Directive Committee and to approve the main initiatives. This is the most important characteristic of this original idea, to formally join all the stakeholders in a body working for a shared goal. At present (October 2013) GSCA has more than 70 members and this number is continuously increasing. GSCA has the role to fix the smart agenda, especially aiming at applying the EU smart idea, and to concretely define actions, projects and initiatives to realize the Genova Smart City transformation process.

GSCA has a dual governance framework, composed by two main boards: the Directive Committee, with the role to define the strategic vision and main development paths, and the Executive Committee, to realize the strategies. GSCA President is the Mayor of Genova, to confirm and enforce the role of the Municipality in driving the smart process.

To support the innovation activity, GSCA has also a Scientific Committee, that has mainly a consulting role: it should examine and ratify—or reject—the proposal of actions, initiatives and projects submitted by the members, and it maintains the relationship between GSCA and the research institutions members.

A deeper analysis of the GSCA members reveals that the composition is very heterogeneous. Indeed, we can count several companies, but also a lot of not-for-profit bodies and public agencies; for example, the Port Authority, The Regional Energy Agency, Trade Unions and so on. We can find also trade associations like Industrial Trade, Commercial Trade, Building Trade; together with cooperative companies and Association of Citizens working in culture, welfare and education sectors. The dimension of company members is very heterogeneous, too: we can find several global, large companies like Toshiba, Siemens, Selex, Ericsson, Erg, Ansaldo; but also a large number of SMEs, mainly working in energy or ICT industry. Also research bodies are represented by several members like University of Genova, CNR (National Research Centre) and IIT (Italian Institute for Technology, settled in Genova).

In Fig. 10 the classification of GSCA members is graphically represented. All the members are classified in one of these categories: Public bodies, Research bodies, Large companies, SMEs, Trade associations and Trade unions, Not-for-profit associations.

This panorama suggests that GSCA is a real connector of different ideas and competences regarding the smart city definition, implementation ad dissemination. GSCA is an important example of quadruple helix and it is the main strength for Genova, to create a smart city being at the same time a smart community. Indeed, we already said that sometime smart projects, especially when focused only on technical implementations, tend to exclude the active role of citizens, considered like the final address of benefits deriving from these implementations, but without an active role in the process. On the contrary, GSCA wants to pursue an inclusive strategy, involving all the stakeholders not only in enjoying the benefits, but also in participating to the picture of their desired smart city.

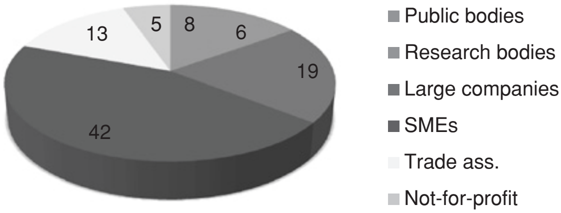


Fig. 10 GSCA members categories

Table 4 Key actors in Genova and Amsterdam smart city

	Genova smart city	Amsterdam smart city
Starting process	Top-down	Top-down
Participation	Open	Closed
Structure	Flat	Hierarchical
First mover	Public body	Public body
Actors	Public, Private and Not-for-profit	Public–private partnership
Governance	Formal organization (Quadruple helix model)	Formal organization (Quadruple helix model)

Moreover, it should be considered that Genova is one of the more aged cities in Europe; citizens over 65 are the 27 % of the inhabitants. It means a low awareness about the smart city idea and a low ICT education level. However, elder people are main stakeholders of smart city initiatives and services; for example, e-health systems, better public transport services, cheaper heating and cooling plants. Therefore, they should be educated and adequately informed and involved in the smart city projects and some not-for-profit members of GCSA are working just for this goal. Only with the higher active participation of all the citizens the smart city could produce and deliver the higher public, economic and social value for all.

In Table 4 we can compare the key players in Genova and in Amsterdam. There are some similiarities and some differences. Both the cities have a top-down process, driven by a public body, that is, the Municipality. It suggests that a smart city project is complex, requires important plans and funding and it is necessary to well define its development paths to obtain effective results. However, Amsmarterdam choose a hierarchical, closed governance model, Genova a flat and open one. Genova, even if strongly focused on smart initiatives interesting physical infrastructures and less involved in digital initiatives, considers the citizens and not-for-profit associations like key players for its success, and a formal, democratic organization of GSCA like a crucial instrument to drive the development of Genova smart city gaining the higher consensus. At present, it is early to assess which is the best solution; but perhaps all of them are the best solution for each city. Indeed, a smart city comprehensive project, involving completely a city and aiming at transforming its profile, needs to be city-specific and harmonized with the culture and the other characteristics of the urban area.

Both Genova and Amsterdam settled a formal body to govern the smart city strategy, in which the Municipality is a key actor, but juridically separated by the association. It is an important choice, in Italy only Genova made it; it shows the intention to give to the smart city an independent life respect to both the politic local Govern and all the private companies.

4.3 Initiatives

To realize the smart city plan, Genova built a portfolio, composed by three types of elements: large EU projects, other funded projects, smart initiatives.

Large EU projects are three and are the projects winning the EU calls for funding smart city projects. Other funded projects are six and they received funds from both international and national govern bodies, for example from MIUR (Minister for Education, University and Research), from different calls respect to smart city topic, but similar in their contents. Smart initiatives are other actions drove by the Municipality of Genova and regarding especially its own organization.

All the projects and initiatives have been analyzed and classified applying the schema already applied for Amsterdam, showed in Fig. 8 and explained in Sect. 3.3. The results of this analysis are showed in Table 5 regarding the 9 large projects and in Table 6 regarding the 51 initiatives.

Analysing Table 5, 8 out of 9 projects are smart and only one is digital. Among the smart projects, only one has a strong role of ICT to support smart actions. Three projects are no-tech: it is because a lot of calls regarding EU projects in smart city topic are focused on design the guidelines, policies, best practices, but also definitions and main contents of a new and immature research field. EU recognizes that to foster a rapid and efficient smart city implementation all over Europe, it is better to pursue a top-down strategy, defining processes and behaviors and spreading them collected in a sort of white book, explaining what and how to do and what not to do, to save time and money and to prevent mistakes.

Genova Smart City presents a lower rate of digital projects, because the main driver of the Genoese strategy has been to adhere to the EU smart city vision, to win the more EU calls, and this vision is mainly technological and focused on CO_2 emission reduction and building efficiency improvement, also through cooling, heating and lighting innovative systems.

For the same reason, Genova presents a lower rate of people involvement respect to Amsterdam. For the first, the strong technical focus of the majority of projects excludes the participation of citizens; moreover, the low rate of digital projects reveals that Genova considers less important in this phase to use ICT to create people networking. Probably it depends also on the lower literacy rate of Genoese citizens, their less daily use of smart devices and the Internet and the lower readiness of Public Administration in supplying digital services.

Generally, we could conclude that Genova choose to apply to EU calls, completely assuming the EU smart city vision, strongly committed in pursuing CO_2

Table 5 EU smart project in Genova

	Project	Description	Type
1	Illuminate	To realize smart illumination in large urban areas to reduce energy consumption	SC
2	ElihMed	Realising innovative existing building refurbishment to improve the energy efficiency; it regards public dwelling	SC
3	R2Cities	To define innovative strategies and solutions to improve energy efficiency in large buildings	SC
4	CELSIUS	Developing pilot project about district heating and cooling systems and energy networks	SC
5	ICITY	Open Platforms implementation to realize public e-services	DC
6	Peripheria	Developing an innovative approach to involve final users and citizens (especially in suburbs) in planning and implementing new products and services. This approach uses especially ICT and Living Lab	NOTECH
7	HARMONISE	To define EU standards and best practices to support security, resilience and sustainability in urban long-term planning	NOTECH
8	Transform	To define a methodology to transform cities in smart cities collecting both theoretical studies about strategic planning and best practices in six EU implementing cities	NOTECH
9	Very School	Realising a heating system in public schools aiming not only at reducing energy consumption and CO_2 emissions, but also at educating children and their parents to a smarter use of energy	SC

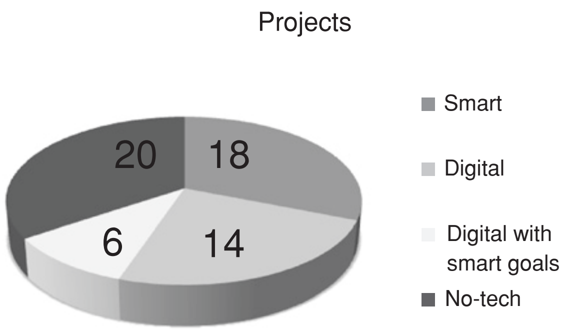


Fig. 11 Smart and/or digital project range in Genova Smart City

reduction in urban areas. However, a stronger focus on the digital side of smart city emerges from the analysis of smart initiatives, showed in Table 6. We can count 10 smart initiatives, 13 digital initiatives, 6 digital initiatives with a strong smart impact and 14 no-tech initiative. These latest mainly regard regulations about the behavior of the Municipality, introducing a smart trend in each act, for example introducing green criteria in procurement, or regard infrastructure initiatives like cycling routes, local public transport, and so on. The projects + initiatives range composition in Genova is showed in Fig. 11.

Table 6 Smart and/or digital project range in Genova smart city

	Title	Description	Type	Spec
1	E3SoHo	To develop an ICT platform to monitor the families' energy consumption in a popular dwelling area, with the aim to extend it all over the city and to educate people in more sustainable behaviors	DC → SC	DATA COMM
2	Diamond Social centre	To develop a social centre built with sustainable criteria and to extent best practices to further similar projects	SC	EFF PEOPLE
3	Molassana civic centre	To develop a civic centre built with sustainable criteria and to extent best practices to further similar projects	SC	EFF PEOPLE
4	Young people centre	To develop a young centre built with sustainable criteria and to extent best practices to further similar projects	SC	EFF PEOPLE
5	Renewable energy plants in civic buildings	To develop a pilot project to collect best practices to convert energy plants in municipal buildings in sustainable plants, based on renewable energy sources	NOTECH	
6	Energy efficiency in public markets	To develop a public market built with sustainable criteria and to extent best practices to further similar projects	SC	EFF
7	SEAP—Action Plan for sustainable energy	The SEAP (Sustainable Energy Action Plan) is the key document signed by the Covenant of Mayors, to guide the city actions to reach its CO_2 reduction target by 2020	NOTECH	
8	Smart Traffic Light	To convert traffic lights in sustainable traffic lights	SC	EFF
9	ELENA European Local Energy Assistance	An EU project to furnish technical support to innovative solutions in cities aiming at reducing the environmental footprint of urban areas	NOTECH	
10	Sun Procurement	To develop rules, administrative and legal instruments, type contracts to produce and distribute solar energy in large co-owner buildings	NOTECH	
11	Servizionline.comune.genova.it	To create a large platform to offer to the citizens e-services, aiming at administrative process efficiency, paper consumption reduction ad mobility reduction	DC	COMM
12	FreewifiGenova	Implementing a lot of public areas in Genova with free wi-fi service	DC	COMM
13	Genova Optical Fibre	To create a proprietary optic fibre to connect all the municipal branches in city	DC	COMM

(continued)

Table 6 (continued)

	Title	Description	Type	Spec
14	FTTH « Fiber to the home »	To offer to citizens broadband services	DC	COMM
15	Municipal Bulding Regulations	To develop a territorial regulation to support a larger use of buildings techniques, to improve building energy efficiency	NOTECH	
16	Smart Energy at Work	To write a handbook to drive best practices regarding energy consumption in workspace	NOTECH	
17	Smart City Management	Organizing a University Master Course for the management and governance of a complex smart city program	NOTECH	
18	Electric Mobility	To create a large urban infrastructure to support private use of electric cars	SC	EFF
19	Infomobility	To create an ICT platform to offer information about the traffic in the urban area in the real time and to reduce traffic and pollution	DC → SC	DATA
20	Mobility Supervisor	To develop an integrated ICT system to collect information about traffic and transport, merging data deriving from different data sources such as sensors, private and public databases, videosystems, and so on	DC	DATA
21	App AMT	To develop a mobile application to supply information about the local public transport systems in Genova and around	DC	DATA
22	Electra—Electric City Transport	To develop an innovative public transport system, based on electric scooter sharing	SC	EFF
23	Car Sharing	To develop an innovative public transport system, based on electric car sharing	SC	EFF
24	Bike Sharing	To develop an innovative public transport system, based on bike sharing	SC	EFF
25	Epistemetec	To realize a digital library to preserve the cultural heritage of some Italian regions	DC	DATA
26	Med-3R	To realize international cooperation between Mediterranean cities to share technical implementations regarding the waste treatment	NOTECH	

(continued)

Table 6 (continued)

	Title	Description	Type	Spec
27	CycleCities	To promote an educational campaign regarding policy-makers, citizens and institutions, about the importance of sustainable transport systems	NOTECH	
28	CATMED	To implement a sustainable district ICT platform, based on sustainability education and citizens involvement, to be gradually applied to all the districts in a smart city	DC	COMM
29	Web Sellers	To realize an ICT platform to sell abroad touristic services in Genova and surroundings	DC	DATA
30	Smart University Energy	To realize a set of sensors and an ICT platform to measure energy consumption and building inefficiency in the University of Genova; these measures will be used to support the smart energy system regarding all the Genoese university	CD → SC	DATA
31	Tecnoedile	To realize a prototype of "near zero energy building", using integrated systems to produce energy from renewable sources	SC	EFF
32	PEAP- Port Energy Plan	To define the energy plan of the Port of Genova, aiming at optimize the energy efficiency	NOTECH	
33	Climate Change! We Change!	This project regards and integrated approach at the problem of reducing the energy consumption in multi-owner building; it involves all the stakeholders: co-owner representatives, energy companies, municipality, etc.	NOTECH	
34	H@H (HEALTH @ HOME)	To offer to elder citizens an ICT application to support medical assistance online	DC	DATA COMM
35	SCOC (SmartCity Operation Security Center	To develop an Open Data platform to integrate heterogeneous information about territorial safety systems	DC → SC	DATA
36	Inset (INTEROPERABLE NATIONAL SYSTEM FOR ETICKETING	To realize an ICT platform to integrate e-ticketing service with municipal policy for tourism in Genova	DC	COMM

(continued)

Table 6 (continued)

	Title	Description	Type	Spec
37	Urbelog	An ICT system to support the efficiency of product delivery in the urban areas	DC → SC	DATA
38	Accessit	To develop an ICT platform to support the design of touristic itineraries in Genova and in the Mediterranean area	DC	DATA
39	Energy Building Business Protocol	To create a Protocol to integrate business with the Municipality of Genova in defining a long term plan about municipal building energy efficiency	NOTECH	
40	Public Transport Business Protocol	To create a Protocol to integrate business with the Municipality of Genova in defining a long term plan about public transport energy efficiency	NOTECH	
41	Genova Smart City Web Site	To create a web site for the Genova Smart City program, able to spread the culture of smarter city among the citizens	DC	COMM
42	Smart Revolution Award	A competition among citizens regarding smart proposal to be submitted to the Municipality	NOTECH	
43	Decision Theatre Partner	To realize an ICT platform to support the strategic planning and governance of long term, integrated smart projects	DC → SC	DATA

The graph shows that the highest number of projects is classified like no-tech. This outline the more comprehensive vision of Genova Smart City. Indeed, not only Genova settled from the beginning a formal association to govern the Smart City initiative, but it gives the highest importance to the context definition. Genova thinks that it is important to define a smart city framework, including governance, processes, best practices, before to implement single initiatives. In this sense, the projects and initiatives are not a sum of independent actions, but a subset of a larger vision including all the smart initiatives in the general framework.

4.4 Analysis

The deep analysis of Genova Smart City case shows to us that different paths could be walked through, to improve the smartness of a city. Genova is an interesting case especially because it demonstrates that each city, even with no experience in smart or digital projects, could become a leader smart city if pursuing a well defined strategy. It shows also that in smart city practices the followers could be better than the first movers, because they will be able to apply the best practices, policies and guidelines developed worldwide to drive the smart city implementation.

Our analysis is useful to outline both the strengths and the weaknesses evidenced in this large smart strategy.

The main strength for Genova Smart City is the key role of the Municipality, able both to start a large implementation of smart actions inside its own organization and to drive the smartness improvement of the whole city area. The settlement of AGSC and the quadruple helix model (even if unconsciously applied) are winning steps towards a comprehensive and shared vision of a smart city capable to sustain and renew its own development over time. The high cooperation between public administration, university and business is the main driver of the future dissemination of smart knowledge.

Another strength is the high international visibility and collaboration and the possibility to collect abroad and to develop smart practices, to be applied in Genova in further projects. Thanks to its nine international projects, Genova participates to a large network of European cities, both large or medium, at different stages in implementing their own smart strategies; this is an inestimable knowledge base.

Finally, Genova has developed a comprehensive vision about the city smartness, regarding not only the technological aspects, but also the regulatory aspects and it has well understood the key role of the Municipality in driving and disseminating smart awareness among companies and citizens.

On the other side, Genova presents also several weaknesses, to be faced not to induce the failure or the low returns of smart initiatives. The more critical weakness is the excessive reliance on the EU funding to implement smart actions; it derives also from the uncritical adhesion to the EU smart city definition, strongly focused on CO_2 emission reduction. This acceptance of the leading role of EU strategy could be a strong obstacle in the future, to develop in Genova its own smart city vision and to replicate best practices, guidelines and innovative technical solution in several smart projects, extending by this way the smartness from one site or areas to several sites and city areas. Surely, the worst obstacle to be overcame is the lack of funding from its own financial resources or the lack of EU funds financing not only pilot projects, but a smart initiative along with its full life cycle.

Another weakness is the low involvement of citizens. It is partially due to the low role of digitalization and smart community development, with a excessive focus on technological aspects. The low digital literacy in Genoese citizens is not a good reason to neglect their digitalization. On the contrary, a stronger effort should be done, to both reduce the digital divide in using digital services and smart devices, and improve digitalization and employees training in planning, using and delivery ICT applications and services.

4.5 Conclusions

The analysis of Genova Smart City projects and initiatives portfolio shows that the profile of these two cities—Amsterdam and Genova—is quite different.

The project portfolio of Genova Smart City contains 8 smart projects, all of them funded by international institutions and especially the EU. The EU vision of a smart city is also the vision assumed by Genova to participate to the EU calls for funding. We could say therefore that the Genova Smart City project portfolio is EU-driven and it reflects the EU smart city idea indeed. Depending on this point of view, Genova is very smart and few digital, very technological and especially based on hard technologies, and less IT based.

The situation is different if we examine the whole portfolio including both large smart projects—funded by EU—and initiatives driven by the Municipality of Genova. In this second case, the portfolio composition is different. Not only we can find several digital initiatives, but also a lot of no-tech small actions, aiming at defining the smart context in the city, regarding a large spectrum of topics. Depending on this point of view, Genova supports a more comprehensive idea of smart city, not only based on environmental urban footprint and sustainability, but more generally on the improvement of the quality of life in the urban area.

Another interesting aspect emerging from the Genova Smart City experience is the more integrated view of smart initiatives and projects. It emerges not only from the role of AGSC in governing the whole process of improving the smartness of Genova, coordinating public and private institutions, business and research bodies, not-for-profit organizations and citizens. It emerges also from the trial to put all the efforts into a unique framework able to measure also the obtained results from not a single project, but the project portfolio. For example, Genova links the smart project portfolio to the SEAP—Sustainaible Energy Action Plan signed by the Covenant of Mayors. Covenant of Mayors is the mainstream European movement involving local and regional authorities, voluntarily committing to increasing energy efficiency and use of renewable energy sources on their territories. By signing the SEAP, the adherent mayors aim to meet and exceed the EU 20 % CO_2 reduction objective by 2020.

Linking smart city projects and SEAP means mainly two thinks:

1. to consider CO_2 one of the most important smart city goals;
2. to collect all the smart projects into a unique basket of actions intended to work together to reach a shared objective, that is, to view all the projects in a comprehensive manner.

However, it means also that it is easier to link smart projects to environmental goals such as energy consumption or pollution, as these goals are measurable; too difficult is to link smart projects to quality of life level, as this goal is fuzzy and often a direct causal impact of smart projects on the quality of life is not granted. Therefore, there is a gap between the ideal definition of smart city assumed by Genova in describing its own aims, as reported in Sect. 4.1; and the pursuing of more restricted goals, such as CO_2, more reflecting the EU definition of a smart city.

It impacts also on the smart portfolio composition; on one side the smart projects are too focused on few topics, especially energy consumption and pollution reduction; on the other, the initiatives portfolio is very broad, potentially including each public or private initiative aiming at an improvement in the quality of life in the urban space.

Therefore, Genova swings between three different smart city ideas:

1. The “close” idea, strictly focused on a smart city definition regarding only the environmental footprint of cities, and consequently it includes mainly the projects and initiatives aiming at reducing pollution and CO_2 emissions and reducing the energy consumption: it reflects the EU smart city idea;
2. The “medium” idea, including both smart city as conceived above and digital city, that is, it joins both the use of hard technologies and ICT, the first to reduce the infrastructure impact on the environment and the latter to connect people through open data, information sharing, broadband connections and digital e-service: it reflects the more accepted smart city idea, both in the academic world and by companies;
3. The “large” idea, including into the smart city definition all the initiatives aiming at improving the quality of life in the urban area, both technology-based or not; this broad definition is quite fuzzy and it makes difficult to really understand what a smart city is.

For these reason, the analysis of Genova Smart City interesting case, even if helps us to enlarge our understanding of the contents and scope of a smart city strategy, doesn’t help us to define a smart city thanks to the empirical analysis.

In the further and last paragraph these two empirical cases—Amsterdam and Genova—will be compared each other to extract a smart city evidence from the overlapping of these two leader experiences in Europe.

5 Conclusion, Lessons Learned and Further Works

The analysis of these two case studies—Amsterdam and Genova—has been carried out with the aim to compare smart and digital city each other and to understand which are similarities and differences between these two urban strategies. From the beginning, the hypothesis under our survey has been that, even if they are often overlapped or confused, smart city and digital city are not the same thing and cities implementing smart city programs implement indeed a mix of smart and digital actions. Finally, after our study, we can say that our hypothesis have been confirmed; even if smart city and digital city have a lot of common aspects, they should not be confused as they need different strategies to be successfully implemented. The outcomes of our research are shown item by item below.

The historical analysis of both the literature and the business cases shows that digital city has born before smart city; like the Amsterdam case study demonstrates, digital city has been developing during several years—and till now—like an instrument to empower citizens respect to government, political issues and the public administration. It establishes itself along with the diffusion of the Internet among people, business and public administration.

Digital city is strongly based upon the ICT and especially the Internet, and therefore the communicational content is its more important aspect; other main aspects of a digital city implementation are data availability, information diffusion and e-services. It emerges from both Amsterdam and Genova that digital actions are mainly focused on improving the relationships with public bodies by delivering digital services or using the web site to spread information and create a more direct relationship with citizens. It means that a digital city strategy somewhat pursues the same goals of e-Government, but with a specific accent on the urban life. Thanks to this strict link with the ICT, the digital city perimeter and boundaries are well defined and its contents are easy to qualify. Also the required infrastructures are well identified, based on broadband connection, open data and web-based public services.

Citizens are actively involved not only in digital city implementation, but especially in the daily use of digital facilities; therefore the role of citizens is not only to receive or to enjoy the results and benefits of a digital city strategy, but to participate to its concrete functioning; without the active, daily use by citizens, a digital city cannot fully exploit its role and its success is limited by the insufficient returns obtained from the digital investments. It means that a digital urban strategy requires a high attention to the digital education of citizens and a strong contrast against digital divide like one of the most important barriers to a digital city full success. For this reason the educational level of citizens in using smart devices or ICT is one of the main drivers for the successful implementation of a digital city plan; as the digital culture has ever been better in The Nederland respect to Italy, no surprise that Amsterdam is a pioneer city in implementing such strategy, whereas Genova has been starting to implement a smart city strategy before.

Smart city has born several years after respect to digital city and had a boom in 2009 after the EU strongly committed to support and fund smart initiatives in European cities, aiming to reduce CO_2 emissions and to govern energy consumption, waste treatment and building efficiency. It appears clearly in Amsterdam but especially in Genova, where all the big smart projects are funded by the EU and the smart strategy has been planned just to catch the opportunity of EU calls. Smart city is nowadays a fuzzy idea, but its original core focus is on environmental impact of urban areas and activities. These topics emerge from the urbanization happened during the latest twenty years and the increasing problems it produces, like pollution, traffic congestion, high dwelling price, inequality and poverty. These goals are easily to individuate in both Amsterdam and Genova smart initiative portfolio.

Also the smart city considers technology like a core component, but in this case we haven't only one technology, like ICT in digital city, but a large set of innovative technologies like for example smart grid, renewable energy sources, new types of fuel for transports, new materials for building, and so on. Respect to digital city, we could say that smart city is based on hard technologies, a digital city on soft technologies.

The role of citizens in smart city is not necessarily active; for example, to reduce pollution by electric buses is a choice made by the local transport companies, and the citizens are the beneficiaries of this urban transport policy. They gain the benefits, but they are not actively involved. Obviously there are also smart actions requiring the citizens commitment, but it is not ever necessary in a smart city strategy, unlike in a digital city. Respect to this aspect, also a different orientation by a specific city can deeply modify the involvement of citizens in smart plan. For example, analysing Amsterdam and Genova smart initiative portfolio, we discover that in Amsterdam the involvement of citizens generally plays a more important role than in Genova, where the technical content of several smart projects prevails respect to the human side.

Despite these differences, smart city and digital city are not completely separable. As we have seen in examining Amsterdam and Genova, both these cities are developing their urban strategy mixing smart and digital actions. The main reason is that both smart and digital strategies have the same final goal, that is, to improve the quality of life and the citizens' satisfaction in their city. Smart and digital initiatives are joined in the strategic vision of local governments and these development paths are often defined in the same long term plan. As smart city is a more recent idea, it tends to absorb also digital city, combining both these strategies in a mixed, city-specific roadmap.

One of the negative effect to include digital into smart and to enlarge the smart city scope is that smart city has more fuzzy perimeter and boundaries respect to digital city. The main reason is that smart city tends to include all the initiatives aiming at improving the quality of life, that is, digital initiatives, but also green actions, inclusive actions, cultural programs and so on. For example, Genova defines its own urban city plan like a smart action, because it tries to incorporate also some trends like to reserve areas for parks or green areas and so on.

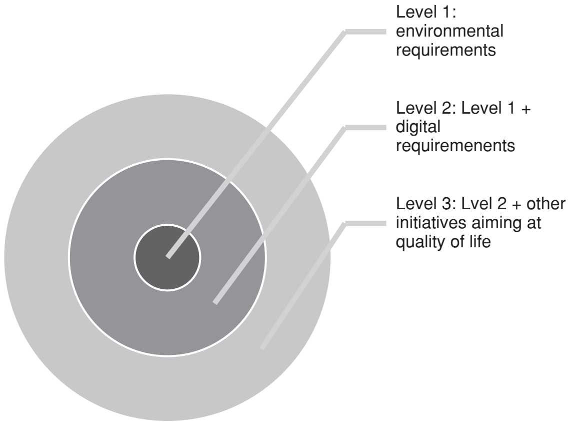


Fig. 12 A three level smart city definition

Moreover, digital city is based on only a technology, that is, ICT, whereas smart city is based on several innovative technologies, but sometimes it includes also initiatives without technological basis: for example, to educate parents to accompany their sons at school by foot instead that by car is a smart initiatives (in a large sense) because it aims at reducing pollution and CO_2 emissions, but using no technologies. In both Amsterdam and Genova initiative portfolio there is a certain percentage of no-tech projects indeed. Therefore, to define what a smart city is becomes more and more difficult.

The case studies show to us that the concept of smart city has indeed different contents, depending on the meaning a city attributes to it. Both Amsterdam and Genova merge in a large smart city strategy a large set of initiatives, contributing to the quality of life in their urban area through different aims. To summarize the evidences emerging from both the literature review and the case studies about the multi-level definition of a smart city, we can define a three-level smart city concept (Fig. 12):

- the smaller concept is represented by the actions, initiatives and strategies aiming at improving the quality of life in city, through the reduction of its environmental footprint, especially using innovative technologies applied to building efficiency, energy production and consumption, transport systems efficiency;
- the intermediate concept merges the smaller one—with environmental goals—with the digital city, that is, the digitalization of data, information and services, and the empowerment of citizens' communication with government and other public bodies;
- the larger concept adds to the intermediate one other initiatives, aiming at improving the quality of life in city, but not based on ICT or hard technologies; for example, green, inclusive, cultural initiatives, and so on; these latest actions are the more city-specific respect to the strictly smart and digital actions, that are more similar in several cities.

Table 7 Comparing smart city and digital city

	Digital city	Smart city
Year	Nineties	Boom in 2009
Technology	ICT	Hard technologies, especially applied at energy production, distribution and consumption
Focus	Information and communication by digital devices	Environmental impact of urban areas and activities
Process	Bottom-up	Top-down
Citizens	Active involvement	Active involvement not required, it depends on both the city vision and on the specific smart actions
Governance structure	No formal governance structure	Different governance structures, driven by public bodies and especially municipalities

A comprehensive comparison between smart city and digital city and their different characteristics are exposed in Table 7. Even if smart city is absorbing digital city, these two different urban strategies need different processes and practices to successfully be implemented and to gain the best results from them. For this reason, even if they are concretely merged into a unique city plan, they should be implemented taking into consideration their different nature.

One of the effects of fuzzier and larger boundaries is that the smart city output and impacts are more difficult to measure, the larger and heterogeneous its perimeter its. Indeed, it is quite easy to link and measure the effects of smart actions impacting on environmental aspects such as energy use and CO_2 emission reduction or cleaner energy production by renewable sources. But more difficult is to measure the impact of digital policies; indeed, it is necessary not to confuse the readiness of a policy with its impact. It is easy to measure the digital infrastructure or facilities realised by a city, measuring the broadband extensions or the number of citizens using smart devices or e-services. But more difficult is to evaluate the benefits or the public value produced by an integrated smart and digital strategy; these measures are only a proxy of the strategy effects. Both smart and digital city, in a large sense, present a high difficult to evaluate the returns they produce. It is an important barrier to smart and digital initiative implementation, because both of them often require a large amount of public investment and therefore also the need to justify the expenses and to demonstrate the reached results.

More generally, the large smart city scope negatively impacts on all the life cycle and governance framework of this urban strategy. Indeed, with very heterogeneous aims, technologies, stakeholders, it is difficult to support investment decisions, funding of projects, priorities demonstration and expenses justification, outputs measurement and performance evaluation. For this reason, to find a sound and shared smart city definition, with clear boundaries and delimited goals, it is necessary to better support the further smart city planning and implementation. As seen in our two case studies, at present all the cities, also the pioneer ones, are at an early stage in smart city development; nowadays all the projects have mainly the role to experiment initiatives and to collect best practices, but in the future these projects should become daily work to improve the quality of life in cities. Therefore, to be able to govern the smart city will be the most important weapon to reach substantial results. Further works will therefore use this study about the contents of smart and digital city to support the definition of a governance framework for their effective realization.

References

1. Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart cities in Europe. *Journal of Urban Technology, 18*(2), 65–82.
2. Anthopoulos, L., & Fitsilis, P. (2010). From digital to ubiquitous cities: Defining a common architecture for urban development. In *2010 Sixth International Conference on Intelligent Environments* (*IE*). IEEE.
3. Dameri, R. P. (2013). Searching for Smart City definition: a comprehensive proposal. *International Journal of Computers and Technology, 11*(5), 2544–2551.
4. Shin, Y., & Shin, D.-H. (2012). Community informatics and the new urbanism: incorporating information and communication technologies into planning integrated urban communities. *Journal of Urban Technology, 19*(1), 23–42.
5. Morganti, L., & Donders, K. (2013). A digital agenda in search of evidence. *info 16*(1), 1.
6. Kim, J., & Steenkamp A. L. (2013). Analysis of smart city models and the Four-Foci taxonomy for smart city design. *The Visibility of Research,*(2013), 637.
7. Dameri, R. P. (2012). Defining an evaluation framework for digital cities implementation. In *2012 International Conference on Information Society* (*i-Society*). IEEE.
8. Dameri, R. P., & Ricciardi, F. (2014). Using social networks in smart city: organizational challenges, synergies and benefits. In *Proceedings of European Conference on Social Media*, Brighton.
9. Dameri, R. P., D'Auria B., & Ricciardi F. (2014). Knowledge and intellectual capital in smart city. In *Proceedings of European Conference on Knowledge Management, Santarem Portugal*.
10. Bosker, M., Buringh, E., & van Zanden, J. L. (2013). From Baghdad to London: unraveling urban development in Europe, the Middle East, and North Africa, 800–1800. *Review of Economics and Statistics, 95*(4), 1418–1437.
11. Zhang, L. Y. (2013). City development strategies and the transition towards a green urban economy. In *The Economy of Green Cities* (pp. 231–240). Netherlands: Springer.
12. Stigt, R., Driessen, P. P. J., & Spit, T. J. M. (2013). Compact city development and the challenge of environmental policy integration: a multi-level governance perspective. *Environmental Policy and Governance, 23*(4), 221–233.
13. Bowerman, B., et al. (2000). The vision of a smart city. In *2nd International Life Extension Technology Workshop*, Paris.
14. Paskaleva, K. A. (2009). Enabling the smart city: the progress of city e-governance in Europe. *International Journal of Innovation and Regional Development, 1*(4), 405–422.
15. Ramaswami, A., et al. (2012). A social-ecological-infrastructural systems framework for interdisciplinary study of sustainable city systems. *Journal of Industrial Ecology, 16*(6), 801–813.
16. Nam, T., & Pardo, T. A. (2011). Smart city as urban innovation: Focusing on management, policy, and context. *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance*. ACM.

17. Ricciardi, F., Rossignoli, C., & De Marco, M. (2013). Participatory networks for place safety and livability: organisational success factors. *International Journal of Networking and Virtual Organisations, 13*(1), 42–65.
18. Yuan, Y. M., et al. (2012). Architecture and data vitalization of smart city. *Advanced Materials Research, 403*, 2564–2568.
19. Lombardi, P., et al. (2012). Modelling the smart city performance. *Innovation: The European Journal of Social Science Research, 25*(2), 137–149.
20. Cocchia, A. (2013). Smart and digital city: a systematic literature review. In: NOSTRO LIBRO.
21. Jenks, M., & Dempsey, N. (Eds.). (2005). *Future forms and design for sustainable cities*. London: Routledge.
22. Komninos, N. (2006). The architecture of intelligent clities: integrating human, collective and artificial intelligence to enhance knowledge and innovation. In *2nd IET International Conference on Intelligent Environments*, *IE 06*. (Vol. 1). IET.
23. Ishida, T. (2002). Digital city kyoto. *Communications of the ACM, 45*(7), 76–81.
24. Komninos, N. (2008). *Intelligent cities and globalisation of innovation networks*. London: Routledge.
25. Camagni, R., Capello, R., & Nijkamp, P. (1998). Towards sustainable city policy: an economy-environment technology nexus. *Ecological Economics, 24*(1), 103–118.
26. Sorrentino, M., Niehaves, B. (2010, January 5–8). Intermediaries in E-Inclusion: a literature review. *Paper presented at the 43rd Hawaii International Conference on System Sciences* (HICSS-43), Kauai, Hawaii. IEEE.
27. McCarthy, S. (2007). Planning for health and wellbeing: city of greater Dandenong wonders of Dandenong and the walking revolution. *Planning News, 33*(5), 14.
28. OECD. (2013). *Functional urban areas in OECD countries*. Paris.
29. *Il riordino delle province e delle città metropolitane, Camera dei deputati-XVI Legislatura—Dossier di documentazione.*
30. Guerrini, A., Romano, G., & Campedeli, B. (2013). Economies of scale, scope, and density in the Italian water sector: a two-stage data envelopment analysis approach. *Water Resource Management, 27*(1), 4559–4578.
31. Guerrini, A., Martini, M., & Campedelli, B. (2013). Measuring the efficiency of the Italian construction industry. *International Journal of Business Performance Management, 14*, 307–325.
32. Zardini, A., Rossignoli, C., Mola, L., & De Marco, M. (2014). Developing municipal e-Government in Italy: the city of Alfa case. In M. Snene & M. Leonard (Eds.), *Exploring Services Science* (Vol. 160, pp. 124–137)., Lecture Notes in Business Information Processing Berlin: Springer.
33. Caragliu, A., Del Bo, C., & Nijkamp, P. (2009). Smart cities in Europe. 0048.
34. Etzkowitz, H., Loet, L. (1997). The dynamics of innovation: from National Systems and Mode 2 to a Triple Helix of university–industry–government relations. *Research policy*, *29.2*(2000), 109–123.

Smart, Smarter, Smartest: Redefining Our Cities

Claire Thorne and Catherine Griffiths

Abstract The UK Government, like many national governments, has the creation of Smart Cities high on its agenda. This interest is triggered by the promise that as a significant part of its national armoury, smart cities can drive economic leverage and possibly even deliver economic salvation. This aspiration has fuelled a growing global competition to attract entrepreneurs, talented people, and investment. The race to create smart cities is on. This chapter will describe the components that make a city smart, and examine the emerging 'need' to create a smart London. While only governments can create policy that enables scale to be achieved, policy so often only follows where the green shoots have already emerged. This is the situation in the UK, in London, and currently policy is not keeping pace with the vibrancy of the initiatives. The evidence of so many successful initiatives and national achievements shows (e.g. the transformation of Singapore to an Intelligent Island,[1] the successful delivery of the London Olympic Games,[2] the aim of President Kennedy to put a man on the moon[3]) that much is driven by vision as well as time, skills and funding. It is this lack of a clear vision that if addressed would help unlock new potential and reinvigorate many existing and older investments. Policy and initiatives could then work cohesively to help deliver what is currently so often only empty rhetoric. This chapter is a clarion call for a vision that focuses

[1] Beyond 2000: A Source Book for Major Projects. Major Projects Association, Templeton College Oxford.

[2] Research by Prof Andrew Davies.

[3] http://history.nasa.gov/moondec.html

C. Thorne
Digital City Exchange, Imperial College London, London, UK

C. Griffiths (✉)
Data Science Institute, Imperial College London, London, UK
e-mail: c.griffiths@imperial.ac.uk

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_4, © Springer International Publishing Switzerland 2014

on delivery of pervasive integration so that peak congestion is removed from urban systems and the enhanced quality of life in cities can bring benefits to all. Lastly, it suggests 'smartness' may not come purely from technological solutions after all, but from the mechanisms used to engage and deliver a new kind of city.

Keywords Smart city • Digital revolution • London smart city • Economic leverage • Data science

1 A Smart City Landscape: The Digital Revolution

The 21st Century is arguably producing a third industrial revolution.[4] The first was the mechanisation of textile manufacturing in the 18th Century, the second was Henry Ford and assembly line mass production in the 20th Century, and the third is the digitization of manufacturing currently in progress. As before, the third industrial revolution is opening up huge areas for innovation, producing seismic changes in the development time, associated costs, and the applications of new technologies. The new world of big and open data is being driven by and simultaneously driving this innovation. Technologies that create and collect data, that manage data, and that enable us to extract value from data, are all being developed and adopted at a rapid pace. Outputs include new platforms, products and processes such as the Internet of Things, next generation sensors, social media, (almost) ubiquitous mobile connectivity, the cloud, along with visualisation, personalisation and miniaturisation. With the advent of these tools, the possibilities for making better use of resources and increasing standards of living in cities are being tantalisingly glimpsed.

In parallel another factor driving the need for innovations in cities is the huge increase in urban population, which is predicted to continue rising. Globally, for the first time in history more people live in urban centres than live in the countryside and it is predicted that by 2040 two thirds of the global population will have moved to cities.[5] Where there are existing historic cities these added numbers are increasing the pressure on already constrained logistics, resources and infrastructures. Addressing these challenges of how to make old systems work effortlessly with new technologies, while designing cities that can offer enhanced quality of life for citizens, requires entirely new thinking and vision, capitalising on the outputs and opportunities offered by the digital revolution.

A 'bigger, faster, easier' city of the future may only be realised however through the integration of platforms, products and processes (i.e. the 'digital') with city infrastructure and the urban space, such as buildings, parks and roads (i.e. the 'physical'). In this scenario, real-time, personalised information could flow within and across cities, and the peaks increasingly seen in city systems—in consumption,

[4] The Economist. 21st April 2012.

[5] Forum for the Future: Report: Megacities on the Move, by Ivana Gazibara.

capacity, congestion, and inefficiency—could be diffused. Through access to, and an understanding of, real-time information flows, and through seamless data integration across networks, dynamic city peaks such as traffic jams, hospital waiting times,[6] surges in utility uses and power outages could all be better managed, in a coordinated 'systems of systems' way. As a result all citizens, whether residents, commuters or tourists, would stand to benefit.

2 A Smart City: Pervasive Integration

'Smart' is often used interchangeably and frequently naively, with 'connected' (i.e. to the Internet) and 'interconnected'. It is used to identify and badge something as meaning leading edge and with seamless communications. 'Smart' is already a global brand. It is routinely being misapplied, inflating the real level of technicality and complexity of a (sometimes digital) system: Examples such as deterrents used to protect security vans being labelled—'smart water'; Teachers delivering lessons to pupils with next-generation blackboards—'smart boards'; Miniature vehicles buzzing around the road network—'smart cars'. Governments are proud of providing 'smart villages' (e.g. in Egypt),[7] 'smart cities' (Rio),[8] and smart initiatives are pushed and promoted as though they are innovative and new. But are all these really smart? In reality, there is little difference between these and older, now mainstream, technologies facilitating communications such as the basic telephone network.

The words are familiar, and increasingly used by public officials and industry consultants to establish credibility and verify their status, but exactly what is smart, where is the vision, and what is the reality? Smartness must be justified by real-world need in order to avoid switching off potential adopters too early otherwise there is a risk that the full promise of this development will not be realised. In a smart city engagement is key, and language matters.

Sensors, applications and devices requiring connectivity should in fact be categorised within 'digital' because in isolation they do not constitute 'smart'. Technologies may be smaller, more sophisticated and more portable than ever before, but often many innovations are simply about revised presentation, rather than new content, application or service. What would really innovate and take these 'digital' tools beyond, into the world of 'smart', is not enhanced connectivity, it is: seamless interconnectivity and interactivity (to and with other devices or systems) and real-time responsiveness (to events and preferences). A new language is therefore required to accurately describe and fully capture what cities need. In summary, the meaning of smart has been hijacked, and its application is limited. Instead, pervasive integration of digital devices and platforms across

[6] http://www.theguardian.com/society/2013/aug/15/nhs-hospital-waiting-lists

[7] The Smart Village Company, Egypt http://www.smart-villages.com/en/page/page/147.

[8] How to Transform a City. IBM Smarter Cities White Paper, March 2012.

city infrastructures and resources with real-time data streams, which people can engage with, should be the smart and more comprehensive goal. This involves the use of data to better understand and inform, to change behaviours (at the micro and macro level), to manage and control more efficiently, and to respond in real-time. With such innovation, major shifts in the (individual and shared) roles and responsibilities of citizens, business and governments can be anticipated.

The Almere Smart Society[9] is an example of a wider pervasive integration initiative to build a smart city. Amsterdam is backing a project working with a consortium of Cisco, IBM, Liander, Living PlanIT and Philips to make smart a reality by focusing on pervasive integration. Their vision is to create a ‘smart society’, along with the Almere Economic Development Board, through the realisation of an ICT facility to promote smarter deployment of ICT, people and resources, and more efficient urban management and innovation. The ultimate aim is economic growth, strong social cohesion and sustainable development. They have focused the strategic plan around five thematic areas:

- Living
- Working
- Mobility
- Public Facilities
- Open Data

Each theme is being developed in close collaboration with all the others. Encouragingly, integration is an aim and part of the project plan from the outset. The advantage of Almere, a new suburb of East Amsterdam, is being a city closely associated with new development. The integration of old and new systems, services and infrastructure in retrofitting large, complex cities such as London is undoubtedly more challenging.

Establishing a planned structure for cities, that may also apply to older cities, has been studied and one proposal is to work along the following six identified (and ranked) main axes[10]:

- Smart economy
- Smart mobility
- Smart environment
- Smart living
- Smart governance
- Smart people

These six axes connect with traditional regional theories of urban growth and development. In particular, the axes are based respectively on theories of regional

9 http://amsterdamsmartcity.com/projects/detail/id/30/slug/almere-smart-society

10 Giffinger, Rudolf; Christian Fertner, Hans Kramar, Robert Kalasek, Nataša Pichler-Milanovic, Evert Meijers (2007). “Smart cities—Ranking of European medium-sized cities”. http://www.smart-cities.eu/. Vienna: Centre of Regional Science. Retrieved.

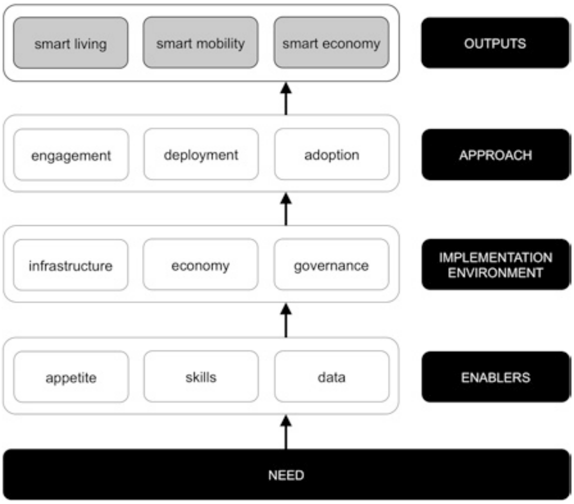


Fig. 1 A framework for defining and realising pervasive integration in cities

competitiveness, transport and ICT economics, natural resources, human and social capital, quality of life, and participation of citizens in the governance of cities.

The latter—'smart people'—is a crucial element and too easily played down by those operating in the smart cities space. Data scientists and engineers have an invaluable opportunity to learn from (and collaborate with) established disciplines that are traditionally focused on human factors, such as medicine, design, and architecture.

The six smart cities axes are easy to grasp and indeed acknowledge the core elements to build capacity for the 'what' and the 'who'. However, this framework does not venture to describe the 'how'.

3 Making a City Smart

Figure 1 presents a revised smart city framework—an attempt to map the constituents and provide the 'how'.

This revised framework is comprised of five descriptors, which offer a way of realising pervasive integration: Need; Enablers; Implementation Environment; Approach; and Outputs. It is important to note that all of the ingredients (i.e. skills, appetite, data) within each of these descriptors are required: Each component is integral, and the approach needs to be holistic.

- **Need**—Driven by growth in urban centres and technical innovation
- **Enablers**
  - *Skills*—Fostering a pool of specialist talent.
  - *Appetite*—Preparedness and willingness of citizens to adopt new technologies and smart interventions. This requires a digitally literate and engaged population.
  - *Data*—Tailored and personalised information streams, accessible to individuals, business and government, in real-time.

- **Implementation environment**
  - *Infrastructure*
  - *Economy*
  - *Governance*
- **Approach**
  - *Engagement*
  - *Implementation/deployment*
  - *Adoption*
- **Outputs**—The core principles are based around facilitating *quality of service*, *quality of life*, and *engagement*.
  - *Smart living*—Including smart health, e.g. How can a smart London enable me to access quality healthcare services locally and timely? How can a smart London enable me to foster (or belong to) a community?
  - *Smart mobility*—Encompassing issues of both transport and mobility, e.g. How can a smart London enable me to reduce my commuting time and/or cost?
  - *Smart economy*—At the least, this alone should be a motivating factor for a renewed and cohesive UK 'smart city' vision and implementation policy.

It is clear that some (but not all) of the smart city ingredients outlined in this framework are already coming together in smart initiatives around the world, albeit in a piecemeal way. In the UK currently there are sprinklings of so-called smart initiatives. However, it is possible that these will fizzle out; their contributions lost, and much well-intentioned investment will be wasted longer term. This is because although a cohesive vision and strategy at a regional or national level is expected to emerge, with a focus on translation, scalability and sustainability, it is yet to be clearly defined. There is as yet no clearly articulated common goal.

4 A Smart City Standard?

It is argued here that taking stock of what London has already achieved, and the expertise it is developing, can help define this much needed common goal and aspiration. Assuming a vision and strategy for London has been announced and a smart London is already emerging: then what? Figure 1 may assist but a discussion around how best to capture the value and the learning, and how to capitalise on the momentum, is also needed. In this case, there are key implementation questions to consider around translation, scalability and sustainability.

As London mostly requires 'retrofit' solutions to age-old infrastructure, transferring any unique 'smartness' to other UK cities and beyond is unlikely to be straightforward. With each city in the world varying in age, design and layout, regulations, existing or planned infrastructure and services, and geography, the smart solutions designed in response to any particular city are, by definition, unique. This means that translating or transferring any technologies or interventions mapped in a like-for-like fashion is unfeasible and most probably futile. In addition, each city should question whether standard solutions are the aspiration, however plausible. How appropriate is it to create a network of 'identic-cities' globally? Should we consider the preservation of other, non-technical attributes of a city, such as its distinct 'personality' and atmosphere, amid the ubiquity of universal devices, platforms and data?

While the specific solutions themselves, the new technologies and services, are not necessarily exportable and applicable to cities elsewhere, the knowledge economy created along the way may well be. This is the case, for example, of Aberdeen. Due to the challenging and unique environment of the North Sea, Aberdeen has amassed significant technical knowledge and industrial capability in deep-sea oil exploration.[11] This has made Aberdeen a world centre and the applications of, and demand for, complex innovative solutions and expertise now reach far beyond the UK.

5 London: The State of the (Sm)Art

How is London performing with respect to the two driving descriptors (enablers and implementation environment) of this new framework for pervasive integration (outlined in Fig. 1)? How is a smarter London realising its promise of better use of resources, more effective investment, increasing standards of living, and integrated services?

- **Enablers**
  - *Skills*—The UK Government, through its UK Trade and Investment arm, is helping to promote intelligent and green buildings, smart grids, the Internet of Things, sustainable urban transportation technology and networks, and telemedicine—it is focusing on the UK's ICT and smart technology and environmental expertise and how this can underpin the creation of low carbon future cities. This is promoting *national* capability—not just city capability.

    Higher Education Institutions are responding to the need with 'digital' and 'smart' educational and research offerings. For example, Imperial College London has launched an Executive Education[12] course on 'smart cities', in collaboration with Arup,[13] and hosts the Digital City Exchange[14] smart cities research project, the Digital Economy Lab,[15] and the Intel Collaborative

[11] http://www.bbc.co.uk/news/business-23490586

[12] www.imperial.ac.uk/business-school/executive-education

[13] http://www.arup.com

[14] www.imperial.ac.uk/dce

[15] http://www3.imperial.ac.uk/digital-economy-lab

Research Institute for Sustainable Connected Cities[16] (co-hosted with UCL). The LSE has established LSE Cities[17]—an international research and teaching centre, focusing on how the design of cities impacts on society, culture and the environment. In addition, London has recently seen the emergence of technology clusters, such as Tech City[18] in the East.

Collectively, this is encouraging, and builds on the existing skills base and experience within London, which has been developed through its historical large-scale technology and infrastructure projects (e.g. through Crossrail,[19] the mass deployment and adoption of the Oyster Card system, the delivery of the London 2012 Olympic Games, and the introduction of the congestion zone[20]). Through these, some of the *individual* issues common to smart cities—such as high demand on existing public services and networks, concerning CO_2 emission levels, and public engagement and community buy-in—have already been partially targeted. Now, there is an opportunity for London to take this learning and apply it to address similar challenges in a more holistic way, and on a larger scale across the country.

– *Appetite*—The adoption of connected devices (such as 'smart phones') has never been so high. Nevertheless there remains a real educational need and 'marketing' exercise to inform the public about pervasive integration and to highlight the benefits of London's smarter future. Most citizens may still view 'smart London' as—at best—irrelevant, or a waste of public funds because the real application has not been effectively conveyed.
– *Data*—In order to realise pervasive integration, access to real (not simulated) real-time (not historic) data is essential. Encouragingly, initiatives like the London DataStore,[21] the Open Data Institute,[22] and the European iCity Project[23] are making city data streams available for publication and discovery. The challenge now is in establishing a similar innovation ecosystem in the private sector domain. How can issues of ownership, regulation, IP, competition, legislation, security, privacy and value for privately generated/owned data sets be handled? And what happens when public and private data sets are combined?

- **Implementation environment**

  – *Infrastructure*—Current and planned investment in London's infrastructure is also encouraging. However, London needs to ensure future investment is

16 http://www.cities.io
17 www.lse.ac.uk/LSECities
18 http://techcity.io
19 http://www.crossrail.co.uk
20 http://www.tfl.gov.uk/roadusers/congestioncharging/
21 http://data.london.gov.uk
22 http://www.theodi.org
23 http://www.icityproject.com

really fit-for-purpose by meeting the growing projected demand. Meanwhile the predicted levels of adoption from past investments have not yet been fully realised. For instance, not all households and individuals in UK cities can afford to, or wish to be digitally connected and Internet access still remains out of reach for some socio-economic groups. Multiple blackspots are still to be addressed around the country and the government's rollout of broadband is 2 years behind schedule.[24]

– *Economy*—There is evidence that the UK is grasping some of the opportunities offered by the new digital revolution and supporting industrial activism to become more engaged and in some areas integrated. Over the financial year 2010–2011, the UK attracted one third of Foreign Direct Investments in Europe within the software sector, 129 of the 392 projects, and of these 70 were located in London. Private investment and venture capitalists invested £453 million in 60 technology companies in London during 2010 making it the most attractive region in the UK for private investment.[25]
– *Governance*—The Digital by Default government policy, the Government Digital Strategy[26] to transform the delivery of UK public digital services with a core principle of user centricity (via the Government Digital Service[27] team), the Race Online initiative (now led by Go ON UK[28] since April 2012[29]), and the formation of the Smart London Board and their published plan[30] are extremely encouraging. Initiatives like these have undoubtedly helped pave the way for future mass-scale adoption of other, smarter interventions. However, there are many issues still to be resolved, for example around the ownership and value of data.

With the diversity and momentum of current smart London initiatives, combined with this new framework for pervasive integration, there is a prime opportunity to speculate on what one/more 'end products' might look like. One version may be an accessible and truly inclusive London where physical and digital solutions enable wheelchair users to travel with unparalleled mobility; or where physical and digital provisions make walking a valid alternative to the traditional (bus, tube, tram and car) modes of transport.

---

24 http://www.bbc.co.uk/news/technology-23173157

25 London's Digital Economy, GLA Intelligence, by Margarethe Theseira.

26 http://publications.cabinetoffice.gov.uk/digital/strategy

27 http://digital.cabinetoffice.gov.uk/about

28 http://www.go-on.co.uk/about-us/our-mission

29 http://digital.cabinetoffice.gov.uk/2012/04/24/race-online-2012-hands-the-baton-to-go-on-uk/

30 http://www.london.gov.uk/media/mayor-press-releases/2013/03/mayor-announces-smart-london-board-to-realise-london-s-ambition

6 London Exemplar: Digital City Exchange—A Systems of Systems Approach

The Digital City Exchange[31] is an interdisciplinary research project funded by Research Councils UK. It focuses on integrating and repurposing real-time, cross-sectoral (transport, energy, water and waste) city data sets to enable business model innovation and to transform the planning, management and use of city services and resources.

The Digital City Exchange is a novel platform, aiming to allow individuals and organisations to combine, trade and exchange city data sets. Added value comes from the associated analytics; uniquely, the platform will facilitate the trade and exchange of real-time predictive models.

Through connecting citizens, business and government to real-time intelligence and enabling ‘smart’ decision-making, almost every aspect of our everyday lives could be revolutionised, leading to:

- Better monitoring and controlling of city services and networks to reduce emissions and waste;
- Improved health and well-being for the population;
- Increased public service productivity and quality;
- The creation of new business models, new data-centric industries, and jobs;
- Improved quality of life in cities.

Research commenced activity in September 2011 and by the end of the 5-year project, new technologies and services are expected to have been deployed and trialled in test-beds across London.

7 Summary

This chapter has argued for a larger vision to be articulated so that the existing sprinkling of initiatives and investments can be maximised by being part of a wider common aspiration. It has also stated that while smart is part of this, it is only part. What is needed is the aim of pervasive integration through the better collection, management and extraction of data.

It has emphasised the need for vision and focus. While London is being seen as a hub that might provide the flywheel for innovation in products, technologies and services, it does not yet have an overarching strategy.

This chapter suggests that there are two key areas to incorporate into this strategy:

- **A cross-disciplinary and cross-sectoral approach**

The real opportunity areas for smarter cities lie at the boundaries where the traditional research disciplines and industry sectors intersect. At these, there are opportunities to

[31] Digital City Exchange http://www.imperial.ac.uk/dce.

foster news types of collaborations, and devise new funding models. Hence, Higher Education Institutions, government departments, the London Boroughs and industry must all address the key question: Is London equipped for this new way of working?

Future cities must be places in which people want to live, work and play. Therefore, quality of life (and happiness) metrics should be seriously factored into any smarter strategy.

- **Avoiding the easy**, '**vanilla vision**'… **but still aiming for simplicity**

Any vision for a future London should be truly ambitious and revolutionary. Perhaps a 'grand challenge' should be set to capture the imagination of citizens, entrepreneurs and the private sector, and to galvanise the public sector. Technically, a vision is needed that goes far beyond app development as examples of a smart solution or an output of pervasive integration. Geographically, we should look far beyond the UK for examples of ambitious innovation (e.g. in Masdar City[32]). Once London has found something truly game-changing, it should avoid the temptation to imitate—it must trump it.

Finally, there is huge opportunity for environmental and sustainability issues to be more visible in smart city debates, and more integral to any smarter city master plan. There are many synergies between the challenges of realising pervasive, city-wide integration and the pressing energy, emissions and sustainability questions. These include: the specialist cross-disciplinary expertise to be fostered; the coordinated 'systems of systems' approach required; the necessary development of large-scale technological and infrastructural solutions; the citizen engagement central to any initiative's success; and the cohesive public-private cross-sectoral collaborations and responses demanded.

Shaping a future London with pervasive integration and people at its heart would be the smartest action to take.

[32] http://masdarcity.ae/en/

Recommendations to Improve the Smartness of a City

Elsa Negre and Camille Rosenthal-Sabroux

Abstract The concept of "smart city" has not yet been clearly defined. However, there are six characteristics/categories for classifying this kind of cities and compare them: smart economy, smart mobility, smart environment, smart people, smart living and smart governance. However, being "smart" is a challenge increasingly important for many cities or communities. This is of particular interest in the domain of Information and Communications Technology (ICT) and for such systems where there are economic, social, and other issues. To the best of our knowledge, there are no studies that attempt to help identifying the actions to be implemented to improve the smartness of a city. Recommending such actions is an emerging and promising field of investigation. Usually, recommender systems try to predict the rating that a user would give to an item (such as music, books, …) he has not yet considered, using a model built from the characteristics of an item (content-based approaches) or the user's social environment (collaborative filtering approaches). In this chapter, we present a framework for a recommender system for cities. The scope of this research work is to take advantage from recognized "smart cities" and to make same actions for city who wants to become "smart". The followed method is: having a list of characteristics of a "smart city", and having a city which wants to become "smart", which actions must be implemented to become "smart" regarding the characteristics of "smartness". This framework uses the actions already implemented in smart cities to enhance the smartness of a given city. The main idea is to recommend to the city the actions already implemented in those smart cities that are similar (the similarity between two cities is based on some indicators such as air quality, water consumption, etc.)

E. Negre (✉)
LAMSADE, Université Paris-Dauphine, Paris, France
e-mail: elsa.negre@dauphine.fr

C. Rosenthal-Sabroux
Université Paris IX Dauphine, Paris, France
e-mail: camille.rosenthal-sabroux@dauphine.fr

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_5,

as the actions to be implemented in the said city. This is done by (1) Pre-treating the indicators values of a given smart city category (only one among the six), (2) Matching the indicators corresponding to this category, (3) Returning to the city the actions to be implemented in a given order (according to the preferences of the city which needs help, for example). Thus, the city will be able to improve its smartness.

Keywords Information systems • Recommender systems • Smart cities

1 Introduction

Under the influence of globalization and the impact of Information and Communication Technologies (ICT) that modify radically our relationship with space and time, the city increasingly develops its activities in a planetary space with three dimensions:

- A global space covering the set of countries that are the geographic places of implantation,
- A local space corresponding to the subset of cities situated in a given geographic zone and,
- A space of influence that covers the field of interaction of the city with the other cities.

The hierarchical city locked up on its local borders is transformed into an Extended City, opened and adaptable. Furthermore, this Extended City is placed under the ascendancy of the unforeseeable environment that leads towards uncertainty and doubt. The Extended City meets fundamental problems of information exchange and knowledge sharing among, on the one hand, its formal entities distributed in the country (offices, core competencies, business units, projects) and on the other hand, the city's people (nomadic or sedentary), bearers of diversified values and cultures according to the places of implantation.

Two networks of information overlap:

- A formal information network between the internal or external entities, in which circulate data and explicit knowledge; this network is implemented under intranet and extranet technologies.
- An informal information network between nomadic or sedentary employees; this network favors information exchange and tacit knowledge sharing. It is implemented through converging Information and Communication Technologies (for example the new *Apple-ipod*© with Web 2.0).

The problems occur when nomadic employees placed in new, unknown or unexpected situations, needs to get "active information" that are information and knowledge they need immediately to understand the situation, solve a problem, take a decision, and act (Fig. 1).

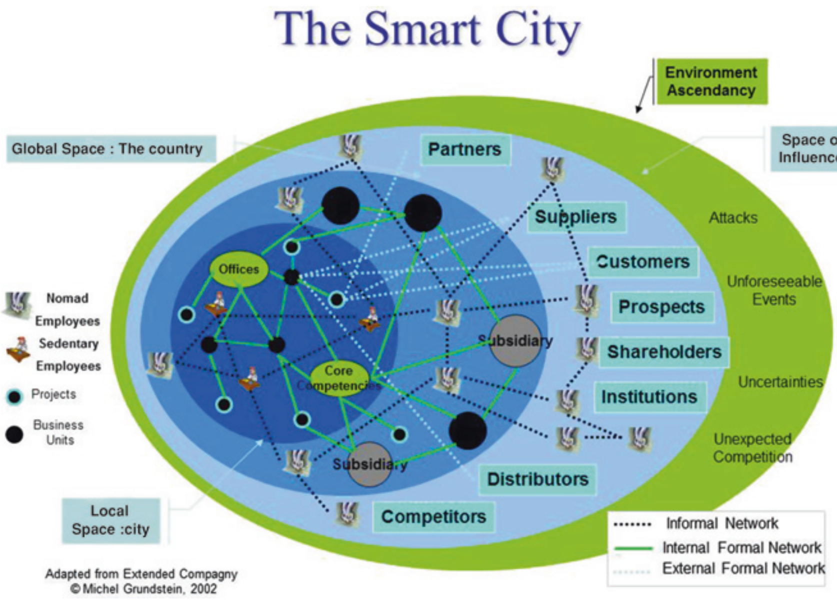


Fig. 1 The Extended city

The concept of “smart city” has not yet been clearly defined. From our point of view the concept of “smart city” is related to ICT and then to Extended City. However, according to [9], there are six characteristics/categories for classifying this kind of cities and comparing them: smart economy, smart mobility, smart environment, smart people, smart living and smart governance. They define for each of the six characteristics a number of factors (in total, 33 factors), each factor being described by a number of indicators (in total, 74 indicators). Each indicator is associated with actions that have been done to attain these indicators. However, being “smart” is a challenge increasingly important for many cities or communities. This is of particular interest in the domain of Information and Communications Technology (ICT); and for such systems where there are economic, social, and other issues. To the best of our knowledge, there are no studies that attempt to help identifying the actions to be implemented to improve the smartness of a city.

This chapter attempts to fill this gap. For our research, the starting point is [9]. Based on the fact that “One of the biggest surprises for the IBMers is how much cities have in common. Whether they’re overgrown towns or giant metropolises, fast-growing or mature, the problems cities face are amazingly similar. And so are the potential solutions” [10]. Recommending actions is an emerging and promising field of investigation. Usually, recommender systems try to predict the rating that a user would give to an item (such as music, books, etc.) he has not yet considered, using a model built from the characteristics of the said item (content-based approaches) or the user’s social environment (collaborative filtering approaches). In this chapter, we present a framework for a recommender system for cities. This framework uses the actions already implemented in smart cities to enhance the smartness of a given city. The main idea is to recommend to the city the actions already implemented in those smart cities that are similar (the similarity between two cities is based on some indicators such as air quality, water consumption, etc.) defined by [9], as the actions to be implemented in the said city. This is done by (1) Pre-treating the indicator values of a given smart city category (only one among the six), (2) Matching the indicators corresponding to this category, (3) Returning to the city the actions to be implemented in a given order (according to the preferences of the city which needs help, for example). Thus, the city will be able to improve its smartness. We based our work on the classification of [9], but the concept can be applied for other classifications as described in [7].

The chapter is organized as follow: Sect. 2 presents the related work about, smart cities, recommender systems and the link between smart cities and recommender systems, Sect. 3 presents the recommendation process with the presentation of the algorithm, finally we conclude with perspectives in order, for example, to take into account social and environmental aspects in the matching.

2 Related Work

2.1 Smart Cities

Based on [9], "A Smart City is a city well performing in a forward-looking way in these six characteristics, built on the "smart" combination of endowments and activities of self-decisive, independent and aware citizens. [...] Each characteristic is therefore defined by a number of factors. Furthermore, each factor is described by a number of indicators. [...] Finally 33 factors were chosen to describe the 6 characteristics": Smart economy (competitiveness, including innovation, entrepreneurship, trademarks, productivity, flexibility, international embeddedness and ability to transform), smart governance (participation, including participation in decision-making, public and social services, transparent governance, political strategies and perspectives), smart environment (natural resources, including attractivity of natural conditions, pollution, environmental protection and sustainable resource management), smart people (social and human capital, including level of qualification, affinity to life long learning, social and ethnic plurality, flexibility, creativity, cosmopolitanism/open mindedness and participation in public life), smart mobility (transport and ICT, including local accessibility, inter-national accessibility, availability of ICT-infrastructure, sustainable, innovative and safe transport systems), smart living (quality of life, including cultural facilities, health conditions, individual safety, housing quality, education facilities, touristic attractivity and social cohesion). These six characteristics and factors form the framework for the indicators and assessment a city's performance as smart city. The indicators that "describe the factors of a smart city are derived from public and freely available data" [9]. We consider that a city is "smart" for a given category if we have information for this category that means that the city is categorized "smart" in this category. If we do not have information for the given category that means that the city is not considered "smart" in this category. In a given category, we know that in reality all the cities do not have the same ranking; in this chapter, we consider that the cities are equivalent. That means that a city is not more "smart" that another in a given category. In the future we can introduce thresholds in a given category [9]. Focuses on medium-sized cities, and on the analysis of characteristics and factors decisive for a successful forward-looking city development, using data from official, public and freely available sources, on the basis of 74 indicators. On the basis of the indicators we mentioned the actions[1] that have been done in order to obtain a given indicator. For example: for *smart environment* and for the factor *sustainable resource management*, the smart city with a high level of indicator makes the actions: *after midnight shut down the public lights and fountains*.

2.2 Recommender Systems

Recommender systems are a particular form of information filtering designed to present information items (movies, music, books, images, web pages, …) that may interest the user.

Recommender systems have been studied in many fields, cognitive science, information retrieval [4, 11, 15], web [3, 20], e-commerce [16], web usage mining [2, 8, 14, 18], data warehouse [13] and many others. The problem of recommendation can be summarized by the problem of estimating scores for items that have not been seen by a user. Indeed, the number of items and the number of users of the system can be very important, it is, therefore, difficult for each user to see all items or that each item is evaluated by all users. It is therefore necessary to estimate the scores for items not yet evaluated.

Intuitively, this valuation is usually based on the scores given by a user to other items and other information that will be formally described below. When it is possible to estimate the scores for items not yet evaluated, then the items with the highest scores may be recommended to the user. More formally, [1] formulates the problem of recommendation in the field of e-commerce as follows.

Definition 1 (*Recommendation for e-commerce*) Given P the set of all users and M the set of all possible items that can be recommended (such as books, movies, restaurants, …). Given u a function measuring the utility of an item m for a user p, i.e., $u : P \times M \rightarrow \mathbb{R}$. Then, for each user $p \in P$, we want to choose the item $m' \in M$ that maximizes the utility for the user: $\forall p \in P, m'_p = argmax_{m \in M} u(p, m)$.

In recommender systems, the utility of an item is usually represented by a score that indicates how a particular user liked a particular item. For example, the user *Michel* gave the score 3 (the maximum score being 10) to the movie *Harry Potter*.

[1] An action is something done so as to accomplish a purpose [21], i.e., it is an operation which produces an effect on something and it is run/operated by a person or a group acting in a particular way [12].

Example 1 In this example, items are movies that the users Elsa, Camille, Michel and Nicolas have given a score. We obtain the matrix $P \times M$:

u(p, m)	Harry Potter	Ice age	Ice age 2	Hulk	Transformers
Elsa		8			7
Camille	9	8			6
Michel	3	5		5	5
Nicolas	5	3		3	3

Note that a cell (i, j) of this matrix corresponds to the utility score given to the movie j by the user i.

The central problem of recommender systems is that the utility u is not usually defined on the full $P \times M$ space, but only on a subset of it. This means that u must be extrapolated to the entire $P \times M$ space.

In recommendation systems, the utility is typically represented by the scores and is first defined over the items previously rated by users. Therefore, the recommendation engine should be able to estimate/predict the scores of item/user unevaluated combinations and to propose relevant recommendations based on these forecasts.

Adomavicius [1] propose that extrapolations from known to unknown ratings are usually done by (a) specifying heuristics that define the utility function and empirically validating its performance, and (b) estimating the utility function that optimizes certain performance criterion, such as the mean square error.

Once the unknown scores are estimated, actual recommendations of an item to an user are proposed by choosing the highest score among all the scores provided for the user, according to the formula given Definition 1.

A recommendation in e-commerce, as defined Definition 1, is the item $m \in M$ [set of all items (movies, books, …)] such as the utility for a user $p \in P$ (set of all users) is maximum.

2.3 Smart Cities and Recommender Systems

Due to the fact that the concept of "smart city" is still not clear and that this concept is emerging and has been developed in numerous publications [7, 9, 10, 17, 19], there are no studies that attempt to help identifying the actions to be implemented to improve the smartness of a city. Recommending such actions is an emerging and promising field of investigation.

To the best of our knowledge, this is the first work dealing with the problem of recommending actions in smart cities context and specially to improve the smartness of a city. Our contribution is to adapt the information retrieval techniques to our context.

By analogy, we can define a recommendation for smart cities as an action $a \in A$ (set of all possible actions) to implement such as its utility for a city $c \in C$ (set of all possible cities) is maximum.

Definition 2 (*Recommendation for smart cities*) Given A the set of all possible actions and C the set of all cities, given a log of cities and a city that wants to improve its smartness and given u a function measuring the utility[2] of an action a for a city c, i.e., $u : A \times C \rightarrow \mathbb{R}$. Then, for each city $c \in C$, the recommended action $a' \in A$ is the one that maximizes the utility for the city: $\forall c \in C, a'_c = argmax_{a \in A} u(c, a)$.

Example 2 In this example, we just illustrate the obtained matrix $C \times A$ where a score indicates that the action has been implemented and if the city has considered this action as efficient:

u(c, a)	Action 1	Action 2	Action 3	Action 4	Action 5
City 1	8	7			
City 2	9		3		
City 3	3	5		5	5
City 4	5	3		3	3

Note that a cell (i, j) of this matrix corresponds to the utility score given to the action j by the city i. Note also that an example of *Action*, as presented Sect. 2.1, can be *after midnight shut down the public lights*. Finally, note that the scores are obtained by giving an overall rating score but it could be obtained by combining the concepts of cost, time to implement, …

3 Recommendation Process

In this section we detail the framework for recommending actions.

First, we restrain our workspace by hypothesizing some assumptions:

- The city, which wants to improve its smartness, chooses, at the beginning, only one category of smart cities (among the six) it wants to join. Future work will automatically select the efficient category by searching into the log[3] similar cities to the one looking for recommendations.
- Indicator values are numerical. Future work will include semantic similarity between indicator values.
- Indicators are the same for a given category/factor for each city (only the values change and can be null). Future work will relax this constraint by including, for example, semantic similarity between indicators.

[2] Ratings/scores are given by a person authorized to make the decision of implementing different actions and through this score indicates whether the action is (or was) relevant.

[3] The log may be a database or other data structure. It will be powered by the recommender system as and when the use of the system by cities. However, for the initial data (the so-called cold start problem), we hope to use the data of the official, public and freely sources, possibly enriched with the participation of volunteers.

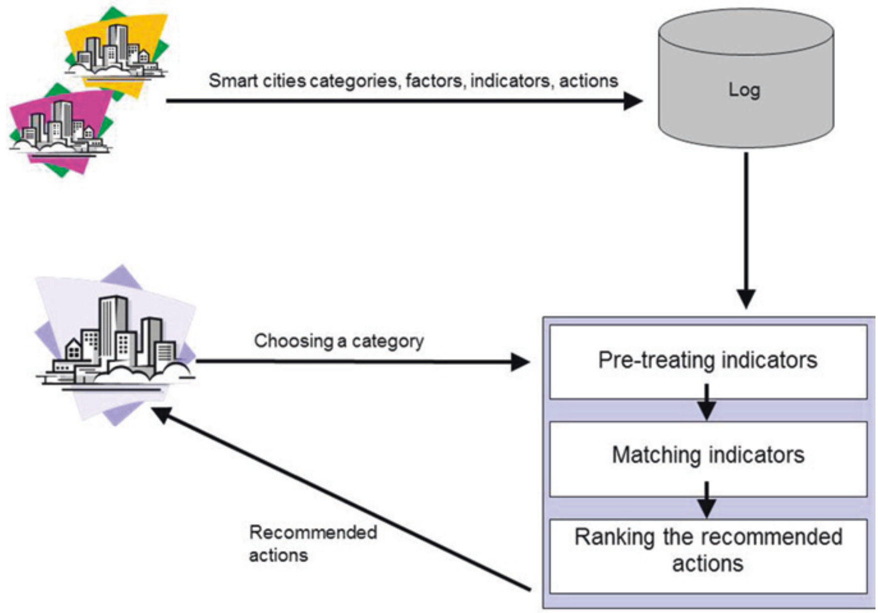


Fig. 2 Overview of the recommendation framework

The framework uses both the characteristics of the city which wants to improve its smartness, and the log containing for each smart city, its category, factors, indicators and implemented actions. It consists of the three following steps, as illustrated in Fig. 2:

1. The first step consists in pre-treating indicators of the smart cities into the log according to the category of smart cities chosen by the city which wants to improve its smartness (the current city),
2. The second step consists in comparing the indicators of the current city and the ones of the logged smart cities and extracting the corresponding actions,
3. The last step consists in ranking the candidate recommended actions.

Each step of this process is presented into more details below.

A city can be seen as a tuple containing the city description and the city information. The city information is a set of smart categories where each category is a set of factors and each factor is a 3-uple specifying the corresponding indicators and their values and for each indicator, the set of implemented actions. Note that the set of categories can be empty if the city is not "smart" for this category. Note also that the set of implemented actions can be empty if no action has been implemented.

So, we have, for each city C_i, $\forall i, j, k, n \in \mathbb{N}^{+*}$:

$$C_i = \langle Description_i, \{Category_{ij}, \{Factor_{ijk},$$
$$\left\{\left\langle Indicator_{ijk1}, Value_{ijk1}, \left\{Action^1_{ijk1}, \ldots, Action^{m_1}_{ijk1}\right\}\right\rangle, \ldots,\right.$$
$$\left.\left\langle Indicator_{ijkn}, Value_{ijkn}, \left\{Action^1_{ijkn}, \ldots, Action^{m_n}_{ijkn}\right\}\right\rangle\}\}\}\rangle$$
$$\text{where } \forall l \in \mathbb{N}^{+*}, m_l \in \mathbb{N}^{+}.$$

The simplified algorithm 1 represents our process. Such an algorithm has a complexity of $O(mn^2)$ where m is the number of cities and n the max number of indicators and/or factors. Some choices of implementation have to be done in partnership with the stakeholders.

3.1 Pre-treating Indicators

Using the log L of smart cities (with their descriptions, categories, factors, indicators, indicator values and actions) and the category G of smart city chosen by the current city, this first step consists in pre-treating indicators of the cities of the log. In fact, we propose to compute intervals of indicator values. Note that our computation is limited to indicators corresponding to the given/chosen category G. In our algorithm, we use the *Pretreat* function. This function is used to compute a set of intervals of indicator values. In fact, this *Pretreat* function is used to search among the smart cities (of the log L) having factors and indicators corresponding to the given/chosen category G, for each indicator $I_{ijkl}, \forall i, j, k, l \in \mathbb{N}^{+*}$, the minimal recorded indicator value $V^{\min}_{ijkl}$ and the maximal recorded indicator value $V^{\max}_{ijkl}$. This function outputs a set of intervals (one per indicator I_{ijkl}) where $V^{\min}_{ijkl}$ and $V^{\max}_{ijkl}$ are the bounds of each interval $\left[V^{\min}_{ijkl}, V^{\max}_{ijkl}\right]$.

Note that, computing such intervals of indicator values is possible because of our hypotheses, i.e., indicator values are numerical and indicators are the same for a given category/factor.

3.2 Matching Indicators

The pretreat function of the previous step outputs a set of intervals of indicator values (one per indicator). The goal of this step is to verify if the indicator values of the current city match the indicator values of smart cities of the log. In fact, according to the previous step, the idea is to verify if the current indicator values belong to the intervals previously computed. Thus, one indicator by one, if the current indicator value $V_{current\,jkl}$ does not belong to the computed interval $\left[V^{\min}_{ijkl}, V^{\max}_{ijkl}\right]$, then the corresponding set of actions $\{A_{ijkl}\}$ (of the given indicator) is selected (without duplicates).

Algorithm 1 RecoSmart(L, C_c, G, Pretreat, Match, Extract, Rank, $\prec$)

Require:
L: The log of smart cities,
C_c: The current city which wants to improve its smartness,
G: The category of smart city chosen by the current city,
Pretreat: A function pretreating the cities of the log,
Match: A match function,
Extract: A function extracting actions,
Rank: A function ranking actions,
$\prec$: An action ranking.
Ensure: An ordered set of recommendations

```
CandCities ← ∅ // for the candidate cities
CandAct ← ∅ // for the candidate actions
Interval ← ∅
for each city c_i ∈ L do
    if c_i.category = G then
        CandCities ← CandCities ∪ c_i
    end if
end for
Interval ← Pretreat(CandCities)
for each factor f_j of G do
    for k ∈ [1..n] // n is the number of indicators of f_j do
        if Match(C_c.G.f_j.indicator(k), Interval_{G,f_j,k}) = False then
            for each city c_l ∈ CandCities do
                CandAct ← CandAct ∪ Extract(c_l.G.f_j.indicator(k))
            end for
        end if
    end for
end for
if CandAct ≠ ∅ then
    Rank(CandAct, ≺)
end if
return CandAct
```

The goal of the extract function is to extract a set of actions that will be the basis for the recommendation, i.e., the extracted actions can contribute to help the current city to enhance its smartness by joining the chosen category G. The obtained set of actions, *CandAct*, is the set of unordered recommendations which is returned.

Note that we are aware that such a set of actions can be voluminous.

Naturally, it would be interesting to take into account the environmental, economic and social aspects during our matching step, as pointed by [5, 6].

3.3 Ranking Actions

In the previous step, a set of recommendations is obtained. The purpose of this next step is to select, when the returned set *CandAct* is not empty, the most relevant one w.r.t. a satisfaction criterion expressed by the city which wants to improve its smartness. To this end, an action ranking is needed, that orders the candidate recommendations. There are many ways of ranking the candidates, from very basic to sophisticated ones. Because it is currently difficult to integrate knowledge in a computer system in order to automate decision making, this step is done by stakeholders. Our future work will try to automate this task.

3.4 Toy Example

In this section, we illustrate our proposition with a toy example (data are synthetic). Suppose there is a log containing some informations about two cities: *Smallville* and *Metropolis*. An extract of this log could be:

Smallville = ⟨small city of USA, {smart environment, {sustainable resource management, {⟨water consumption, 300 (liter per year), {after midnight shut down the public fountains, do not water plants during summer}⟩, ⟨electricity consumption, 3,000 (kW per year), {after midnight shut down the public lights, solar street lamps}⟩}, pollution, {⟨air quality, $\frac{10}{10}$, Ø⟩}}}⟩.

Metropolis = ⟨big city of USA, {smart environment, {sustainable resource management, {⟨water consumption, 100,000 (liter per year), {after midnight shut down the public fountains, do not wash cars during summer}⟩, ⟨electricity consumption, 200,000 (kW per year), {after midnight shut down the public lights}⟩}, pollution, {⟨air quality, $\frac{8}{10}$, {restrict vehicles access into town center, lower vehicle speed on big roads}⟩}}}⟩.

Now suppose that the city named *Gotham* (C_c in Algorithm 1) wants to improve its smartness for the category *smart environment* (G in Algorithm 1) where:

Gotham = ⟨city of Batman, {smart environment, {sustainable resource management, {⟨water consumption, 150,000 (liter per year), Ø⟩, ⟨electricity consumption, 100,000 (kW per year), {after midnight shut down the public lights}⟩}, pollution, {⟨air quality, $\frac{8}{10}$, Ø⟩}}}⟩.

Table 1 resumes informations for each city.

The corresponding utility matrix, as illustrated by Definition 2, could be:

u(c, a)	A_1	A_2	A_3	A_4	A_5	A_6	A_7
Smallville	9	7		8	8		
Metropolis	8		6	7		5	5
Gotham				8			

where the action A_1 is *after midnight shut down the public fountains*, A_2 is *do not water plants during summer*, A_3 is *do not wash cars during summer*, A_4 is *after midnight shut down the public lights*, A_5 is *solar street lamps*, A_6 is *restrict vehicles access into town center* and A_7 is *lower vehicles speed on big roads*.

Into the log, informations about the category *Smart environment* are known for cities *Smallville* and *Metropolis*. So, the candidate cities are *CandCities* = {*Smallville*, *Metropolis*}.

Table 1 Cities values

	Smart environment					
	Sustainable resource management				Pollution	
	Water consumption		Electricity consumption		Air quality	
	Value	Actions	Value	Actions	Value	Actions
Smallville	300	• After midnight shut down the public fountains, • Do not water plants during summer	3,000	• After midnight shut down the public lights, • Solar street lamps	$\frac{10}{10}$	Ø
Metropolis	100,000	• After midnight shut down the public fountains, • Do not wash cars during summer	200,000	• After midnight shut down the public lights	$\frac{8}{10}$	• Restrict vehicles access into town center, • Lower vehicles speed on big roads
Gotham	150,000	Ø	100,000	• After midnight shut down the public lights	$\frac{8}{10}$	Ø

Table 2 Intervals and Gotham values

	Smart environment		
	Sustainable resource management		Pollution
	Water consumption	Electricity consumption	Air quality
Intervals	[300; 100,000]	[3000; 200,000]	$\left[\frac{8}{10}; \frac{10}{10}\right]$
Gotham	150,000	100,000	$\frac{8}{10}$

The pretreat function outputs for the given category G, for each indicator of each factor, an interval of values. In our example, for the category *Smart environment*, we have two factors: *sustainable resource management* and *pollution*. For the factor *sustainable resource management*, we have two indicators: *water consumption* and *electricity consumption*. Values of *water consumption* are 100,000 for *Metropolis* and 300 for *Smallville*. So, the corresponding interval is [300; 100,000]. Values of *electricity consumption* are 200,000 for *Metropolis* and 3,000 for *Smallville*. So, the corresponding interval is [3000; 200,000]. For the factor *pollution*, we have here, only one factor: *air quality* and the values are $\frac{8}{10}$ for *Metropolis* and $\frac{10}{10}$ for *Smallville*. So, the corresponding interval is $\left[\frac{8}{10}, \frac{10}{10}\right]$. Table 2 resumes intervals and values of *Gotham*.

Then, the matching function returns the set of actions (implemented by the candidate cities) corresponding to indicators which Gotham's values are out of bounds defined by the pretreat function. In our example, see Table 2, Gotham's water consumption value is the only one being out of bounds. So, candidate actions that can be interesting (no redundancies) are the actions already implemented by *Smallville* and *Metropolis* for the indicator *water consumption* such as, $CandAct = \{$do not wash cars during summer, do not water plants during summer, after midnight shut down the public fountains$\}$.

Finally, these candidate actions have to be ordered. For example, they can be ordered according to the ease of implementation. It is easier and faster to decree that public fountains have to be shut down after midnight than to force people to not use water during summer. So, a possible ordered set of actions to be implemented in *Gotham* to improve its environmental smartness is: $\{$after midnight shut down the public fountains, do not water plants during summer, do not wash cars during summer$\}$.

The corresponding utility matrix could be:

u(c, a)	A_1	A_2	A_3	A_4	A_5	A_6	A_7
Smallville	9	7		8	8		
Metropolis	8		6	7		5	5
Gotham	9	7	6	8			

where the action A_1 is *after midnight shut down the public fountains*, A_2 is *do not water plants during summer*, A_3 is *do not wash cars during summer*, A_4 is *after midnight shut down the public lights*, A_5 is *solar street lamps*, A_6 is *restrict vehicles access into town center* and A_7 is *lower vehicles speed on big roads*.

4 Conclusion and Perspectives

We based our work on the classification of [9] focusing on medium-sized cities, and on the analysis of characteristics and factors decisive for a successful forward-looking city development, using data from official, public and freely available sources, on the basis of 74 indicators.

In this work, we restrain our workspace by hypothesizing some assumptions:

- The city which wants to improve its smartness, chooses: at the beginning, only one category of smart cities (among the six) it wants to join.
- Indicator values are numerical.
- Indicators are the same for a given category/factor for each city (only the values change and can be null).

Thus, we propose a kind of recommender system. Given some information on smart cities, our framework returns some actions to implement for a city which wants to improve its smartness.

This is a work in progress, it is a theoretical approach, and our future works include real cases with experts. We are aware that in some evaluations it is possible to expect a bias, because we are in human actions. Our main objective is to propose digital information system for decision aid. The final decision is done by decision makers and must be take with all the stakeholders, taking into account economical, social, financial aspects and all others stakes of this kind of project.

Our future work include (but are not limited to):

- Automatically selecting the efficient category by searching into the log similar cities to the one looking for recommendations.
- Including semantic similarity between indicator values.
- Relaxing some constraints by including, for example, semantic similarity between indicators.
- There are many ways of ranking the candidates, from very basic to sophisticated ones. Because it is currently difficult to integrate knowledge in a computer system in order to automate decision making, this step is done by stakeholders. Automating this task is also a challenge.

References

1. Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering, 17*(6), 734–749.
2. Baeza-Yates, R., Hurtado, C., Mendoza, M., Dupret, G. (2005). Modeling user search behavior. *LA-WEB '05: Proceedings of the third Latin American Web Congress*, (p. 242). IEEE Computer Society, Washington, DC, USA.
3. Baeza-Yates, R.A., Hurtado, C.A., Mendoza M. (2004). Query recommendation using query logs in search engines. In W. Lindner, M. Mesiti, C. Türker, Y. Tzitzikas & A. Vakali (Eds), *EDBT Workshops, volume 3268 of Lecture Notes in Computer Science* (pp. 588–596). Springer.
4. Baeza-Yates, A. R., & Berthier, B. A. (1999). *Modern Information Retrieval*. New York: ACM Press.
5. Bettencourt, L., & West, G. (2010). A unified theory of urban living. *Nature, 467*(7318), 912–913.
6. Bettencourt, L. M. A., Lobo, J., Strumsky, D., & West, G. B. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLoS ONE, 5*(11), e13541.
7. Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J., Mellouli, S., Nahon, K., Pardo,T.A.,& Scholl, H.J.(2012). Understanding smart cities: An integrative framework. In *Proceedings of the 2012 45th Hawaii International Conference on System Sciences, HICSS '12* (pp. 2289–2297). IEEE Computer Society, Washington, DC, USA.
8. Fu Y., & Shih, M-Y. (2002). A Framework for Personal Web Usage Mining. In *International Conference on Internet Computing*, pp. 595–600.
9. Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic', N., & Meijers, E. (2007). *Smart cities: Ranking of European medium-sized cities*. Centre of Regional Science SRF: Vienna University of Technology.
10. IBM. Smarter cities. http://smartcitieschallenge.org, 2013.
11. Kent, A. (1971). Information analysis and retrieval. Wiley, Inc.
12. Larousse. Action. http://www.larousse.fr/dictionnaires/francais/action/924, 2013.
13. Marcel, P., Negre, E. (2011) A survey of query recommendation techniques for datawarehouse exploration. In EDA.

14. Pierrakos, D., Paliouras, G., Papatheodorou, C., & Spyropoulos, C. D. (2003). Web usage mining as a tool for personalization: A survey. *User Modeling and User-Adapted Interaction, 13*(4), 311–372.
15. Salton, G.(1983). Introduction to modern information retrieval (McGraw-Hill Computer Science Series). New York: McGraw.
16. Schafer, J. B., Joseph Konstan, A., & Riedl, John. (2001). e-Commerce recommendation applications. *Data Mining and Knowledge Discovery, 5*(1/2), 115–153.
17. Schaffers, H., Komninos, N., & Pallot, M. (2012). Smart cities as innovation ecosystems sustained by the future internet. Technical report.
18. Srivastava, J., Cooley, R., Deshpande, M., & Tan, Pang-Ning. (2000). Web usage mining: Discovery and applications of usage patterns from web data. *SIGKDD Explorations, 1*(2), 12–23.
19. The Committee of Digital and Knowledge-based Cities (CDC) of United Cities and Local Governments (UCLG). Smart cities study: International study on the situation of ict, innovation and knowledge in cities. Chaired by Inaki Azkuna, Major of the City of Bilabo, 2013.
20. White, R.W., Bilenko, M., & Cucerzan, S. (2007). Studying the use of popular destinations to enhance web search interaction. In SIGIR, (pp. 159–166).
21. Wiktionary. Action. http://en.wiktionary.org/wiki/action, 2013.

The Smart City and the Creation of Local Public Value

Federico Fontana

Abstract The creation of public value in a financially sustainable way, which is the distinctive function of local authorities, is becoming increasingly complex. This is due to many reasons, and most notably to the roles played by different stakeholders, such as citizens, businesses, other public authorities and not-for-profit organizations. To address this difficulty, many local authorities state that they wish to become 'smart'. A smart city is meant to be actively engaged in improving the quality of life of its citizens and in pursuing sustainable growth, thanks to the wide use of ICT. The aim of this chapter is two-fold. On a theoretical level, it aims at contributing to the definition of smart city and at critically analyzing its role in the creation of public value. On a practical level, it assesses the adoption of the smart city model by a significant number of large and medium-size Italian cities, in order to draw useful recommendations for the future.

Keywords Smart city • Urban strategic planning • Creation of local public value

1 Introduction

The distinctive function of local authorities is the creation of public value in a financially sustainable way. In other words, they are expected to effectively meet the public needs of their citizens, to generate a positive spread between social benefits and costs and thus to contribute to the prosperity of their constituencies. At the same time, they are expected to pursue financial stability by efficiently using the increasingly scarce and therefore precious public resources [12, 34].

F. Fontana (✉)
Department of Economics and Business Studies, University of Genoa, Genoa, Italy
e-mail: fontana@economia.unige.it

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_6,

The fulfillment of this function is very complex, due to both the growing proliferation of public needs and the progressive lack of available resources, but also because it is significantly affected by the roles played by many other actors, including citizens, businesses, other public authorities and not-for-profit organizations. These aspects are the main reasons for the wide and growing interest in urban strategic planning, which, in fact, may offer a useful contribution to local government, as long as it is set and carried out in an authentic and substantial way.

To face the difficulty of creating public value, the most innovative local authorities state in their urban strategic plans that they wish to become ‘smart’. A smart city identifies an urban environment that is actively engaged in improving the quality of life of its citizens while pursuing sustainable socio-economic development, thanks to the wide use of information and communication technologies (ICT).

Although widely used, so far the concept of smart city has not displayed a consistent meaning and therefore needs to be deepened and better defined. To address this problem, the chapter seeks to clarify the definition of smart city and to identify fields of action in which the city can be smart. It does so by analyzing the potential benefits that a smart city brings to the quality of life, environmental protection and economic development of its community, and at the same time by looking into the possible obstacles and solutions that characterize the relationship between the local authority and other actors in the social system.

Another problem is that even the application of the smart city model displays a large variation. In this regard, the chapter aims at describing the state of the art of Italian regional capitals, seen as a significant sample of large and medium-size cities in the country. Specifically, the chapter tests whether or not the smart orientation is taken into account in their urban strategic plans, analyzes their common and different features, strengths and weaknesses, and suggests some solutions to overcome weaknesses and exploit strengths.

To sum up, the aim of this chapter is two-fold. On a theoretical level, it aims at contributing to the definition of smart city and critically analyzing the relationship between this concept and the creation of local public value. On a practical level, it assesses the adoption of the smart city model by a significant sample of large and medium-size Italian cities, in order to draw useful recommendations for the future. In essence, the chapter aims at providing a critical and empirically informed analysis of the potential success as well as possible failure of various smart city projects.

2 The Creation of Public Value Through Urban Strategic Planning

The creation of public value in a financially sustainable way, which is the distinctive function of local authorities, is becoming increasingly complex, starting with the possibility of divergence between the community socio-economic development and the institution equilibrium [19].

Indeed, creating a positive difference between the benefits that are produced for and the sacrifices that are required from citizens is not in itself guarantee of financial sustainability for the local authority. This is because benefits and sacrifices are partly economic, but mostly non-economic in nature [45] and also because they often correspond to accounting records of opposite sign.

After all, financial sustainability does not necessarily imply the creation of public value, due to the multiple modes of remuneration of local authorities, which only partly require users to pay the nominal value of the services they are offered. Most often, local authorities are rewarded through political prices or taxation, either direct or transferred [38].

Nevertheless, the creation of public value and financial sustainability need to be pursued jointly: the non-transitory absence of either one or the other would in fact deprive local authorities of their own reason or even possibility to exist. Hence, the need to achieve appropriate levels of effectiveness and efficiency, favored by the new public management model and essential to the reasonable satisfaction of public needs, on the one hand, and the convenient use of scarce public resources, on the other [3, 18, 40].

To continue, the creation of public value is characterized by the degree of operational diversification of local authorities [4]: let's just think of the plurality of functions performed and of the services produced, which are very significant in terms of areas and groups to be targeted, of content and modalities of intervention and, finally, of the multi-disciplinary and multi-sectorial skills and of the composite nature of the problems to be tackled (which often involve aspects that are at the same time environmental, social, economic and technical).

Also the targeted geography is variable and often does not correspond to the administrative boundaries. One classical example is to be found in the field of transport but also in the public services of water distribution or tourism promotion.

Even more important, in terms of operational complexity, is the framework of relationships of opposite sign, sometimes co-operative sometimes competitive, that come into play. To begin with, the governing bodies of local authorities are the expression of the ideas, values and claims of only one section of the community. Secondly, in many cases there is no overlap between the citizens who use the services and products of a given local authority and those that contribute to their funding. Finally, even among the actors that use those services there are often divergent interests, which are functionally antagonistic (e.g. consumers and businesses, pedestrians and motorists) or compete in the allocation of scarce resources. The systemic process of bringing together different expectations is therefore a fundamental and critical condition for the creation of local public value.

Another element of complexity is the dynamism that significant changes in the socio-economic, scientific-technological, and political-cultural domains impress on public needs and public policies [2]. This means that the true identity of the city, of the territory and of the local community is often questioned, if not completely doubted, because of phenomena that give it uncertainty and discontinuity. At the same time, local authorities are assuming roles and features that are more and more composite: the productive role that pertains to them in their quality of service units, the directing role that belongs to them in their quality of public holdings, and the regulatory role that fits them in their quality of local governing bodies.

No less significant are the complementary institutional, political and business dimensions of local authorities [6]. The institutional dimension refers to the set of rules that constitute its statutory principles, defining both the areas of activity and the degree of autonomy [50]. The political dimension refers to the systematic search for consensus that characterizes all government entities, which has to be harmonized with the managerial function expressed by the administrative and technical structure [18, 35, 48]. Both affect the business dimension, influencing both the form of financial sustainability and the creation of local public value.

Last but not least, it should be noted that socio-economic development only partially depends on local authorities. The roles played by other actors in the system—citizens, businesses, other public authorities and not-for-profit organizations—are equally determinant, as well as the contributions they make in terms of resources, expertise, ideas and actions [28]. They constitute a rather fragmented framework, but their attitudes and behaviors nevertheless affect the output produced by local authorities and more generally the process of creation of public value. It is therefore critical that local authorities adopt a public governance approach, namely, a willingness and an ability to play the important role of attracting, involving, monitoring, and promoting the activities of other social actors. They need to facilitate and positively orient, in a collaborative and synergistic sense, the individual and collective development of these actors, thus contributing to generate and at the same time draw upon the social capital of the whole community [39, 42].

The above-mentioned aspects of specificity and complexity that characterize the distinctive function of local authorities are the main reasons for the wide and growing interest in urban strategic planning [33].

Urban strategic planning, in fact, may offer a useful contribution to local government, as long as it is set and carried out in an authentic and substantial way. It is necessary that the plan does not limit itself to only internal and external analyses—even if they are to a degree indispensable. It also needs to identify a fair model of development that is guided by a long-term and far-reaching vision and is able to make clear the meaning of its foundational choices and, on this basis, identify possible courses of action, projects to be given priority and related operational solutions.

More specifically, the effectiveness of the urban strategic plan requires some appropriate conditions for both the object and the subjects of planning.

With regards to the contents of the plan, it is necessary to have a selective and integrated approach, which is both far-reaching and perspective, sustainable and flexible [43, 46].

A selective approach, limited to a few themes, objectives and projects that are relevant to the socio-economic development of the territory and the community, is essential to focus attention and actions on crucial and decisive issues, those that are able to have a greater impact on future scenarios, and to avoid instead dispersion of energies and dissipation of resources.

At the same time, an integrated approach that is mindful of the interdependence and co-determination of the various policy interventions is essential to make them consistent and coordinated, to generate useful synergies and to create systemic value. From a space-time perspective, a far-reaching and long-term horizon is necessary to take account of complementarities (both as sources of constraints and opportunities) between different regions and various levels of government, to achieve important goals, to coagulate significant resources, and to enable innovative processes, overcoming the shortsightedness and constraints of each single administrative mandate.

Further key features to take into account are the sustainability and flexibility of the contents of the urban strategic plan. The former, which is the result of the beneficial correlation between goals and resources, makes the urban strategic plan rational and realistic, at once ambitious and feasible, avoiding idealistic temptations as well as the propensity to give up. The latter, which corresponds to the dynamism of the context, makes the plan adaptive and constantly updated (both in terms of geographical and operational contents).

In summary, all these characters of the urban strategic plan allows to identify (1) the areas in which the city, on the basis of its identity, vocation and resources can (and should) try to excel autonomously, (2) the areas in which, in order to be successful, the city must weave collaborative relationships with other entities (and on which it would be appropriate to invest), and, finally, (3) the areas where the city does not have or can acquire the conditions to play a significant role (and which it would be reasonable to give up).

As for the role of the actors involved in the planning process, the principles of openness, partnership and leadership are fundamental [14, 41].

The drafting and subsequent implementation of the urban strategic plan requires an open and transparent approach, which is at time relational and communicative, engaging and participatory. This would promote the fruitful interaction of the plurality of key public and private subjects and prevent both the self-referential attitude of the former and the opportunism, indifference or exclusion of the latter. In this way, it is possible to strengthen democratic participation as well as the accountability of local authorities, to balance all powers involved, to positively deal with conflicts of interest, and to promote mutual trust and a sense of belonging of all different actors, thus encouraging collaborative and proactive approaches.

Growing importance is also attributed to the development of partnership relationships between public and private actors. These alliances, which are the result of voluntary agreements governed by fair rules and by negotiation skills, allow for a clear distribution of responsibilities, tasks, risks and benefits among all relevant stakeholders.

What is essential, in any case, is the exercise of the function of leadership by local authorities, which presupposes their competence and legitimacy and results in the construction of truly shared and consensual scenarios.

Under these conditions, the plan can be a real and high-impact instrument for public governance and strategic management, able to dynamically integrate the needs for economic development and social and environmental protection with the management tools that are necessary to achieve the shared goals, on which it is then possible to gravitate interests, generate resources and promote the assumption of responsibilities [8, 21].

3 The Smart City Model for the Creation of Public Value

In the previous section we showed how the creation of local public value in a financially sustainable way is a very complex function, which can find a useful governance tool in the urban strategic plan.

To address this complexity, many local authorities state that they wish to become smart. A smart city is meant to be actively engaged in improving the quality of life of its citizens and in pursuing sustainable socio-economic development, thanks to the wide and innovative use of ICT.

However, so far the concept of smart city, although widely used, does not have a consistent meaning and therefore needs to be deepened and better defined.

The concept of smart city was first mentioned in the mid-90s [5], although its use became prominent at the beginning of the third millennium, due, on the one hand, to the interests of multinational companies operating in the ICT sector, such as IBM and Cisco, and, on the other hand, to the attention that international bodies such as the European Commission and the OECD devoted to the subject [23, 25].

It is no coincidence that, even in the scientific literature on the subject, different schools of thought have developed around the concept of smart city [11, 13, 36, 44].

Among the most prominent of these schools of thoughts, there is the one that focuses on ICT applied to the redesign of every aspect of urban life. In this sense, the smart city is considered an urban environment at the same time equipped, interconnected and intelligent [26]. An appropriate hardware, software, and network equipment composed of sensors, kiosks, personal devices, smartphones, tablet PCs, GPS devices, the web, social networks, etc. can detect massive amount of data on the life of the city in real-time [37]. Their interconnection, that is, their integration on a platform of enterprise computing, allows for the exchange of information between the various municipal services [31]. The intelligent use of such information allows to perform complex analyzes, to develop conceptual models, and to visualize and optimize critical processes, in order to take the most rational operational decisions [49].

This meaning of smart city can be seen as an extension and evolution of other concepts of the city, such as the ‘digital city’ and the ‘ubiquitous city’.

The digital city, the most dated among these concepts, was created to refer to any digital initiative undertaken by a city, starting with the provision of Internet access (in this case we also refer to the ‘wired city’) up to the 3D representation of the city (the so-called ‘cyber city’). In the most general sense, the digital city is identified as an information system that collects digital information on the real city and makes it available in a virtual public space, where citizens can consult it, but also interact with the system and with other users (hence the oftentimes used term of ‘intelligent city’) [16, 32].

The ubiquitous city (also referred to as ‘U-city’) further develops the idea of the digital city, creating a new generation of urban space, which results from the convergence between physical world and virtual reality. The U-city is defined as an innovative model designed to improve the management of the city, the quality of life and economic development by identifying the critical success factor in the attention given to the end user. Nonetheless, there are projects that focus only on certain categories of citizens (e.g. young people), which mitigates the user-centric nature of the ubiquitous city [10, 29].

In addition to the current of thought that focuses on ICT, another one worth mentioning defines smart cities as those cities that thoroughly innovate their governance and their own conditions of socio-economic development. This meaning, although it does not renounce the support that comes from a wide and innovative use of ICT, focuses on the proper fulfillment of the needs of citizens, businesses and other organizations. From this point of view, a smart city, by monitoring and integrating its critical infrastructure, whether it is the physical capital (roads, bridges, etc.), technological capital (hardware, software, network) or intellectual and social capital (resulting from the relationships between the members of the community), plans the activities of prevention, maintenance and management, makes an efficient use of its resources, and optimizes the effectiveness of its services. Under these conditions, a smart city is an urban context that is at the same time innovative, competitive, effective, efficient, as well as safe, livable, equitable, and sustainable [24, 47].

The main difference between this meaning of smart city and the previous one consists in the role attributed to ICT. In the first case, ICT is an indispensable element around which everything revolves; in the second case, it is only one of the pillars of the model, of which it represents an important enabling factor, but not necessarily the only one, and, sometimes, not even the most important one.

The non-ICT-centered meaning is characterized by a broader, more flexible and open vision. A vision that appears more consistent and convincing, certainly more coherent with the objective of creating local public value. Like for the ICT-centered approach, even this meaning can be related to other recently developed concepts of the city, and in particular to that of ‘knowledge city’ [1, 9].

In essence, a knowledge city is purposely designed to encourage and nurture the collective knowledge, that is, the intellectual capital of the community, seen as a determinant factor for the sustainable creation of local public value [15]. This city-model derives its social, environmental and economic success by a series of factors, notably [17]: the allocation of facilities, networks and tangible and intangible assets for the production of goods and services based on knowledge (in the broadest sense of the word and, thus, potentially in its scientific, technological, cultural, and artistic manifestation); the development of conditions able to promote talent, creativity, innovation and entrepreneurship; the availability of technologies, instruments and services for the systematic, effective and efficient dissemination of knowledge; the presence of actual and virtual places that can facilitate interpersonal relations, the exchange of information and the sharing of experiences; and, finally, the ability to generate, attract and retain citizens who are not only highly qualified from a professional point of view but also engaged with the political-institutional life and environmentally-conscious.

The similarities between the concept of knowledge city and that of smart city are very apparent, although the former is characterized by a greater focus on intellectual and social capital, and the latter by a broader, more open and flexible perspective. The concept of smart city, therefore, is more complete and more easily applicable to the majority of urban areas, since it is respectful of their identity, their distinctive characteristics and their evolutionary paths.

On this basis, it seems interesting to identify the areas of intervention in which the city can be smart, that is, able to contribute to the quality of life of its citizens, to the protection of the environment and to economic development. A systemic approach allows to identify six relevant dimensions [22]: smart economy (i.e. competitiveness), smart people (in terms of social and human capital), smart living (i.e. quality of life), smart environment (i.e. attention to the natural resources), smart mobility (which refers to both transport and telecommunication networks and services), and, of course, smart governance (with its features of openness, transparency, participation, and accountability).

These are the same dimensions that the European Commission takes into account when designing programs to give financial support to smart cities.

In this regard, it should be noted that the reference to such dimensions, each of which can be further articulated, has the advantage of making the model very encompassing, covering all the areas of intervention of the city. Yet it is unlikely that a single city can excel in all the above-mentioned areas. It is more likely that each urban reality can be smart in one or more areas of intervention (e.g. economic development, the protection of the environment), but not necessarily in all of them.

In other words, there is no single model of smart city, but rather as many variations as there are possible meanings and contexts of ‘smartness’, with all their possible nuances and combinations [27].

Despite its conceptual variety, the smart city, to be considered such and to become successful, must prove to be genuinely creative. This means that it must develop an original model of socio-economic development through a clear strategic direction, a model that makes the most of its identity, its vocation and its specificity, avoiding improvised or unrealistic approaches as well as to give into emulative practices—unfortunately fairly common.

Obviously, some contextual conditions are essential, notably the concentration, variety and variability of the community of reference [7, 30]. This concentration, which is defined as the presence of a significant number of people in a given geographical space, is an essential factor from both a qualitative and a quantitative perspective: it ensures the necessary population density, but especially the high intensity and frequency of interpersonal and inter-organizational relationships within which smart ideas can grow and spread.

Variety, in the broadest sense of the term, refers both to the community (i.e. diverse people, knowledge, activities and needs) and to the territory (a combination of different uses of the urban area, e.g. residential, touristic, administrative, manufacturing, commercial, recreational). Variety determines the wide array of opportunities for interaction and promotes the development of creativity, innovation and entrepreneurship.

Variability, in the double meaning of instability and dynamism, is also very significant, since it is from situations of uncertainty and struggle that important innovations might emerge (especially when the fear of a crisis overcomes the aversion to change). Likewise, it is from the opening and consequent evolution of the urban environment that the cognitive capital can be increased and new opportunities for development can materialize.

However, in order to accrue the benefits deriving from the creation of public value, a smart city not only must (try to) be such, it must also be able to communicate its objectives and be perceived as a smart city by all relevant stakeholders. The construction of an image that is at the same time recognizable and attractive, credible and distinctive plays a decisive role in determining the success of a smart city.

4 An Empirical Study About Smart City and Urban Strategic Planning

In view of the contribution that the smart city model gives to the creation of local public value, it may be interesting to assess if and how this model is included in urban strategic plans—and if consideration is given to its various meanings, fields of activity and contribution to the socio-economic development. Specifically, we intend to analyze whether or not the smart orientation is taken into account in the urban strategic plans of the Italian regional capitals, as reflected in the documents published on their institutional websites.

Focusing on the regional capitals allows us to analyze a relatively limited but significant sample of institutions that, although characterized by some common features, differ in several aspects, ranging from the size, geographical characteristics and territorial and socio-economic aspects. By covering substantially all of the significant areas of the country, they constitute a sufficiently representative sample of the variety that characterizes the system of local authorities in Italy.

The documents considered, despite the variety of denominations and methodological approaches, include all urban strategic planning tools, but also any other document that specifically refers to the concept of smart city and that is published on the web as of July 31, 2013. The reference to the documents available online provides useful information on the degree of sensitivity of the specific local administration to the wider dissemination—in terms of accountability—of the information included in its plans.

In general, there is a significant commitment to urban strategic planning (the data shown represent an update to those reported in [20]): 20 out of 21 municipalities (95 %, with the exception of Trieste) have started a strategic process. There is also a high level of disclosure, since 18 municipalities out of 20 (90 %, with the sole exceptions of Potenza and Catanzaro) publish online their urban strategic plans (Table 1).

No less significant is the reading of the data at a demographic level, according to the classes identified by the Ministry of the Interior, and with reference to the geographical areas identified by the Italian Institute of Statistics—ISTAT (Table 2).

From a demographical point of view, the classes of municipalities with the highest level of strategic elaboration and dissemination are those with more than 250,000 inhabitants, where all the institutions establish and publish on the web their strategic planning documents. The next smaller size class (between 100,000 and 250,000 inhabitants) is still characterized by a high degree of strategic disclosure, since this class of cities make available online all the plans they formulated (5 out of 6 bodies, representing 83 % of total local authorities). Relatively smaller, however, is the commitment of the regional capitals of smaller size (up to 100,000 inhabitants), in which only 60 % of the plans are published.

Geographically, the cities in the North–West and Center of the country plus the Islands are those that, overall, are characterized by a larger strategic development and transparency, with the formulation and online publication of strategic plans by all regional capitals. The North–East area still displays a substantial level of strategic planning and disclosure (all 4 strategic plans formulated by the 5 municipalities included in the analysis are published online, representing 80 % of the total). The Southern area, even in the presence of a high level of strategic planning commitment (all regional capitals have begun the process of strategic planning), is characterized by a lower level of disclosure (about 67 %).

In addition to the number of strategic plans produced and disclosed, it's interesting to analyze some other qualifying aspects.

First of all, even if all the documents are characterized by a strategic breath and a medium to long-term perspective (usually 10 years long), 2 out of 17 plans (representing 11 % of the total) focus exclusively on urban-regulation aspects (it's the case of Ancona and Milan), even if they are the result of participative decision-making processes.

It is also important to point out that, although in most cases the process of urban strategic planning and implementation was directly promoted by the local authority, there are cases, like those of Turin and Florence, where the process was initiated, implemented and disseminated by a separate organization (namely '*Strategic Turin Foundation*' and '*Future Florence Association*') gathering both public and private actors and with no management power. In these cases, the plan may contain highly sophisticated analyses and proposals, be perceived as the privileged site for the meeting and engagement of all key-players and for the establishment of an effective communication strategy, but is hardly seen as an authentic instrument of local government.

Table 1 Urban strategic plans of Italian regional capitals: overall framework

Cities	Demographic classes and geographical areas	Title of urban strategic plans	Year
Aosta	*Up to 100,000 inh North–West*	*Future of Aosta: Strategic Plan of Aosta and of La Plaine*	2010
Turin	*From 500,000 to 1,000,000 inh North–West*	*1) City Strategic Plan—International Turin*	2000
		2) 2° Strategic Plan of the Metropolitan Area	2006
Genoa	*From 500,000 to 1,000,000 inh*	*1) Plan of the City of Genoa*	2002
	North–West	*2) The City Changes (UrbanLab)*	2009
Milan	*Over 1,000,000 inh North–West*	*Government Plan of the Territory*	2011
Trento	*From 100,000 to 250,000 inh*	*1) Strategic Plan 2010*	2003
	North–East	*2) Strategic Agenda 'Trento 2020'*	2007
Bolzano	*From 100,000 to 250,000 inh North–East*	*Ideas for 2015: Thinking the City*	2004
Venice	*From 250,000 to 500,000 inh North–East*	*Venice Metropolitan Area*	2004
Trieste	*From 100,000 to 250,000 inh North–East*	*N/A*	N/A
Bologna	*From 250,000 to 500,000 inh North–East*	*Metropolitan Strategic Plan*	2013
Florence	*From 250,000 to 500,000 inh*	*1) Strategic Plan Florence 2010*	2002
	Center	*2) There is More than One Florence*	2009
Ancona	*From 100,000 to 250,000 inh Center*	*A Plan for Ancona: the Changing City*	2009
Perugia	*From 100,000 to 250,000 inh Center*	*Perugia—Europe from 2003 to 2013*	2004
Rome	*Over 1,000,000 inh Center*	*Strategic Plan for the Development of Rome Italian Capital*	2009
L'Aquila	*Up to 100,000 inh South*	*L'Aquila 2020*	2008
Campobasso	*Up to 100,000 inh South*	*Territorial Strategic Plan*	2008
Bari	*From 250,000 to 500,000 inh South*	*BA2015—Metropolitan Area of Bari*	2008
Naples	*From 500,000 to 1,000,000 inh South*	*Strategic Plan*	2006
Potenza	*Up to 100,000 in South*	*Strategic Project of Potenza's Hinterland*	2005
Catanzaro	*Up to 100,000 inh South*	*Strategic Plan*	2011
Palermo	*From 500,000 to 1,000,000 inh Islands*	*Palermo, Capital of the Euro-Mediterranean Area*	2010
Cagliari	*From 100,000 to 250,000 inh Islands*	*Strategic Plan*	2008

Table 2 Urban strategic plans of Italian regional capitals: data by demographic classes and geographical areas

	N. of cities	N. (%) total population		N. (%) of plans approved		N. (%) of plans online	
Total	21	9,732,740	(100 %)	20	(95 %)	17	(81 %)
Demographic classes (interior ministry)—N. of inhabitants:							
Up to 100,000	5	319,897	(3 %)	5	(100 %)	3	(60 %)
From 100,000 to 250,000	6	853,516	(9 %)	5	(83 %)	5	(83 %)
From 250,000 to 500,000	4	1,342,822	(14 %)	4	(100 %)	4	(100 %)
From 500,000 to 1,000,000	4	3,130,918	(32 %)	4	(100 %)	4	(100 %)
Over 1,000,000	2	4,085,587	(42 %)	2	(100 %)	2	(100 %)
Geographical areas (ISTAT):							
North–West	4	2,874,628	(30 %)	4	(100 %)	4	(100 %)
North–East	5	1,076,927	(11 %)	4	(80 %)	4	(80 %)
Center	4	3,403,925	(35 %)	4	(100 %)	4	(100 %)
South	6	1,564,897	(16 %)	6	(100 %)	4	(67 %)
Islands	2	812,363	(8 %)	2	(100 %)	2	(100 %)

One further consideration to make is that most of the urban strategic plans are fairly recent (the oldest one was approved in 2000 and only 9 out of 17 plans, representing 53 % of the total, are more than 5 years old). It would therefore be premature to assess the impact they had on their socio-economic environment. Among the older experiences, four are fairly significant, having already moved to the second generation of urban strategic plans. In the cases of Turin, Trento and Florence the second-generation plan stems from a critical analysis of the structure, content, status of implementation and impact of the first-generation plan. In the case of Genoa, however, the two documents are not sequential and rather highlight a discontinuity of both strategic and administrative nature.

Within this framework and considering the overall high levels of strategic planning and disclosure recorded by the generality of Italian regional capitals, it is interesting to analyze if, how and what of the smart city model is reported in their urban strategic plans (Table 3).

The first thing that can be evidenced is that only 4 of the 18 analyzed urban strategic plans contain specific references to the smart city model. Moreover, these 4 plans refer to individual areas of activity, such as the security of infrastructure, eco-friendly construction activities, energy efficiency, sustainable mobility and the use of ICT in the delivery of services to citizens. They all lack an overall strategic vision of the smart city.

Another thing that can be highlighted is the demographic and geographic distribution of the collected data: first, the 4 cases that cite the smart city model belong to different demographic classes (with weights ranging from 20 % to 50 %), with the sole exception of the class between 100,000 and 250,000 inhabitants; second, each of them belong to a different geographical area (with weights ranging between 17 % and 25 %), with the sole exception of the Islands.

In essence, the distribution of the few urban strategic plans containing specific references to the model of the smart city is numerically rather homogeneous, both demographically and geographically.

Nevertheless, all the plans of the Italian regional capitals contain frequent references to aspects that are considered typical of the smart city model, such as change (e.g. the plan of Bolzano '*Ideas 2015: Thinking the City*', the second plan of Genoa '*The City Changes*', and '*A Plan for Ancona: the Changing City*') and innovation (e.g. '*Venice—City of higher education, research and innovation*', '*Bari—Research & Innovation, The metropolis in a bit*', and '*Cagliari—Knowledge, innovation and development*').

Consequently, it seems that at the time of the preparation of these urban strategic plans, the reference to the smart city model was not yet sufficiently robust and widely known, so as to remain largely unexpressed or marginal. In view of these findings, it is interesting to see whether, beyond the content of the urban strategic plans, the smart city model finds confirmation in other planning documents that the regional capitals have approved and published online (Table 3).

This approach leads to substantially different results, since 13 of these municipalities (62 % of the total) publish on their websites documents where they declare their intention to become smart (mostly in response to bids for funding at the

Table 3 The smart city model in the urban strategic plans and other specific documents of Italian regional capitals

	N. (%) of online documents containing references to the smart city model	Urban strategic plans	Other specific documents
Total	***Overall situation***	**18 (86 %)**	**13 (62 %)**
	of which: without significant content	14 (67 %)	–
	specific fields of activity	4 (19 %)	8 (38 %)
	overall model	–	5 (24 %)
Demographic distribution	***Up to 100,000 inhabitants***	**3 (60 %)**	**2 (40 %)**
	of which: without significant content	2 (40 %)	–
	specific fields of activity	1 (20 %)	2 (40 %)
	overall model	–	–
	From 100,000 to 250,000 inhabitants	**5 (83 %)**	**2 (33 %)**
	of which: without significant content	5 (83 %)	–
	specific fields of activity	–	2 (33 %)
	overall model	–	–
	From 250,000 to 500,000 inhabitants	**4 (100 %)**	**3 (75 %)**
	of which: without significant content	3 (75 %)	–
	specific fields of activity	1 (25 %)	3 (75 %)
	overall model	–	–
	From 500,000 to 1,000,000 inhabitants	**4 (100 %)**	**4 (100 %)**
	of which: without significant content	3 (75 %)	–
	specific fields of activity	1 (25 %)	2 (50 %)
	overall model	–	2 (50 %)
	Over 1,000,000 inhabitants	**2 (100 %)**	**2 (100 %)**
	of which: without significant content	1 (50 %)	–
	specific fields of activity	1 (50 %)	–
	overall model	–	2 (100 %)

(continued)

Table 3 (continued)

	N. (%) of online documents containing references to the smart city model	Urban strategic plans	Other specific documents
Geographical distribution	***North–West***	**4 (100 %)**	**4 (100 %)**
	of which: without significant content	3 (75 %)	–
	specific fields of activity	1 (25 %)	2 (50 %)
	overall model	–	2 (50 %)
	North–East	**4 (80 %)**	**3 (60 %)**
	of which: without significant content	3 (60 %)	–
	specific fields of activity	1 (20 %)	3 (60 %)
	overall model	–	–
	Center	**4 (100 %)**	**2 (50 %)**
	of which: without significant content	3 (75 %)	–
	specific fields of activity	1 (25 %)	1 (25 %)
	overall model	–	1 (25 %)
	South	**4 (67 %)**	**3 (50 %)**
	of which: without significant content	3 (50 %)	–
	specific fields of activity	1 (17 %)	1 (17 %)
	overall model	–	2 (33 %)
	Islands	**2 (100 %)**	**1 (50 %)**
	of which: without significant content	2 (100 %)	–
	specific fields of activity	–	1 (50 %)
	overall model	–	–

national and EU level). Within these 13 cases, the majority (8 out of 13, 62 %) talks of smart interventions in specific fields of activity (the same that were mentioned above), but there are also cases (more specifically 5, which account for 38 % of the total) that refer to a comprehensive model of smart city.

As pointed out when talking of urban strategic planning, even in the development of smart city projects there are both initiatives launched directly by the local authorities (8 out of 13 cases, 62 % of the total) and initiatives launched by separate organizations (mostly associations or foundations) promoted by the same municipalities (5 cases out of 13, 38 % of the total). It should be noted, however, that in no case the launching association or foundation is the same organization that is involved in the urban strategic planning—at the most there are forms of collaboration that develop between the two entities (e.g. '*Turin Smart City Foundation*' vs. '*Strategic Turin Foundation*').

For completeness, it is also worth noting that in four other urban areas (Florence, Potenza, Trento and Trieste), some initiatives aimed at developing smart projects have even been initiated by organizations to which, at least so far, the local administration does not participate directly.

This multifarious framework allows us to develop some critical considerations.

The fact that only one-fifth of the urban strategic plans formulated by the Italian regional capitals refer explicitly to the smart city model can have two complementary meanings. On the one hand, it may signal the weakness or, more simply, backwardness (even only from a terminological point of view) of most of the analyzed urban strategic plans, which overlook a relevant and critical model for the creation of local public value. On the other hand, it may signal the lack of strategic importance that was attributed to the smart city model, at least until the time these plans were approved (which is pretty recent). This could have happened despite the smart city model is formally identified as instrumental in improving the quality of life, safeguarding the environment and promoting economic development.

The latter interpretation seems to be confirmed by the fact that 62 % of the analyzed municipalities pursue, in fact, smart city projects, but mainly in the context of their participation in specific bids for public funding. These are certainly positive for the innovative opportunities they offer, but nevertheless expose to the risk of undertaking occasional or sporadic initiatives that are not included in a clear strategic vision. Although it is too early to evaluate the results that can be achieved in this way, another risk worth mentioning is that these projects, once the funds allocated to them are exhausted, get abandoned, making their socio-economic impact extremely modest and ephemeral.

Moreover, the fact that in a significant number of cases, the pursuit of smart city projects is delegated to organizations outside the local administration (not to mention those cases in which the initiative is promoted by entities to which the municipality does not even participate) may, in turn, be variously evaluated. On the one hand, it is a solution that can support the wide and open involvement of the plurality of public and private stakeholders. On the other hand, it is a situation that, in the absence of specific managerial powers attributed to the delegated organization, can hardly be an effective form of local government (as already noted on the subject of urban strategic planning). This can result in an excellent design of smart city, which, however, cannot be concretely implemented outside some random occurrences.

Finally, the fact that there is a predominance of projects focused on specific fields of activity rather than on a comprehensive model of smart city can, also, be interpreted in different ways. On the one hand, this can be a strength, if it means that only the aspects considered most relevant and critical to the specific urban context are selected. On the other hand, it can be a point of weakness, if these projects are not part of a clear strategic vision.

The latter interpretation seems, unfortunately, more likely, as the areas of activity that are addressed in the smart city projects are often common to several cities, not assuming, at least apparently, a character that is tailored to the specific urban situation. In addition, as already mentioned, since these projects substantially correspond to the activities that are financed with public funds, they seem to reflect an opportunistic behavior rather than strategic choices that are broad and forward thinking. This adds to the fact that in several cases the only chosen area of activity is the development of ICT.

As mentioned earlier, technological innovation is an essential condition for any smart city project. However, such projects run the risk of failing if they are designed to respond to a technological innovation rather than to an actual need. In other words, these projects are likely to propose answers to needs that are not felt by the citizens, perhaps neglecting others that are of greater importance for everyday life.

Even the usability factor of the technological tools that are developed assumes a certain importance. In countries where the average age is rather high, like Italy, it needs to be considered that large segments of the population are not familiar with digital solutions and therefore will tend not to use them, despite having them available, even when they respond to actual needs.

After all, it should not be overlooked that citizens must be made aware not only of the existence and availability of a service, or its ease of use, but also of the concrete benefits that the service itself can bring to each of them individually and to the community. For example, equipping bus stops with digital panels providing passengers with real-time information on the arrival times of buses can even be counterproductive, if first the efficiency and proper frequency of the public transport service is not ensured.

5 Conclusion

To sum up the main points touched in this chapter, a city can be defined smart when the investments in physical, technological, intellectual and social capital nurture a sustainable economic development and a high quality of life, while at the same time wisely managing natural resources and using a participatory model of government. It is important to remark that the quality of being smart does not have to be uniquely related to the presence of ICT, but also to the recognition that the intellectual and social capital as well as the physical capital are important factors in the creation of local public value.

From an infrastructural point of view, it is important that the available resources are used together to improve economic and political efficiency and enable social and urban development. From a social point of view, a smart city is a city whose community has found out how to learn, adapt and innovate, with a particular focus on achieving social inclusion and citizen participation in urban governance. From an environmental point of view, sustainability emerges as a priority; this is a very important aspect in a world where resources are scarce and cities increasingly base their development also on the availability of natural resources. From an economic point of view, a city can be considered smart if, thanks to its competitiveness, is able to attract new businesses and thus to increase local prosperity.

Consequently, research on the smart city is both complex and fascinating and may represent one of the main areas of urban innovation and development in the coming years.

To be effectively set up and implemented, however, the smart city model requires competence and the ability to follow through. It cannot be managed in an improvised or episodic way. It requires a strategic vision that is specific, clear and selective and a system of governance that is authentic, open and engaging.

To this end, it is necessary that the smart city model is clearly stated in the urban strategic plan and, in an integrated and convergent way, in the operational programs and budgets of the local authority. This condition is, in fact, essential to make the municipality's overall system of governance meaningful, relevant and functional and to avoid the proliferation of a multitude of independent and distinct planning tools. The latter could perhaps be singularly well-designed, but likely to compose a too crowded instrumental framework, which can be redundant and wasteful, inevitably rigid, costly and of little value, since its results are essentially alien to the effective processes of government and management.

With specific regard to the smart city orientation in urban strategic planning, there are many other weaknesses that should be adequately addressed and that concern both methodological and substantive issues. For what concerns the latter, at least two perils must be avoided: first, the excessive generality of the strategic objectives, which is typical of settings that tend to be all-inclusive; second, the opaque definition of the contents of the plan, which is the result of non-rational or non-transparent choices. For what concerns methodological issues, especially the way in which decisions are taken, the main risks and limits concern the only apparent openness of the planning process and the purely fictitious involvement of civil society. This corresponds to a decision-making process that is circumscribed to the narrow political and administrative boundaries or, no less seriously, to a privileged and non-transparent relationship among strong powers.

Another risk that is not to be underestimated is the lack of coherence, both in terms of harmony and synchrony and of horizontal and vertical integration, between the smart orientation of a local authority and (1) that of contiguous territorial contexts (either geographically close or more generally united by the same socio-economic problems) and (2) that of other levels of government (provincial, regional, national). This aspect is particularly important for the urban realities of smaller size, which are increasingly, and per se praiseworthy, testing smart solutions. If the need for an integrated approach is not taken into account, these initiatives might be characterized more by their audacity than by their probability of success.

In all such cases, the governance tools that have been adopted are often only formally 'for governance'. In reality, they are dominated by rhetoric, fashion or fiction, they can be self-referential, shortsighted, emulative, unrealistic, bent to particular interests, and, in any case, unable to contribute to the creation of local public value. In other words, they tend to be irrelevant to the directions of change of the corresponding socio-economic system.

On the contrary, in order to be useful to the development of a smart city, urban strategic planning requires the prior definition of appropriate rules concerning openness, transparency of information and communication flows, solutions for the involvement and participation of social actors, partnership arrangements and the exercise of leadership. These are essential rules to try to reduce and overcome—with the awareness of never succeeding completely—many areas of risk inherent to the innovation process of urban contexts. These risks include actors not being open to dialogue and exchange, information asymmetries, power imbalances, divergence of interests, unstated priorities, lack of resources, inertial activities, and unforeseen emergencies.

Despite these risks, if carried out according to the above-mentioned system of rules, the urban innovation process allows giving answers to the problems that the vast majority of stakeholders consider most appropriate. In other words, it provides answers that are largely shared across all interested actors. In order to do so, it is necessary to build a clear, strong, distinctive and long-term vision and to formulate specifically selected yet at once flexible and adaptive goals and projects.

The result will be a smart agenda for local government that is significant enough to make a difference and streamlined enough to be effectively implemented and shared among relevant actors. This will allow the municipality to mobilize interests, build consensus, attract resources, and produce positive results. The actual achievement of positive results—obtained through the implementation of strategies, the activation of processes of collective learning, the higher cohesion among social actors, better ownership of new policy initiatives, and the progressive realization of the desired idea of smart city—can effectively contribute to the creation of local public value.

References

1. Amato, G., Varaldo, R, & Lazzeroni, M. (Eds.) (2006). *La Città nell'Era della Conoscenza e dell'Innovazione*. Milano: Franco Angeli.
2. Anselmi, L. (Ed.). (2005). *Principi e Metodologie Economico-Aziendali per gli Enti Locali. L'Azienda Comune*. Milano: Giuffrè.

3. Borgonovi, E. (Ed.). (1984). *Introduzione all'economia delle amministrazioni pubbliche*. Milano: Giuffrè.
4. Borgonovi, E. (2004). *Principi e Sistemi Aziendali per le Amministrazioni Pubbliche*. Milano: Egea.
5. Bright, J. (1995). The Smart City: Communications Utopia or Future Reality? *Telecommunications, 29*(9), 175–181.
6. Buccellato, A., Asquer, A., & Spano, A. (2004). *Il Governo delle Aziende Pubbliche*. Milano: Giuffrè.
7. Camagni, R. (1993). *Principi di Economia Urbana e Territoriale*. Roma: Carocci.
8. Capello, R., & Nijkamp, P. (Eds.). (2004). *Urban Dynamics and Growth*. Amsterdam: Elsevier.
9. Carrillo, F. J. (Ed.). (2006). *Knowledge Cities*. Boston: Elsevier/Butterworth-Heinemann.
10. Choi, J. H. J. (2010). The City is Connections: Seoul as an Urban Network. *Multimedia Systems, 16*(1), 75–84.
11. Cretu, L. G. (2012). Smart Cities Design using Event-driven Paradigm and Semantic Web. *Informatica Economică, 16*(4), 57–67.
12. Deidda Gagliardo, E. (2002). *La Creazione del Valore nell'Ente Locale*. Milano: Giuffrè.
13. Doherty, P. (2013). *Smart Cities. How to Build Sustainable and Resilient Environments in an Increasingly Urbanized World*. McGraw-Hill Financial Global Institute. www.mhfgi.com
14. Donna, G. (2010). La Pianificazione Strategica nell'Ambito Pubblico. *Impresa Progetto—Electronic Journal of Management*, *2*. http://www.impresaprogetto.it
15. Edvinsson, L. (2006). Aspects on the city as a knowledge tool. *Journal of knowledge management, 10*(5), 6–13.
16. Ergazakis, E., Ergazakis, K., Askounis, D., & Charalabidis, Y. (2011). Digital cities: Towards an integrated decision support methodology. *Telematics and Informatics, 28*(3), 148–162.
17. Ergazakis, K., Metaxiotis, K., & Psarras, J. (2004). Towards knowledge cities: Conceptual analysis and success stories. *Journal of Knowledge Management, 8*(5), 5–15.
18. Farneti, G. (1995). *Introduzione all'Economia dell'Azienda Pubblica*. Torino: Giappichelli.
19. Fontana, F. (2009). Capitale Intellettuale e Creazione di Valore Pubblico Locale. *Impresa Progetto—Electronic Journal of Management*, *2*. http://www.impresaprogetto.it
20. Fontana, F. (2012). The intellectual capital in urban strategic planning. *Journal of U.S.-China Public Administration, 9*(7), 791–811.
21. Gibelli, M. C., & Curti, F. (Eds.). (1996). *Pianificazione Strategica e Gestione dello Sviluppo Urbano*. Firenze: Alinea.
22. Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanović, N., et al. (2007). Smart Cities: Ranking of European Medium-sized Cities. http://www.smartcities.eu
23. Granelli, A. (2012). *Città Intelligenti?*. Roma: Luca Sassello Editore.
24. Hall, R. E. (2000). The vision of a smart city. In *Proceedings of the 2nd International Life Extension Technology Workshop*, Paris, France, 28 Sept.
25. Harrison, C., & Donnelly, I. A. (2011). A theory of smart cities. In *Proceedings of the 55th Annual Meeting of the International Society for the Systems Sciences*. Hull.
26. Harrison, C., Eckman, B., Hamilton, R., Hartswick, P., Kalagnanam, J., Paraszczak, J., et al. (2010). Foundations for smarter cities. *IBM Journal of Research and Development, 54*(4), 1–16.
27. Hospers, G. J. (2003). Creative cities in Europe. *Intereconomics, 38*(5), 260–269.
28. Kickert, W. J. M. (1997). Public governance in the Netherlands: an alternative to Anglo-American 'Managerialism'. *Public Administration, 75*(Winter), 731–752.
29. Kwon, O., & Kim, J. (2007). A Methodology of identifying ubiquitous smart services for U-city development. In J. Indulska, L. T. Yang, J. Cao, J. Ma, & T. Ungerer (Eds.), *Ubiquitous Intelligence and Computing Proceedings* (pp. 143–152). Berlin: Springer.
30. Landry, C. (2000). *The Creative City: a Toolkit for Urban Innovators*. London: Earthscan.
31. Lee, J., Baik, S., & Lee, C. (2011). Building an integrated service management platform for ubiquitous cities. *Computer, 44*(6), 56–63.

32. Loukis, E., Charalabidis, Y., & Scholl, J. (2011). Editorial of the special issue on digital cities. *Telematics and Informatics, 28*(3), 144–147.
33. Mazzara, L. (2009). *Il Piano Strategico nell'Ente Locale*. Milano: Ipsoa.
34. Moore, M. H. (1995). *Creating Public Value*. Cambridge: Harvard University Press.
35. Mussari, R. (2002). *Economia dell'Azienda Pubblica Locale*. Padova: Cedam.
36. Nam, T., & Pardo, T. A. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th Annual International Digital Government Research Conference*: *Digital Government Innovation in Challenging Times* (pp. 282–291). New York.
37. Naphade, M., Banavar, G., Harrison, C., Paraszczak, J., & Morris, R. (2011). Smarter cities and their innovation challenges. *Computer, 44*(6), 32–39.
38. Passaponti, B. (1986). *I Prezzi Politici nei Servizi di Pubblica Utilità*. Pisa: Servizio Editoriale Universitario.
39. Pendenza, M. (2000). *Cooperazione, Fiducia e Capitale Sociale*. Napoli: Liguori Editore.
40. Pollitt, C., & Bouckaert, G. (2000). *Public Management Reform*. Oxford: University Press.
41. Pugliese, T., & Spaziante, A. (Eds.). (2003). *Pianificazione Strategica per le Città*. Milano: Franco Angeli.
42. Putnam, R. (2005). *The Evolution of Social Capital*. Oxford: University Press.
43. Rur/Censis. (2006). Strategie per il Territorio. In *Proceedings Urbanpromo 2006*. Venezia.
44. Schuurman, D., Baccarne, B., De Marez, L., & Mechant, P. (2012). Smart ideas for smart cities: Investigating crowdsourcing for generating and selecting ideas for ICT innovation in a city context. *Journal of Theoretical and Applied Electronic Commerce Research, 7*(3), 49–62.
45. Sòstero, U. (2003). *L'Economicità delle Aziende*. Milano: Giuffrè.
46. Tanese, A., Di Filippo, E., & Rennie, R. (Eds.). (2006). *La Pianificazione Strategica per lo Sviluppo dei Territori*. Soveria Mannelli: Rubbettino.
47. Toppeta, D. (2010). The smart city vision: How innovation and ICT can build smart, "livable", sustainable cities. *The Innovation Knowledge Foundation*. http://www.thinkinnovation.org
48. Volpatto, O. (1987). *Amministrare Oggi l'Ente Locale*. Milano: Giuffrè.
49. Washburn, D., Sindhu, U., Balaouras, S., Dines, R. A., Hayes, N. M., & Nelson, L. E. (2010). *Helping CIOs Understand "Smart City" Initiatives: Defining the Smart City, Its Drivers, and the Role of the CIO*. Cambridge: Forrester Research Inc.
50. Zangrandi, A. (1994). *Autonomia ed Economicità nelle Aziende Pubbliche*. Milano: Giuffrè.

Performance Measurement in the Smart Cities

Mara Zuccardi Merli and Elisa Bonollo

Abstract A successful smart city needs an adequate performance measurement system to have all the information required to develop an effective involvement of stakeholders. Indeed the concept of smart city is connected no longer just to the presence and use of digital infrastructure but also to the role of human, social and relational capital and to the participation of all stakeholders, who, to be really involved, must be adequately informed about goals, activities and results achieved. In this work, after an introduction of the smart city concept, a new model to measure the performance of a smart city is proposed and the results of an empirical study on a sample of smart cities in Italy and Europe are reported. The empirical study aims to analyze how smart cities included in the sample are used to measure their performance and the capability of the new model to meet all the information needs.

Keywords Smart city • Performance indicators

1 Smart Cities from the Perspective of Participatory Government

The term "smart city" was coined in the 2000s, as part of a marketing idea of US multinationals, IBM and Cisco. In coordinating the marketing of their products and services, they came up with a vision of the perfect city, which featured a high

---

This research work is carried out from all two authors; however the paragraphs n. 2 and 4 can be attributed to M. Zuccardi Merli, paragraphs n. 1 and 3 to E. Bonollo.

M. Zuccardi Merli (✉) · E. Bonollo
Department of Economics and Business Studies, University of Genoa, Genoa, Italy
e-mail: zuccardi@economia.unige.it

E. Bonollo
e-mail: bonollo@economia.unige.it

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_7,

level of automation and "intelligence", thanks to the widespread use of information communication technology (ICT) tools [1].

Over time, smart city has become a widely used term in the vocabulary and in the actions of European Union policy, influencing its priorities as well as the mechanisms to allocate community funds; it has also been the object of interest in management studies.

In the Communication of the European Commission no. 519 of 2009, smart cities are defined as cities based on intelligent networks, on a new generation of buildings and on low emission transport systems that can change the future of our energy.

More specifically, as part of the "Strategic Energy Technology Plan", and then the "Smart Cities and Communities-European Innovation Partnership" in 2012, the European Commission identified the concept of smart city as a place to catalyse progress in energy production, distribution and use, transport and ICT, in order to reduce energy consumption [2].

The Italian Government's "National Research and Competitiveness Operational Programmes of 2007–2013" also headed in the same direction, and identified the area of action of smart cities in operations spanning many areas, including mobility and logistics, health services, education and training system, e-government, cultural and tourism services, energy efficiency, use of renewable energy sources and rational use of natural resources.

On the other hand, in management studies, the significance of the smart city concept is attributed to the vision of the local government, in other words a new way of interpreting the institutional purposes of government, leading to recognition of certain priority areas of action [3].

Still, the literature shows that there is no single, widely accepted vision of smart city.

Hence, in the 2007 study carried out by the Vienna University of Technology in collaboration with the University of Ljubljana and the Delft University of Technology, the smart city has this definition: "a city well performing in a forward-looking way in these six characteristics, built on the 'smart' combination of endowments and activities of self-decisive, independent and aware citizens" [4]. The characteristics that distinguish smart cities are identified in the areas of economy, mobility, environment, people, living and governance, thereby taking in many of the activities of a local government.

Many other definitions have also been published, including:

- the "smart community" where "public administration, enterprises and residents have understood the opportunities offered by IT and attempt to use those technologies to improve their day-to-day life and work in a significant and efficient manner" [5];
- the cities that exploit the opportunities of information technology to promote their economic and social development [6];
- the places where ICT is incorporated into the living and work environments [7];
- the "areas which have the ability to support learning, technological development, and innovation procedures on the one hand, with digital spaces, information processing, and knowledge transfer on the other hand" [8].

These definitions focus on the technological component of smart cities, while other, more recent studies emphasise the need to adopt a different approach, where there is more involvement of citizens. In this case, smart cities are defined as cities where “investments in human and social capital and traditional (transportation) and modern (ICT-based) infrastructure fuel sustainable economic growth and high quality of life, with a wise management of natural resources, through participatory government” [9].

So, the smart aspect is progressively connected no longer just to the presence and use of digital infrastructure but also to the role of human, social and relational capital. It is therefore linked to the active participation, already in the planning phase, of everyone who lives and works in the territory in order to integrate the applications, suggestions and needs of the various actors of the local government context, as part of the strategic path inspired by the smart city vision [10–13].

Empowerment and participation of citizens, enterprises and other stakeholders, in a bottom up approach, therefore becomes an essential requirement for the success of smart cities [12]. If the community is not involved, the “smart” innovation projects remain the dominion of the few and risk being perceived as elitist [10]. Indeed, it is inconceivable to initiate a process that involves major changes in the lives of citizens without them being adequately informed, prepared and motivated as regards the potential and advantages of those changes. Information is the key both to acquire consent and to disseminate the benefits to the whole community.

Local governments that want to be a smart city must act as mediator/coordinator/director/lead actor with the numerous subjects that operate in their territory, including public administrations, enterprises, Universities, research centres, but more than anything else, the community, to be able to implement value coproduction systems in public services [14, 15]. In actual fact, coproduction of value occurs for smart cities as part of the services provided by the local government for the territory, and as a consequence for the citizens, which means the local government is forced to operate in terms of public governance and specifically participatory government [16–21].

Indeed in public administration, participatory government means involving stakeholders—the first being the community—in the activity of producing services, and the programming and control process is one of its inevitable spinoffs. In other words, participatory government means that the local government performs its activity using a different basis from the past, so participatory government leads to the development of:

- participatory planning;
- relational control, expressing the relationships between the various actors involved in the production of public services;
- participatory control, expressing the relationships with citizens in the perspective of getting them more involved in progressive and final monitoring activities.

Lastly, participatory government also requires evolved forms of external accountability on the operations of the local government and on its real capacity to involve the community and stakeholders through the use of appropriate tools in the reporting phase.

Therefore, planning, relational and participatory control and forms of external accountability means the local government that aspires to be a smart city needs to develop a solid performance measurement system in terms of concept, with shared and clear functions.

2 Performance Measurement for Accountability

A successful smart city needs an adequate performance measurement system to have at its disposal the information required to create and develop effective involvement of all the actors operating in its context. Therefore, it is necessary to involve actively stakeholders in the planning phase (as contacts in setting priorities, strategies and goals), in the implementation phase of the prepared smart projects (as partners in the coproduction of public services) and in the reporting phase (as recipients of the communication on performance targets achieved).

The definition of an adequate performance measurement system aimed at activating a procedure to collect and make available data and information is closely connected to the "why" and to the "what" to measure [22].

The prevailing theory is that the "why", in other words the motivations that the local government wants to achieve by collecting and processing performance measurements, is mainly ascribable to the following requirements [23]:

- learning, to support improvement at the strategic and operational level in order to redefine the priorities and solutions adopted by the local government to meet the needs of the community;
- planning and control, to make decisions on the allocation of resources, on making the organizational units and the individual employees responsible for achieving results and on the ways services are provided, evaluating possible outsourcing;
- external accountability to all stakeholders operating outside the local government so the activities performed can be explained and verified as well as its responsibility towards the various stakeholders, so they can judge the work of the local government and take action to influence its decisions.

With reference to the "what" to measure, the prevailing theory is that performance in local government, like in all public administrations is a complex concept due to the variety of activities performed, the impossibility of linking the value of production to earnings achieved, the many interests that gravitate around public administrations and the interconnection of their activities. This makes it necessary to appreciate the concept of performance in multi-dimensional terms, distinguishing between the depth and the span of the performance [24].

The depth of the performance in local governments concerns the different levels in which it can be observed and measured: at the level of local government as a whole, at the level of organizational unit and at the level of the single individual. On the other hand, the span of the performance refers to the possible performance content dimensions, classified in various ways, but generally ascribable to activities, resources, effectiveness, efficiency, outcomes, etc.

As far as the motivations mentioned earlier are concerned, external accountability becomes particularly important in a smart city, because in order to develop involvement of the stakeholders in its smart projects, the local government must provide prompt and transparent communications about the performance achieved.

The stakeholders, potential users of this information, are represented by the actors of the local government context, including other public administrations, enterprises, citizens, etc. These players can take the role of:

- customers/users and/or co-producers of the smart services provided;
- electors who use their democratic vote to express their opinion on the priorities of politicians, therefore also as regards smart projects;
- financing entities through taxation, with the need to be informed on the use of financial resources taken from the public administration.

All of these stakeholders must be given detailed information about the resources used, the activities performed and the results achieved: the more this accountability meets the various information needs, the more the smart city local government strengthens relations with the various actors in its context and the more it increases its social legitimization and the probability of lasting success of its smart city projects (that is not tied to obtaining considerable but occasional financing).

On the other hand, as far as the "what" to measure is concerned—so the depth and span of the performance—the level of the local government as a whole becomes important; this is because, as we stated earlier, the local government plays both the lead role and acts as the director in the smart city.

With reference to the span of the performance, to define the so-called content dimensions, we need to examine some observations. In actual fact, in recent years, studies and research propose the adoption of different performance measurement models for smart cities, characterized by different dimensions to consider as the object of measurement.

A succession of models has emerged over time, and basically they are reported in the 2007 study by the Vienna University of Technology, "Smart cities-Ranking of European medium-sized cities" [4], in the Komninos study published in 2008 [25] and in the research carried out in 2008 by "The European House-Ambrosetti" for ABB [10].

As part of its research performed in 2007, the Vienna University of Technology, in collaboration with the University of Ljubljana and the Delft University of Technology developed a tool to rank the degree of smartness of 70 medium-sized European cities, with a population of under 500,000 inhabitants. Specifically, six performance dimensions were associated with the six characteristics of a smart city already nominated in this research (economy, mobility, environment, people, living and governance): competitiveness, social and human capital, participation, transport and ICT, natural resources and quality of life. In turn, 33 factors were selected to describe the dimensions and 74 indicators were chosen to analyse the performance in each factor. For all the dimensions considered, the purpose of the indicators was to measure and then to compare the impact of the smart projects on the competitiveness of enterprises, on the cultural level and quality of life of the population, on the participation of citizens in public life and on the environmental conditions.

Instead, in his study, Komninos [25] identifies four dimensions of the smart city: three are relative to inputs (skills, knowledge and digital spaces), and the fourth concerns outputs (innovation performance). The purpose is to define an ideal smart city model, pinpointing what makes a local government smart, and therefore what its internal dynamics might be, its weaknesses and the impacts in terms of innovation, economic development and wellbeing of the community.

For each of the dimensions identified, Komninos proposes constructing a total of 35 indicators. These are mainly outcome indicators to measure the impact that the smart city projects may have on the variables characterizing its context: for example, "business R&D expenditure (as per cent of GDP)", "number of incubators (per million of population)" and "researchers in industry and services (per cent of total workforce)", measured before and after the creation of a smart city should increase significantly.

Lastly, the study of "The European House-Ambrosetti" [10] points out three dimensions that express the benefits to citizens deriving from the actual creation of a smart city, in order to evaluate the progress and/or the criticalities detected by the local governments in their pathway of development towards the smartness. The study refers to a representative sampling of the most populous cities in Italy, in which the concept of smartness refers to the ability of the urban fabric—not just the infrastructure and services provided by the local government but also, and above all, the social and economic fabric—to free up and manage resources in an efficient, shared way, also by applying innovative processes and technological options. So, the dimensions that are particularly relevant in terms of influence on the level of smartness are: mobility management, resource management and quality of life for citizens. Each dimension is then associated with a total of nine performance indicators. These performance indicators focus mainly on outcome, to express the benefits that the creation of a smart city should give to citizens. Smart services or infrastructure facilities are not considered in the metric, since according to this research the services offered or specific infrastructure do not necessarily translate into real benefits for the daily lives of citizens. Each performance indicator is then connected with two drivers, considered to be relevant as policy indices for local governments to improve their level of smartness.

The performance measurement models described above have been constructed to compare the smart cities in one country or in several European and non-European countries, or to propose an ideal smart city model, defining its distinguishing features (Table 1). As we said earlier, these models include mainly outcome indicators that, by their nature, involve medium-long term observation and detection times.

The performance measurements identified in the different models do not appear to be exhaustive. Indeed, outcome indicators alone cannot be considered enough in terms of participatory government for adequate external accountability; this requires a continuous interactive process between the various stakeholders, supported by a timely, accurate information system.

Table 1 Performance measurement models of smart city

Models	Vienna University of Technology—2007 [4]	Komninos—2008 [25]	The European House-Ambrosetti—2012 [10]
Dimensions	Competitiveness	Education and skills of the population	Mobility management
	Social and human capital	Knowledge and innovation institutions	Resource management
	Participation	Digital infrastructure and e-services	Quality of life for citizens
	Transport and ICT	Innovation performance	
	Natural resources		
	Quality of life		

Table 2 Proposal of a new performance measurement model

Dimensions	Focus	Types of indicators
Production	Quantity and quality of smart public services provided and resources used	Input Activity Quantitative effectiveness Qualitative effectiveness Efficiency
Technological innovation	Innovative outputs	Activity of innovative outputs Effectiveness of innovative outputs Efficiency of innovative outputs
Quality of life of the community	Living conditions of population and local economic development	Outcome
Eco-sustainability	Environmental impact	Environmental outcome Environmental outcome/ economic and financial variables

In order to have a performance measurement system that can provide timely, accurate, externally-oriented information and to harmoniously combine the strategic and short term aspects, it becomes significant to create a new performance measurement system model that considers at least the following dimensions (Table 2):

- production, focused on the number of smart services provided, on the level of liking by the community, on the resources used (inputs) and, therefore on the construction of indicators relative to the inputs used and the activities performed (in quantitative terms) by the local government, to the quantitative and qualitative effectiveness and to the efficiency;
- technological innovation, to stress the innovative outputs and to measure their effectiveness in terms of improving quantity and quality of smart services, and

the efficiency, comparing the costs incurred with the technological solutions introduced and then arriving at the construction of indicators of activity, effectiveness and efficiency;
- quality of life of the community, both in terms of improving the living conditions of the population and of local economic development thanks to the smart projects, and therefore, arriving at the construction of outcome indicators;
- eco-sustainability, to highlight both the environmental impact of the activity carried out with environmental outcome indicators and any trade-off between these outcome indicators and economic and financial variables.

These dimensions are therefore closely connected. The production dimension allows, in the absence of market exchanges, to program, monitor and communicate the co-production of smart services and the level of efficiency. The technological innovation dimension is connected to the production because it considers the exploitation of the technological skills to generate service innovations for the community served. Finally, the remaining dimensions are linked together, but also to the preceding dimensions as in a smart city the production of innovative services should result in an improvement in the quality of life in term of social, economic and environmental sustainability.

As regards the suggested indicators of the dimensions proposed, they can be classified into 5 types (Table 2):

- input indicators designed to report the amount of financial, human and material resources used for smart services;
- activity indicators, concerning the quantity of smart services provided by a smart local government or the amount of work performed; they can be measured also for innovative outputs;
- effectiveness indicators, measuring the degree which predetermined goals of a particular activity or program are achieved; they are related to the smart city capacity to satisfy citizens' needs expressed in a quantitative and qualitative manner (quantitative and qualitative effectiveness); they can be measured also for innovative outputs;
- efficiency indicators, deriving from the relationship between inputs and outputs, so highlighting the ability of maximizing the quantity and/or quality of the smart services provided in relation to the resources used; they can be measured also for innovative outputs;
- outcome indicators, describing the positive and negative effects on stakeholders; they can be referred to social and economic aspects (outcome) or focused only on environmental aspects (environmental outcome); if we consider the relationship between environmental outcome and economic and financial variables, we highlight the environmental cost/benefit of a smart city initiative.

We must point out that this performance measurement system proposed for smart cities must not consist of simply gathering data. It must be characterized by measurements that express the diversity and complexity of what is being measured, but at the same time are simple to understand, easy to communicate and to satisfy the information needs of stakeholders.

3 An Empirical Study: Evidence from Some European Smart Cities

In recent years, thanks to funds provided by the European Union, and by national and local governments, we have seen an increase in smart city projects. Indeed, many local governments are developing activities to save energy and to produce energy from renewable sources, for sustainable mobility, for improvements to the quality and quantity of public services through the extended use of ICT, all with the active involvement of citizens.

In relation to the description of participatory government provided above (participatory planning, relational control, participatory control and external accountability) and to the connected performance measurement system, this work provides the results of an empirical study whose object is to measure performance in smart cities.

In particular, the study aimed to detect whether achieved performance was measured and communicated in smart cities, and if it was, what defining or differentiating elements can be found in this activity. All of this aims to establish whether it is possible to identify any emerging performance measurement models to use as a reference for external accountability and if these models differ from or are traceable to the model proposed previously.

In other words, the study focused on “how” and “what” smart cities communicate in terms of performance measurement.

The work is developed in two parts and looked at the local governments in Italy as well as at significant smart cities at the European level, carrying out an empirical study in May–July 2013 on the websites of the local governments involved in the study.

At first, the study concerned Italian local governments involved in smart city projects, as:

- winners of invitations to apply for European Union or Italian government funding for smart projects;
- participants in the smart city project entitled “Le città ad alto potenziale di innovazione” (Cities with high potential for innovation) promoted by the Associazione Nazionale dei Comuni Italiani (ANCI—National Association of Italian Municipalities);
- participants in “City Protocol”, an agreement amongst local governments around the world to create the first smart city certification system;
- winners of the 2013 “Smart City Road show” prize linked to the realisation of a project inspired by the vision of a smart city, awarded by Smau Observatory-Milan University of Technology.

This adds up to 24 local governments of very different sizes, with populations of between 27,000 and 2,600,000.

With reference to “how” these local governments communicate the performance measured as part of the smart projects, we found that this took place through:

- dedicated website;
- organisation of and/or participation in workshops and conferences on this theme;
- the use of traditional planning documents (Strategic plan, Forecasting and programming report, Performance plan, Management executive plan) and reporting documents (Annual report, Performance report);
- the use of social reporting (social report, sustainability report, environmental balance sheet).

These channels and document types for external accountability must be chosen in relation to the audience to be reached, the accessibility of information, the attractiveness and the speed of the message. In all cases, the aim is to develop a dialogue with the stakeholders, in the perspective of participatory government [15, 26].

The use of a dedicated website makes information accessible to the community and can become a tool of strong interaction with citizens at a relatively low cost [27]. Conversely, the organization of and/or participation in workshops and conferences make it necessary greater investments in human and financial resources and require more technical information because of the specialist audience of these events [26].

The use of traditional planning and reporting documents for external accountability of a smart city highlights the smart vision of the local government, leading to integration between the smart projects and the other programs of the administration. However the literature shows different opinions on the role of traditional planning and reporting documents as tools of external accountability. Some authors highlight their importance to make local governments accountable [28–30], but others stress their complexity and their not easy availability. On this matter, the social reporting can be an effective tool for communicating results and outcome of the smart city projects in an easy-to-read manner, in order to allow citizens to evaluate the impact of the smart projects on the economic and social context [31–33].

The study showed that only nine of the Italian local governments examined (37 %) have a dedicated website, in other words, one created specifically for communications about smart city projects (Bari, Bologna, Florence, Genoa, L'Aquila, Milan, Pavia, Potenza and Turin).

The way of communicating through speeches and participations at workshops and conferences was the most frequent, probably because the local governments felt the compelling need not only to communicate the performance achieved to the outside world, but also to share and compare with other realities that were heading in the same direction.

Conversely, communication through traditional planning and reporting documents does not appear to be particularly common. Indeed, it was found that objectives and results measured by performance indicators specifically connected to programmes relative to the realisation of a smart city were communicated in the Strategic plan, Forecasting and programming report, Performance plan and/or Performance report, only in five cases (Genoa, L'Aquila, Modena, Reggio

Emilia and Turin). In other fifteen cases, these documents only showed indicators generally referring to typical smart city themes (such as ICT application in providing public services, energy saving, environmental sustainability etc.), as part of the local government's various programmes of activity. Lastly, in the remaining four cases, the traditional planning and reporting documents do not mention smart projects at all or are not available on line.

Most of the Italian local governments involved in the study prepare the "Sustainable Energy Action Plan" and make it available on their website. This plan includes strategies, objectives, timeframes, resources and responsibilities specifically relating to the environment, to energy and to quality of life. This planning document is not compulsory and is prepared as part of a European Initiative called the "Covenant of Mayors", which aims to involve European cities in a pathway to energy and environmental sustainability and that for many of the Italian local governments considered, actually represented their first step towards smartness.

With reference to the social reporting documents, only the municipality of Venice published, in its social report, the objectives and the results relative to projects referring to typical smart city themes (environmental sustainability, energy saving, mobility, e-government).

As regards the "what", the contents of the dedicated websites, the traditional planning and reporting documents and the social reporting documents were studied in order to identify which of the performance dimensions identified in the paragraph above (production, technological innovation, quality of live of the community, eco-sustainability) were considered and what type of indicators were constructed (indicators of input, activity, qualitative and quantitative effectiveness, efficiency, outcome, etc.).

The study showed that only the municipality of Turin regularly publishes and updates a "smart city dashboard" in its dedicated smart city website. This information is summarised, immediate and easy to read; it aims to describe the progress of some variables considered particularly significant as part of the creation of smart Turin. At the moment, these variables are measured by environmental indicators, only concerning the quality of the air and the energy certifications of buildings (for example, "PM10 air pollution index" and "number of energy performance certificates of buildings according to energy class").

Most of the other local governments under observation only report general information about the organisation of events and the realisation of specific smart initiatives aimed at improving the public services provided on their institutional website.

With reference to which performances are measured and communicated in the traditional planning and reporting documents and in the social reporting documents, it emerged that the main dimensions considered are production and technological innovation (Table 3). Conversely, there is poor disclosure of performance relative to eco-sustainability and to quality of life of the community.

In particular, communication regarding the production dimension mainly concerns indicators relative to the activity produced, expressed in quantitative terms, and to the financial resources used (such as "users of bike sharing service", "number of photovoltaic installations at schools", "contributions to encourage the

Table 3 Types of indicators of the local governments examined (%)

Local governments	Production					Technological innovation			Quality of life of the community	Eco-sustainability	
	Input (%)	Activity (%)	Quantitative effectiveness (%)	Qualitative effectiveness (%)	Efficiency (%)	Activity (%)	Effectiveness (%)	Efficiency (%)	Outcome (%)	Environmental outcome (%)	Environmental outcome/eco-nomic financial variables (%)
Bari	16.30	25.68	8.40	1.48	0.49	11.60	0.49	0.00	13.83	21.73	0.00
Bergamo	27.36	24.53	0.94	0.94	0.00	7.55	1.26	0.00	0.94	36.48	0.00
Bologna	11.14	21.75	0.27	0.00	0.00	10.61	0.00	0.00	0.00	56.23	0.00
Brescia	5.30	15.23	0.00	0.66	0.00	66.23	3.97	0.00	0.00	8.61	0.00
Cesena	9.35	19.42	2.88	6.47	0.00	9.35	2.16	0.00	0.72	41.73	7.91
Cosenza	0.00	47.62	0.00	0.00	0.00	42.86	0.00	0.00	6.35	3.17	0.00
Florence	23.47	30.02	0.63	0.00	0.00	21.78	2.54	0.00	1.90	19.66	0.00
Genoa	6.70	27.88	0.00	5.27	0.00	10.32	0.11	0.00	0.00	49.73	0.00
L'Aquila	18.69	38.32	0.00	0.00	0.00	0.93	0.00	0.00	3.74	38.32	0.00
Lecce	0.00	33.33	0.00	0.00	0.00	66.67	0.00	0.00	0.00	0.00	0.00
Livorno	39.49	17.03	0.00	0.00	0.00	41.67	0.36	0.00	0.00	1.45	0.00
Milan	0.00	33.33	0.00	0.00	0.00	66.67	0.00	0.00	0.00	0.00	0.00
Modena	7.07	32.25	0.00	0.18	0.00	28.80	0.72	0.00	1.63	29.35	0.00
Naples	12.98	49.47	0.00	2.29	0.00	2.29	0.00	0.00	2.44	30.53	0.00
Pavia	0.00	25.00	0.00	0.00	0.00	59.62	7.69	0.00	0.00	7.69	0.00
Potenza	22.03	49.57	0.00	0.29	0.00	0.29	0.00	0.00	0.00	27.54	0.29
Reggio Emilia	10.73	39.02	0.98	1.46	0.00	4.88	1.95	0.00	2.93	38.05	0.00
Rome	8.66	46.06	0.00	0.00	0.00	15.35	0.39	0.00	0.00	28.74	0.79
S.Giovanni Persiceto	0.00	18.75	12.50	6.25	6.25	50.00	0.00	0.00	0.00	6.25	0.00

(continued)

Table 3 (continued)

Local governments	Production					Technological innovation			Quality of life of the community	Eco-sustainability	
	Input (%)	Activity (%)	Quantitative effectiveness (%)	Qualitative effectiveness (%)	Efficiency (%)	Activity (%)	Effectiveness (%)	Efficiency (%)	Outcome (%)	Environmental outcome (%)	Environmental outcome/economic financial variables (%)
Senigallia	18.18	9.09	0.00	0.00	0.00	72.73	0.00	0.00	0.00	0.00	0.00
Turin	22.64	55.97	0.63	0.00	0.00	18.24	2.52	0.00	0.00	0.00	0.00
Treviso	47.22	13.10	0.00	0.00	0.00	13.49	0.00	0.00	0.00	26.19	0.00
Venice	18.09	44.33	0.00	0.00	0.00	11.53	0.00	0.00	0.00	26.04	0.00
Verona	19.61	29.02	1.18	0.00	0.00	15.29	0.00	0.00	0.00	30.59	4.31
Mean	*14.38*	*31.07*	*1.18*	*1.05*	*0.28*	27.03	*1.01*	*0.00*	*1.44*	*22.00*	*0.55*
	47.97					*28.04*			*1.44*	*22.56*	

replacement of old electric household appliances"). Technological innovation is also measured mainly by indicators of activity of the public services involved in dematerialisation and e-government projects (such as "number of services with on-line payment", "users of Wi-Fi area", "number of on-line certificates").

Instead, the eco-sustainability dimension is clarified with environmental outcome indicators, generally relating to harmful emissions and energy saving (such as "reduction of CO_2 emissions related to replacement vehicles Euro1", "electricity production from renewable sources", "electricity savings from new led-based public lighting"). As far as this is concerned, we would like to point out the case of the municipality of Cesena, which also reports indicators that highlight the trade-off between the amount of CO_2 emissions saved through the realisation of certain initiatives and the relative cost incurred, providing an indication of the cost required to reduce a single unit of carbon dioxide.

When looking at the quality of life of the community (in terms of economic development and better living conditions for the population), the cases of the municipalities of Bari and of Reggio Emilia stand out. The former constructed indicators, corresponding to smart initiatives, on the increase in employment levels and on the rise in the added value produced, to highlight the positive direct effects on the economic and social fabric of its territory. On the other hand, as part of the "Reggio Smart" programme, the performances to measure and communicate identified by the municipality of Reggio Emilia included the amount of investments in research and development and the number of start-up enterprises established after smart projects were implemented.

In the part of the study concerning experiences at the European level, Amsterdam, Tallinn (Estonia) and Helsinki were examined because they are cases of excellent smart cities.

The study showed that the growth direction of all of these local governments has for some time been implemented as part of the smart city vision, developing broad concordance with all the stakeholders as regards the policies adopted and the actions taken. The planning and reporting documents, always available online, show what are often also disclosure contents, explaining the best practices realised. All of this is to involve all the stakeholders, based on the logic of participatory government.

The City of Amsterdam in particular is one of the most evolved examples of a smart city due to the many smart projects realised; indeed, all the activity of the local government is developed on the basis of programmes inspired by the smart city vision. The website publishes the planning documents on the subject of mobility using electric-powered vehicles, on sustainable urban development, on actions in the environmental area to reduce CO_2 emissions etc. It also includes reports on the results of the main projects realised in recent years, highlighting how smart projects can provide the opportunity for economic development and are also a way of making technological innovation accessible to citizens, hence improving their quality of life.

For their part, Tallinn and Helsinki focus their smart activities on more restricted areas, sustainable mobility and the development of digital urban services respectively. In particular, the municipality of Helsinki, through a dedicated website, provide information on completed and in progress projects, creating an efficient communication channel with its stakeholders.

Therefore, the dimensions monitored in these three cases are basically eco-sustainability, with reference to environmental outcome, and technological innovation.

In conclusion, we can state that the performance measurement by the local governments examined during the study is traceable to the theoretical model proposed, although in most cases the constructed indicators are activity and environmental outcome indicators that refer to the dimensions of production (for Italian local governments) and eco-sustainability (for European smart cities).

4 Conclusion

In recent years, the realities of smart cities—those local governments, on the basis of a smart vision, that adopt "intelligent and innovative solutions" for energy saving and energy production from renewable sources, for mobility, for environmental sustainability and for providing public services in new ways by using ICT—are becoming more and more important. All of this requires the use of considerable human and financial resources but more than anything else the widespread involvement of the community, in other words the development of so-called participatory government.

The reasons behind considering active participation of the community as an essential premise are connected both to the use of large amounts of public resources to realise smart cities (the same resources that may be used to develop other social, cultural projects etc.), and to the need for the various stakeholders to become coproducers of public services precisely due to the diffusion of technological solutions linked to the implementation of smart projects.

For involving the local government stakeholders is necessary the implementation of a performance measurement system, specifically projected for the smart city, in order to plan the activities to put in place and the goals to achieve, monitor their progress and be accountable for the results achieved.

In a smart city, the measurement of the performance should at least concern the dimensions of production, technological innovation, quality of life of the community and eco-sustainability and should be effected through the construction of a set of input, activity, effectiveness, efficiency and outcome indicators.

The empirical study highlighted that the theoretical model proposed is more complete than the ones applied in practice, which seems to focus only on the dimensions of production (for Italian local governments) and eco-sustainability (for European smart cities). Indeed, the study showed that in the traditional planning and reporting documents and in the social reporting documents the Italian local governments gave information mainly on the activities developed for the smart services provided (activity indicators). Conversely, the European smart cities focused on the environment outcomes of their activities (environmental outcome indicators).

Nevertheless the theoretical model can be easily adapted to most of the smart cities taken into consideration by the study and his application may provide more significant results without an excessive extra work as the additional information to be collected are easily available.

References

1. Harrison, C., & Donnelly I. A. (2011). A theory of smart cities. In *Proceedings of the 55th Annual Meeting of the International Society for the Systems Sciences*. Hull, UK.
2. Granelli, A. (2012). *Città intelligenti?*. Roma: Luca Sassello Editore.
3. Schaffers, H., Ratti, C., & Komninos, N. (2012). Special issue on smart application for smart cities–New approaches to innovations: guest editors' introduction. *Journal of Theoretical and Applied Electronic Commerce Research, 7*(3), II–IV.
4. Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanović, N., & Meijers, E. (2007). *Smart Cities. Ranking of European Medium-Sized Cities*. Vienna, Austria: Centre of Regional Science of Vienna, Vienna University of Technology.
5. Morse, S. (2004). *Smart Communities: How Citizens and Local Leaders Can Use Strategic Thinking to Build a Brighter Future*. New York: Jossey-Bass.
6. California Institute for Smart Communities. (2001). *Smart Communities Guidebook*. California: San Diego State University.
7. Steventon, A., & Wright, S. (Eds.). (2006). *Intelligent Spaces: The application of Pervasive ICT*. Berlin: Springer.
8. Komninos, N. (2002). *Intelligent Cities: Innovation, Knowledge systems and digital spaces*. London and New York: Spoon Press.
9. Caragliu, A., Del Bo, C., & Nijkamp, P. (2009). *Smart Cities in Europe, Series Research Memoranda 0048*. Free University Amsterdam, Faculty of Economics, Business Administration and Econometrics (2009).
10. Fondazione Ambrosetti. (2012). *Smart Cities in Italy: An Opportunity in the Spirit of the Renaissance for a New Quality of Life*. ABB-The European House Ambrosetti.
11. Anttiroiko, A. V., Valiamo, P., & Bailey, S. J. (2013). Smart cities in the new service economy: Building platforms for smart services. *AI & Society*.
12. Schaffers, H., Komninos, N., & Pallott, M. (Eds.) (212). *Smart Cities as Innovation Ecosystems sustained by the Future Internet*. White paper, Firewall.
13. Schaffers, H., Komninos, N., Pallott, M., Trousse, B., Nilsson, M., & Oliveira, A. (2011). Smart cities and the future internet: Towards cooperation frameworks for open innovation. In J. Domingue, et al. (Eds.), *The Future Internet*. Berlin: Springer.
14. Fung, A. (2001). Varieties of participation in complex governance. *Public Administration Review, 66*(S1), 66–75.
15. Donato, F. (2010). *Le amministrazioni pubbliche verso logiche di governo partecipato*. Milano: Giuffrè.
16. Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development, 24*(6), 1073–1087.
17. Kickert, W. J. M. (1997). Public governance in the Netherlands: An alternative to Anglo-American managerialism. *Public Administration, 75*(4), 731–752.
18. Kooiman, J., & Van Vliet, M. (1993). Governance and public management. In K. Eljassen & J. Kooiman (Eds.), *Managing Public Organizations: Lesson from Contemporary European Experience*. London: Sage.
19. Ongaro, E. (2009). *Public management reform and modernizations*. Cheltenham: Edward Elgar.
20. Rhodes, R. A. W. (1997). *Understanding Governance. Policy Networks, Governance, Reflexivity and Accountability*. Maidenhead: Open University Press.

21. Hughes, O. (2010). Does governance exist? In S. P. Osborne (Ed.), *The New Public Governance? Emerging Perspectives on the Theory and Practice of Public Governance*. London: Routledge.
22. Lebas, M. J. (1995). Performance measurement and performance management. *International Journal of Production Economics, 41*, 23–35.
23. Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review, 63*(5), 586–606.
24. Bouckaert, G., & Halligan, J. (2008). *Managing Performance. International Comparisons*. London: Routledge.
25. Komninos, N. (2008). *Intelligent Cities and Globalisation of Innovation Networks*. New York: Taylor & Francis.
26. Levi, N. (2004). *Il piano di comunicazione nelle amministrazioni pubbliche. Dipartimento della Funzione Pubblica*. Napoli: Edizioni Scientifiche Italiane.
27. Mussari, R., & Steccolini, I. (2010). Using the internet for communicating performance information. *Public Money & Management, 26*(3), 193–196.
28. Boyne, G., & Law, J. (1991). Accountability and local authority annual reports: The case of welsh district councils. *Financial Accountability and Management, 7*(3), 179–194.
29. Ryan, C., Dunstan, D., & Mack, J. (2001). Local Government Annual Reports: Australian Empirical Evidence on Recipients, Paper presented at the APIRA Conference, July, Adelaide.
30. Taylor, D. W., & Rosair, M. (2000). The effects of participating parties, the public, and size on government departments' accountability disclosures in annual reports. *Accounting, Accountability and Performance, 6*(1), 77–97.
31. Rusconi, G. (1988). *Il bilancio sociale d'impresa. Problemi e prospettive*. Milano: Giuffrè.
32. Steccolini, I. (2004). Is the annual report an accountability medium? An empirical investigation into italian local governments. *Financial Accountability and Management 20*(3), 327–350.
33. Low, W., & Davenport, E. (2001). Parallel lines—The development of social auditing and triple bottom line reporting in New Zealand. In *Proceedings of the Governance and Corporate Social Responsibility in the New Millennium. Governance and Social Responsibility Conference: Proceedings of the 2001 Conference*, Burwood, Victoria.

Empowered Cities? An Analysis of the Structure and Generated Value of the Smart City Ghent

Bastiaan Baccarne, Peter Mechant and Dimitri Schuurman

Abstract Smart cities have gained momentum as a conceptual model which embodies a fresh wave of techno-optimism and emphasizes the positive effects of ICT and other innovative technologies in a city, often in combination with multidisciplinary collaborative partnerships. This article assesses a series of six smart city initiatives within one local city ecosystem by proposing a conceptual framework which is then used to analyze the architecture, value flows and contextual dimensions of the smart city Ghent. The results of our analysis show the multi-level collaborative value creation potential in a smart city and shed light on the complexity of these processes. The main conclusion is that current smart city initiatives face the challenge of evolving from demonstrators towards real sustainable value. Smart cities often have a technological deterministic, project-based approach, which forecloses a sustainable, permanent and growing future for the project outcomes.

Keywords Smart city • Multi-stakeholder network • Collaborative value creation • Living labs • Innovation ecosystems

1 Introduction

Cities are becoming the main locus of society. Worldwide, population has been steadily concentrating in cities. In Europe, more than 70 % of the population now lives in urban areas [1]. These demographic changes have an impact on the way society is being organized. On the policy level, cities are increasingly positioned as the main center of political action. To quote New York's major Bloomberg: "while nations talk, cities act" [2]. Cities indeed play an increasingly important role in the lives of the vast majority of people and are becoming a central platform for

B. Baccarne (✉) · P. Mechant · D. Schuurman
iMinds—MICT, Ghent University, Korte Meer 11, Ghent, Belgium
e-mail: Bastiaan.Baccarne@UGent.be

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_8,

knowledge exchange and value generation. At the same time, we are facing grand societal challenges such as global warming, congested traffic, ecological challenges, aging populations, economic challenges, etc. Although these challenges transcend regions, nations and continents, cities are often seen as the main driver for change and most relevant when it comes to tackle these challenges.

In this article, cities are approached as organic ecosystems, which strive to become 'greener' (with smart energy, smart environments and smart mobility), and more 'liveable' (with smart health, smart education and smart living/working), increasing the overall quality of life for city inhabitants [3, 4]. Recent technological evolutions have also fostered a fresh belief in the positive effects of ICT and other innovative technologies in a city. The combination of smart (technology enabled) solutions to meet the grand societal challenges and the focus on the city as the main driver of change led to the concept of the 'smart city'. Although its definition is still subject of debate, it has been increasingly stimulated by (trans)national governments (e.g. the European Commission) and international networks (e.g. EuroCities) over the past years. The availability of funding and emerging enthusiasm about the first smart city success stories has led to a boost in smart city initiatives worldwide. Despite the support for these initiatives, however, only little research exists on the actual value creation and value creation potential of smart cities.

This article assesses a series of six smart city initiatives in Ghent (Flanders, Belgium) to determine in which way and to what extent public and economic value is being created. First, the article provides a brief overview of the evolution towards a smart city and the different definitions of the concept. Next, we analyze some of the main dimensions which appear in smart city (related) literature and propose a conceptual framework, mapping the different actors and the setup of smart city initiatives. This enables us to assess six smart city projects within the city of Ghent with a focus on how value is being generated and processed.

2 The Journey Towards a Smart City

In the second half of the 1990s, Internet caused huge optimism regarding the possibilities of ICT for the improvement of everyday life. For the relation between the citizens and the city, the most prominent example of this uncurbed techno-optimism was the 'e-government' hype [5]. Although the concept of e-government is steadily fading away on the academic and public agenda, most of the promises related to this concept were not realized. The emergence of the next generation web platforms [6] fostered a new era of promises, this time focusing on the democratic potential (e.g. transparency and participation) rather than on governmental services [7]. Democratization of data, for example, allows increased transparency and stimulates participation and interaction between governments and their citizens. Also important in the evolution towards a smart city is the emergence of new technologies to measure and interconnect different dimensions of everyday life, the so-called internet of things.

Besides changing demographics, politics, technological evolutions and societal challenges, economic reality is changing as well. Especially in the domain of

Table 1 Main catalysts for the emergence of smart cities

Changing demographics	A strong rise of people living in urban areas
Changing politics	Cities becoming central actors for social, economic and political change
Grand societal challenges	Climate, mobility, ecological challenges, aging populations, economic challenges, etc.
Techno-optimism	Internet, e-government, web 2.0, internet of things, (linked) Open Data, etc.
Pressure to innovate	Open innovation, increased competition, innovation spiral, etc.
Policy support	The importance of funding and governmental support
City marketing	'Smart' as an appealing attribute for the city as a brand

new media & ICT, rapid technological evolutions, shorter product life cycles, globalization and increased competition have put high pressure on companies, forcing them to innovate in order to survive. This has led to an 'innovation spiral', which means that ever more innovations come to the market, although this also implies an increasing amount of failures [8]. Frissen and van Lieshout [9] refer to this phenomenon as an 'interesting mix' between massive market failures and groundbreaking innovations. In this context, smart cities are trying to stimulate innovation and tailor innovations to the needs of their citizens by stimulating collaborative development of innovations with multiple stakeholders.

Another catalyst in the emergence of smart cities is policy support. Smart city projects are most often relying on funding (see later). Also, and finally, the notion 'smart' is becoming a popular attribute which a lot of cities want to identify themselves with, relating the phrase 'smart city' to city marketing as well (see Table 1).

3 Defining Smart Cities

Literature on urban development shows various concepts for labeling the integration of ICT in civic planning and management, such as 'intelligent cities', 'digital cities', 'ubiquitous cities' or 'smart cities'. This section elaborates on these closely interconnected concepts.

The concept '*smart cities*', although often used as a marketing concept by both cities and businesses to envision a city of the 'future', emphasizes the growing importance of digital technologies in the city to make it more 'green', more 'accessible' and more 'liveable'. Caragliu et al. [3, p. 50], state that a city is smart when "investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory governance". In other contexts 'smartness' refers to context-aware systems, ubiquitous computing and Internet-of-Things technologies [10].

Other authors use the concepts '*ubiquitous cities*' or '*U-cities*' to refer to "a next generation urban space that includes an integrated set of ubiquitous services: a convergent form of both physical and online spaces" [11, p. 143], emphasizing the importance of involving the citizens in development of U-City services (e.g. Helsinki's Virtual Village, U-Seoul and the Lower Manhattan project) [12]. Finding a match between the needs of the citizens of U-cities and the right ubiquitous services is put forward as a critical success factor [11, 13].

'*Digital cities*' are "extensive information systems (including network infrastructures and applications running on them) that collect and organize the digital information of the corresponding 'physical cities' and provide a public information space for people living in and visiting them" [14, p. 144]. Ergazakis and Ergazakis [15] state that these 'digital cities' should offer innovative services targeting various stakeholders that are inherent to a city environment (administrations, citizens and businesses), focusing on interactions between different city stakeholders [15, 16]. Similar to the notion of digital cities is the idea of '*intelligent cities*', which aims at uniting, promoting, acquiring and stimulating diffusion of information. In order realize this, an 'intelligent city' should develop and implement electronic and digital technologies in the city [17].

In smart cities, these collaborative digital environments facilitate the development of innovative applications, starting from the human capital of the city, rather than believing that technology as such can transform and improve cities. Another important dimension is the collection of all sorts of data and information through sensors and sensor networks. Under the moniker 'Open Data', this information is made public and put to use in 'smart city' applications and technologies that visualize, transform and utilize this data [18]. Smart cities focus on the involvement of all relevant stakeholders, whereas 'digital cities', 'wired cities' or 'ubiquitous cities' stress the presence of technological infrastructure. In other words, a city needs to be 'digital', 'wired' and 'intelligent' in order to become 'smart', although being 'digital', 'wired' and 'intelligent' does not automatically imply that the city will become 'smart' by itself.

While both research and policy often promise disruptive solutions, improvement of life in the city and economic growth, there is a vast lack of evidence concerning the actual value that is being created in a smart city and the processes that allow the exchange of value and knowledge. In this article, a smart city is considered as a collaborative ecosystem allowing for the co-creation of sustainable, future proof innovations that improve life in the city and boost the economy, in which technology plays an enabling role. Because it is often difficult to assess or define this concept in actionable, tangible elements, we will make this assessment based on six smart city projects in the city of Ghent.

4 A Framework for Analyzing the Structure and Generated Value of a Smart City

4.1 Smart Cities as an Ecosystem

The collaborative nature of smart cities is related to the Living Lab-concept and the quadruple helix-models for innovation. Triple and quadruple helix-models, deal with collaboration between universities, government, industry, and end-users, in this context citizens [19]. Co-operations like these have been claimed to facilitate exchange of ideas and technologies, with fewer barriers between academia, end-users, policy and industry [20]. This approach is very similar to Living Lab literature. Living Labs are ecosystems in which end-users and other relevant stakeholders are involved in the development of an innovation over a longer period of time using a combination of different research methods, following an iterative process [21]. Living Labs facilitate university–industry relationships, but also relationships between large companies and SME's, start-ups, entrepreneurs, and, most importantly, involve the end-users themselves, commonly referred to as public–private–people partnerships (4P's) [22]. Various Living Lab authors stress the importance of collaboration and knowledge support activities as cardinal to a successful Living Lab [23, 24]. These collaborative ecosystems promise to contribute to the facilitation of knowledge exchange among the ecosystem actors. In line with the above collaborative ecosystem literature, this chapter conceptualizes knowledge as both information (e.g. data), expertise (latent) and skills (e.g. coding).

Cosgrave et al. [25] connect the multi-stakeholder aspect of Living Labs to the concept of '*innovation districts*', small regions which cluster innovative actors such as start-ups, creative industries and venture capitalists. These "pockets of growth" are characterized by inter-firm collaboration and governmental support. In EU programs such as i2010 and Europe 2020, the importance of smart cities is highlighted, and the Living Lab-approach is considered a best practice in this context as it enables structuring user interaction by keeping users continuously involved in making better products and services while their expectations are continuously monitored and reflected upon in a systematic process [26]. Consequently, collaboration between all smart city stakeholders requires a user-driven and user-centric research approach to replace a more technology-centric approach. Based on these concepts, the proposed smart city framework includes four types of actors (1) *policy*, (2) *citizens*, (3) *research* and (4) *private partners* (see Fig. 1).

4.2 Policy

The policy actor is present at several levels. The most active policy level is the city government, but smart cities are also being supported on a regional level and on the (inter)national level as well. Smart city initiatives help these governmental levels to reach policy goals. An important actor is the European Commission, which put the idea that European cities should become 'empowered' or 'smarter' forward as one of the core inspirations of the European Digital Agenda, "which seeks to recognize the power of urban planning and the role of ICTs in managing infrastructures" [27]. Horizon 2020, the EU's new research program for 2014–2020 encompasses a €80 billion package for research and innovation funding. Horizon 2020 will support the development of ICT in Science (in future and emerging technologies or e-Infrastructures); in industrial leadership (such as smart systems, robotics, photonics, etc.) and in societal challenges (such as eHealth, eGovernment and eSkills.) Also, international organizations such as the OECD (Organization for

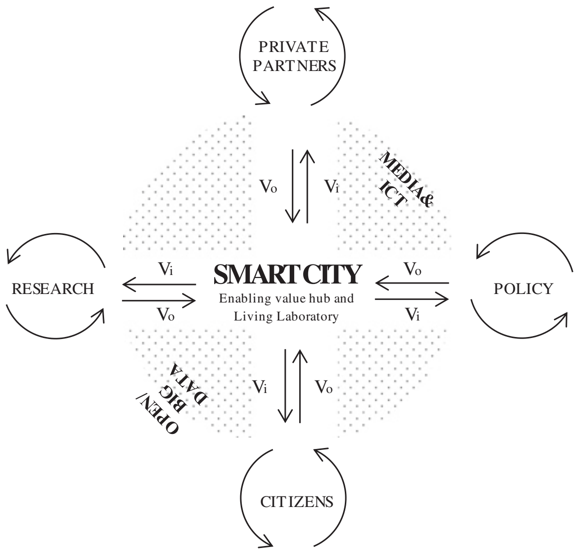


Fig. 1 Conceptual model of a smart city

Economic Co-operation and Development) [28], and UNESCO [29] have started to promote open access to information and knowledge, thus stimulating open innovation and smart city initiatives.

4.3 *Citizens*

End-users and citizens have been increasingly emancipated on different levels. In the domain of new product development, R&D departments rely increasingly on user input and collaboration (e.g. the use of *Lead Users* [30]). Innovations are no longer (solely) developed top-down, but are increasingly shaped and molded bottom-up [31]. End-users and citizens have also become emancipated when it comes to the creation and distribution of products, services and media themselves, indicating a power shift from traditional industries towards the people [32]. Another evolution that supports more citizen or user-centric paradigms and projects is the criticism on technological-deterministic discourses [33]. In smart city projects, one of the challenges is to transcend the technological-deterministic discourse by actively involving all stakeholders that can provide substantial input for developing a more accessible, information based, interactive and participatory urban environment.

In this context, *Web 2.0* is an important medium that creates a new degree of agency in constructing engagement with online resources, with other internet users, with open innovation [34] and with 'collective creativity' [35]. Web 2.0 also demolishes the idea that innovation is a proprietary activity conducted inside organizations in series of managed steps and entails ceding control over decisions about the content of products or services to networks of (online) citizens who interact with one another. Web 2.0 is "characterized by new forms of interaction with users who now play a key role in the content-creation and innovation processes" [36, p. 43], and consists out of a set of tools and a collection of social processes originating out of online communities and networks [37].

4.4 Research

At the academic level, smart cities have been looked at from different domains and backgrounds. It is a cross-disciplinary concept which covers urban studies, economics, political studies, city planning, engineering, sociology, communications as well as user research. This is one of the main reasons why it is difficult to find consensus on the actual definition of the concept. When it comes to the role of the research actor in the smart city ecosystem, [38] consider academic researchers as a necessary actor because they provide expertise in user research and knowledge. The triple and quadruple helix concepts also stress the importance of universities as a distinct actor in the innovation ecosystem [19, 20, 25, 39]. Moreover, the contribution of academia is not limited to user research; it can also include research on technical topics or policy and business related issues.

4.5 Private Partners

Innovation is becoming increasingly important for companies to remain competitive. However, high flop-rates still illustrate the need for an adequate management of innovation, which includes selecting the right tools and methods in order to structure and optimize innovation processes [40]. Traditionally, innovation was viewed as an inherently closed process with most operations running inside the boundaries of the company and R&D processes taking place in secretive in-house laboratories. More recently, this closed, vertically integrated model has been challenged and replaced by a distributed view on innovation and innovation management [41]. Smart cities serve as an innovation broker, connecting different stakeholders, allowing for real-life validation, ideation and co-creation. They create a framework for open innovation [34] continuous innovation [42] and systemic innovation [23].

In the smart city as an ecosystem, value and affordances flow between the different actors (see Fig. 1, indicated as V_i and V_o). In our conceptual framework value consist out of two dimensions: *socio-economic value* and *affordances*. Affordances can be conceptualized as 'what one system provides to another system', in the case of this article, as what a city system provides to its users, its citizens or other smart city actors. An affordance also encompasses the perceived functional significance of that system for an individual. For our purposes, we use the definition of affordances by Norman [43], describing them as: “the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used” (p. 9). An illustration in a smart city context could be the online co-creation of a city service, which enables citizens to interact with their city (affordance 1), opening this process to local entrepreneurs, which enables them to generate business out of this (affordance 2). In our analysis, these affordances are approached as ‘enabling dimensions’.

A relevant conceptual model for ‘*value*’ can be found in literature on business modeling. For example, [44] proposes a tool to model the relationship between the value of new ICT products or services and the control over new ICT products or services. In our analysis, a distinction is made between the generation of public value (e.g. safer streets) and economic value (e.g. generation of revenue). The concept of ‘*public value*’ refers to the value that is generated through the creation and implementation of services and technologies that adequately harness opportunities within the city, tackle societal challenges and/or realize policy goals [33]. It refers to, for example, reducing traffic jams, emancipating citizens, increasing neighborhood cohesion, etc. ‘*Economic value*’ on the other hand covers economic metrics such as the annual economic growth of cities and companies within the city, a decrease in unemployment, the extent to which new businesses (start-ups) are being generated and able to survive, a reduction of bankruptcies, an increased competitive advantage, attracting existing businesses to the city, etc.

As discussed in the introduction, two other frequently occurring smart city attributes are the use of technology (ICT, internet of things, etc.) and the integration of Open/Big Data. Therefore we also take these contextual dimensions into account when analyzing smart city ecosystems.

5 Methodology

In the next section, we will apply this conceptual framework on the city of Ghent as a smart city ecosystem. Because of the long-term nature of smart city projects, the exploratory nature of our research a multidimensional comparative case-study analysis seems the most suitable approach to make the assessment [45]. Case study research excels at bringing an understanding of a complex issue and can extend knowledge or add strength to what is already known through previous research. On top of that, case studies are most suited for processes which are poorly understood and lack a (solid) theoretical foundation [46], allow to analyze the process open-ended and on multiple levels [45] and gain deeper qualitative insights. Yin defines the case study research method as an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used. Given the complexity of the studied phenomenon, the multiple levels of analysis and the participation of the author team in the studied smart city projects, the multidimensional comparative case-study design seems most appropriate.

For our case study analysis, six smart city projects were selected using three criteria, the project had to (a) take place in the city of Ghent; (b) be referred to as a 'smart city project' in the project documents and (c) have a collaborative nature. Both finished projects and running projects were taken into account. As research partners in the selected projects, we were able to use both research results (documents) as well as our own experiences (participatory observation/action research) and lessons learned (soft data). The following hard data sources were used for our analysis: (a) meeting reports of steering committees, (b) the initial project proposal and project reports and (c) deliverables from the projects.

The presented conceptual model is applied in three ways. First, the ecosystem architecture (actors) is studied for each of the six smart city projects. Next, the incoming and outgoing value is studied. Finally, the six cases are analyzed on eleven parameters;

- involvement of the full smart city ecosystem
- intensity of the network collaboration
- reuse of knowledge
- importance of Big Data
- importance of Open Data
- importance of technology
- generated economic value
- generated public value
- potential for civic engagement
- knowledge valorization
- sustainability.

5.1 Research Context

The City of Ghent has developed a long-term strategy until 2020, comprising five strategic goals. Knowledge and innovation is one of these goals. In the light of the development of smart cities and the empowerment of smart citizens, a long-term strategic program 'Digitaal.Talent@Gent' has been set up. This program supports the strategic mission of the administration and the city council: "Ghent, a creating city in the development of a sustainable, solidarity and open society by uniting all creative forces" [47, p. 1] . In this regard, different objectives have been formulated around 'knowledge, innovation and creativity', 'social sustainability', 'economic sustainability' and 'ecological sustainability'. In specific, Ghent is involved in setting up open platforms to help develop innovation ecosystems (for and by active user involvement) accelerating the move towards smart cities and providing a wide range of opportunities for sustainable services that are developed, implemented and used for and by citizens and businesses as co-producers.

5.2 Selected Smart City Projects

Citadel.[1] Citadel (on the Move), is a European project that aims to make it easier for citizens and application developers to use Open Data to create innovative mobile applications they want and need. Currently, open governmental data is often difficult to access and use, even by the developer community. Citadel aims to lower this barrier by (a) creating formats that make it easier for local governments to release data in useable, interoperable formats and by (b) providing templates that simplify creating mobile applications. These templates should provide a simplified route to smart service development for non-developers who have great service ideas.

Ghent Living Lab.[2] Ghent Living Lab (GLL) is an open collaborative network led by the City of Ghent. Key partners include the local government and its service partners, iMinds (Flemish organization that supports innovation in media and ICT), all major colleges and universities in the city and local (developer) networks and community organizations. GLL acts as a facilitator between the different parts of the collaborative network that has been established between the research community, businesses, the public sector, citizens and the wider community. Its primary focus is on smart cities and the development of Future Internet related services to support the further development of smart cities. GLL serves as a learning platform and as a test and development environment in a real-life environment. In this way, GLL becomes a tool to work with researchers, entrepreneurs, citizens, digital creative forces and the City of Ghent on joint trajectories in function of product development, research, service delivery and policy strategy. GLL is also an effective member of the European Network of Living Labs.[3]

Zwerm.[4] Zwerm was part of the European project SMARTiP.[5] This city intervention/game took place in two neighborhoods in Ghent. It wanted to support 'smart engagement' and establish a meaningful and stimulating contact between citizens and their neighborhood. Zwerm had two overarching objectives: (a) activate citizens around urban places of interest and motivate them to carry out assignments that are beneficial to the community, meanwhile emphasizing neighborhoods as the place where citizens can meet each other, and (b) encourage a better take-up and use of ICT while helping to develop the information society.

Mijn digitaal idee voor Gent[6] **(MDIVG)**. In the same SMARTiP project, a crowdsourcing platform was launched to gather and generate 'wild' ideas on smart engagement, but also on mobility and environmental solutions for cities. MDIVG

[1] http://www.citadelonthemove.eu/

[2] http://www.ghentlivinglab.be/en

[3] http://www.openlivinglabs.eu/

[4] https://www.zwermgent.be/

[5] http://www.smart-ip.eu/

[6] For academic research on this project see [65, 66].

involved about 5,500 citizens, the city of Ghent and iMinds (looking to benefit from the crowd input). The intermediation platform or ‘crowdsourcing enabler’: Mijndigitaalideevoorgent.be enabled selective and creative crowdsourcing (see [48]) based on the proprietary software of UserVoice (http://www.uservoice.com). The platform was open to answers on the question ‘How can ICT make it even more pleasant to live in Ghent?’ between April 1st and May 15th 2011. In this period the website was visited by 5,451 unique visitors and counted 17,873 page views. More than 1,400 people registered their e-mail on the platform, enabling them to submit an idea or cast votes on already submitted ideas. A total of 128 ideas were submitted, which received more than 4800 votes.

Apps for Ghent.[7] This hackaton event, organized by the City of Ghent, iMinds Multimedialab (Ghent University) and OKFN[8] wanted to stimulate the use of Open Governmental Data provided by the City of Ghent. By doing so, the city wanted to increase governmental transparency and stimulate citizen entrepreneurship. The idea is that by providing both professional and amateur developers with data, it will fuel the creation of innovative applications. The event was a ‘hackaton’ where developers are challenged to create the best application. Participating teams were brought together and allotted a fixed timeframe to develop a prototype or mock-up of an innovative application within a city context. The winning team of the 2012 edition created an application which connects neighborhoods to cultural activities.

Future legends.[9] Future Legends was a Living Lab project, instigated by the City of Ghent, which capitalizes on the lifestyle of Flemish young people from urban areas and the outskirts of the city. These young people are often low skilled and mostly, but not exclusively, of immigrant origin. Research of the REC Radio Centre and the VRT Radio showed that these ‘urbans’ show limited engagement with the mainstream media landscape. In other words, their own rhythm of life requires an own media pattern and offerings. The bottom up ‘Future Legends’-project resulted in a media platform called ‘Chase—Music From Scratch’ (www.chase.be). This online radio station offers youngsters a platform to express their creativity by participating in the show. Together with professional artists they can compose a playlist and air their own creations.

6 Application of the Framework, Case Study Ghent

In this section, we apply the proposed smart city framework on the different actors for each of the projects (Table 2). We then analyze the different flows of value and affordances between these actors (Table 3). Finally, we compare the selected smart city projects on the different dimensions described before (Table 4).

7 http://appsforghent.be/

8 http://www.okfn.be/

9 http://www.mediatuin.be/projecten/future-legends

Table 2 Application of the framework: actors

	Policy	Citizens	Private partners	Research
Citadel	City government + EU (funding & collaboration)	Low involvement, mainly citizen developers	Providing programming services	Social sciences
Ghent living lab	City government + EU (funding) + ENoLL (EU network)	Focus on citizen participation and empowerment	Involved as project partners (mainly ICT)	Multidisciplinary (technical, creative and social)
Zwerm	City government + EU (funding)	Citizen as a topic of research and co-creator (two neighborhoods)	Providers of technology	Social sciences
Mijn digitaal idee voor Gent	City government + EU (funding)	Citizen as an external source of information	No involvement	Social sciences
Apps for Ghent	City government	Citizen developers	Invited to participate, sponsor and scout talent	Engineering (organizational support)
Future legends	City government + public organizations	Urban youngsters	No involvement	Social sciences

The policy level plays a central role in all six cases, especially the city government, which is part of every project in our selection. The policy actor uses funding schemes and collaborative partnerships to meet its policy goals. Besides the city government, smart city projects are supported by the regional government (Flemish government, through IWT and IWT funded organizations) and the transnational government (European Commission, through special project funding) as well.

Smart city projects do not always approach citizens in the same way. This difference exists on two levels. First, the selected projects *target* different populations. In only two projects, all city inhabitants are being targeted. The other projects target specific niche groups in the city. This subset of citizens can be determined by geographical parameters (such as Zwerm, which was targeted on two neighborhoods), by skills (such as citizen developers) or by socio-demographic profile (such as the Future legends project). Second, *citizen involvement* can be of a different nature. Citizens can be approached as a source of information (GLL, MDIVG), as a provider of services (Apps for Ghent, Citadel) or as a research subject (Zwerm, Future Legends).

Private partners are not always involved in smart city projects (MDIVG, Future Legends). If commercial enterprises are involved, they are providing technological *infrastructure* (Zwerm), *services* (Citadel) or they function as a *partner* that can potentially benefit from the project (GLL, Apps For Ghent). In some of the projects, research actors are *part of the project* and have their own work packages, central in the project (Citadel, Zwerm). In other projects, research activities should rather be considered as a *side track* of the project (GLL, MDIVG, Apps For Ghent, Future Legends). For the latter, research partners only use the data which is generated within the project for academic analysis, but the research results are not being processed within the project.

Table 3 provides an overview of the incoming (Vi) and outgoing (Vo) value flows between actors in a smart city ecosystem. Most of the flows enable other actors to perform certain tasks and can therefore be considered as *affordances*. These enabling flows or affordances differ in nature. Within the studied projects, we distinguish enabling *funding* (financial support), enabling *environments* (an ecosystem or working space), enabling *services* (activities), enabling *knowledge,* (data and expertise), enabling *networks* (brokerage) and enabling *policy* (stimulation by policymakers). Different scenarios exist in which the four smart city actors play a different role in transforming one affordance to another.

Each 'chain of affordances' starts with the instigation of the policy actor (city), which seeks for enabling funding at the European level (except Future Legends, which was supported by city resources), brings together relevant smart city actors (enabling network or environment) and sometimes provides enabling information (e.g. open governmental data). Next, each case follows a distinct chain of affordances, depending on the goals and configuration of the project. *Cities* pursue creation of public and economic value but only generate public value themselves (implementing city improvements). C*itizens* pursue creation of public value and also generate public value themselves (creating apps or services). *Private partners* pursue creation of economic value but generate public value instead (creating free apps and services). R*esearchers*, finally, pursue creation of public value, but generate no value directly, since this actor only plays an enabling role. Overall, the potential generation of economic value is not (yet) realized.

Besides the different roles in a smart city and the flows of affordances between them, some higher level units of analysis remain to be tackled. Table 4 shows the results of our multidimensional comparative case study analysis. We distinguish four main clusters (a) the collaborative nature of a smart city, (b) the role of knowledge and technology, (c) the overall creation of value, and (d) the future of smart city initiatives after the project ends. The performance levels were coded by the author team, based on project documents and insights gathered through project participation.

6.1 The Smart City Ecosystem

The first dimension assesses whether the full smart city ecosystem is involved in the project or not. As was discussed above, an important element in smart cities is the way research, policy, private partners and citizens collaborate and share knowledge and services in order to optimally develop future products and services with a high sustainability. Nevertheless, only three out of six projects involve all four smart city actors. On top of that, one of these two (GLL) has only set up this collaboration on paper and has not yet rolled out full collaborative projects. The role most often

Table 3 Application of the framework: flows

		Policy (city government)	Citizens	Private partners	Research
Citadel	Vi	Applications for an improved city environment (public value and potential economic value) and policy advise (enabling knowledge)	Opportunity to create own applications (enabling environment, enabling knowledge and enabling policy)	EU funding (enabling funding) and research insights (enabling knowledge)	EU funding (enabling funding)
	Vo	Open governmental data (enabling knowledge)	Applications for an improved city environment (public value and potential economic value)	Creation (enabling service) of an empowering platform (enabling environment)	User research (enabling service) + policy and development advise (enabling knowledge)
Ghent living lab	Vi	Ideas of citizens and SME's towards policy (enabling knowledge) + potential economic and public value creation	Potential to transform ideas into reality (enabling network)	Easy access collaboration with policy, citizens, other SME's and research (enabling network) + ideas & user insights (enabling knowledge)	Potential use cases (enabling network and enabling knowledge)
	Vo	Brokerage/networking ideas with creators (enabling network)	Ideas, feedback, suggestions, opinions, willingness to collaborate (enabling knowledge)	Economic growth/entrepreneurship, (potential economic value) products & services (potential public value)	Enabling user research and product development for non-academic projects (enabling service and enabling knowledge)
Zwerm	Vi	Policy advise (enabling knowledge)	Experimental socio-technical environment (enabling environment)	EU funding (enabling funding) + Information on economic potential for enabling technology (enabling knowledge)	EU funding (enabling funding) + Raw data on citizen behavior (enabling knowledge)
	Vo	Facilitating the city as a laboratory (enabling environment)	Participation and behavioral data (enabling knowledge)	Provision of technical components (enabling environment)	Knowledge on human behavior and policy advice (enabling knowledge)

(continued)

Table 3 (continued)

		Policy (city government)	Citizens	Private partners	Research
Mijn digitaal idee voor Gent	Vi	Ideas of citizens (enabling knowledge)	Empowerment platform (enabling environment)	–	Reuse of citizen input for academic analysis (enabling knowledge)
	Vo	City improvements (enabling policy) + city improvement (public value)	Ideas (enabling knowledge)	–	Insights on citizen participation (enabling knowledge)
Apps for Ghent	Vi	Applications for an improved city environment (public value and potential economic value)	Open governmental data (enabling knowledge)	Open governmental data (enabling knowledge)	Low
	Vo	Open governmental data (enabling knowledge) and stimulate app development (enabling policy)	Apps based on governmental Open Data (public value and potential economic value)	Apps based on governmental Open Data (public, and potential economic value)	Low
Future legends	Vi	Policy advise (enabling knowledge)	Experimental environment (enabling environment)	–	Raw data on citizen behavior (enabling knowledge)
	Vo	Project funding (enabling funding)	Project participation (feedback + behavioral data) (enabling knowledge) and creation of their own radio service (public value)	–	Knowledge on human behavior and policy advice (enabling knowledge)

Table 4 Multidimensional comparative analysis of six smart city projects

	Citadel	Ghent living lab	Zwerm	Mijn digitaal idee voor Gent	Apps for Ghent	Future legends
Involves total smart city ecosystem	Yes	Yes	Yes	No	No	No
Network collaboration	Medium	t.b.d.	Medium	Medium	Medium	High
Reuse of knowledge	Yes	No	Yes	No	No	No
Importance of big data	Medium	Low	Low	Low	Medium	Low
Importance of open data	High	Medium	Low	Low	High	Low
Importance of technology	High	Medium	High	High	High	Low
Created economic value	t.b.d.	Low	Low	Low	Low	Low
Created public value	t.b.d.	Medium	High	Medium	Medium	High
Potential for civic engagement	High	High	High	High	High	High
Knowledge valorization	t.b.d.	Medium	Medium	Medium	Medium	High
Sustainability	t.b.d.	Medium	Low	Medium/high	Medium/low	High
Potential for economic growth	High	High	Low	Low	High	Low
Importance of funding	High	Medium	High	Medium	Low	Medium

neglected is that of the private partner. This is challenging when the aim is to create economic value and forecloses the sustainability of the developed products and services. Without a private partner, smart city projects have to rely on 'citizen entrepreneurs' or continuous project support by the city government.

6.2 Collaborations

Besides the involvement of all four smart city actors, it is also interesting to elaborate on the intensity of the collaboration between smart city project partners. The downside of involving the full ecosystem is that collaboration between partners becomes much more difficult and more likely to be less intense. In the selected projects, the city government always acts as the main project coordinator, determining the degree of interaction with the other three actors. Overall, the intensity of collaboration is rather high, which can be explained by the policy goals, which focus more on the collaborative dimension of smart cities than on the technology dimension. For Apps For Ghent, GLL and MDIVG, the main reason for a medium rating on collaboration is the lower interaction with research partners, which are either only using the generated data for academic purposes (MDIVG), only involved for the promotion of the research group (Apps For Ghent) or, as is the case for GLL, have not yet had the chance to collaborate in one of the projects.

6.3 Reuse of Knowledge

This observation brings us to another interesting dimension: the reuse of knowledge. While an increasing amount of smart city projects are being set up, all focusing on efficiency and sustainability, the question rises whether each of these projects generates new knowledge. From this perspective it is important to build upon previous projects and related knowledge. Reuse processes are considered increasingly important for developing high-quality software and ICT projects. As explained by [49], reuse processes can play a crucial role in the success of private entrepreneurial initiatives as well public projects.

Reuse is critical, as it allows working on existing artifacts instead of starting from scratch, thereby enabling the development and deployment of software and services with greater ease. Consequently, time and human effort required to develop software product and pilots can also be effectively reduced. Given the financial crisis across Europe, reuse of ICT-based pilots and products can effectively add to the cost-cutting measures proposed by the public and private bodies. In addition to this, iterative reuse can also have a relevant, verifiable impact on product productivity and quality, as reusing existing artifacts can iteratively improve the quality of the software or pilot. Nevertheless, our analysis shows that only two of the selected smart city projects incorporate reuse of knowledge. Citadel and Zwerm are both part of collaborative European projects in which the reuse of the infrastructure and system logics in other cities is one of the main goals.

6.4 Importance of Big Data

As our society becomes more digital, with key drivers such as social media, mobile devices and sensor networks, we notice a tremendous growth of generated data. This trend is often defined with the phrase ‘Big Data’. There are numerous definitions for the term ‘Big Data’. However, most authors agree that Big Data is a loosely defined term to describe data that has become so large and so complex that they are difficult to process using standard (statistical) software and databases [50]. The analysis of Big Data can help people interact in a more flexible and adaptive way with their environment [51, 52]. Big Data can be a source of competitive advantage presenting new opportunities to create new business models to monetize data or to customize services to individuals. However, Mantelero [53] also points out that these huge amounts of data represent a strategic and economically relevant asset resulting in a centralized power held only by a few subjects. In the context of smart cities, Big Data can be approached as a valuable resource connecting the dimension 'reuse of knowledge' and 'Open Data'. Smart cities often produce huge amounts of data, be it by opening up (governmental) datasets, sharing research results or capturing data by sensors placed throughout the city. In order to optimally tap into this source of raw information, smart city projects must find a way to cope with Big Data. In our analysis, only two out of six smart city projects take this challenge more or less into account (Citadel and Apps For Ghent). Both projects focus on transforming raw data into actionable services and understandable visualizations. Given the increasing importance of this dimension, there are various opportunities for future smart city projects to focus on harnessing this largely untapped potential.

6.5 Importance of Open Data

Open Data is related to the idea that certain data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control. The goals of the Open Data movement are similar to those of other 'Open' movements such as Open Source, Open Content, or Open Access. The term 'Open Data' itself is recently gaining popularity with the rise of the Internet and World Wide Web and, especially, with the launch of open-data government initiatives such as Data.gov. In order to become more innovative and transparent, Public Administrations worldwide are starting up Open Data Portals stimulated by the idea that open government data (OGD) can open up economic opportunities, can promote transparency and accountability or can support the reform of public services and innovation [54]. Similarly, the EU's 'Open Data Strategy for Europe' emphasizes the fact that public administrations are sitting on a goldmine of unrealized economic potential. Therefore, it is not surprising that Open Data is a central dimension in two out of six smart city projects (Citadel and Apps For Ghent). Open Data is increasingly becoming important for smart cities. The market value of the reuse of public governmental data in the European Union is estimated at €27 billion, each year [55]. Similar to Big Data, this potential is nevertheless largely untapped and the actual economic valorization of this estimation still remains to be proven. We should also be vigilant of privacy and security issues concerning open data since these data sources can also be used for malicious purposes. More specifically, triangulating different data sources can pose a threat for the privacy of the individual and revealing governmental data might help to find weak spots in security systems.

6.6 Importance of Technology

A lot of smart city projects have a technological-deterministic nature. They build upon the belief that (new) media and ICT solutions can improve life in the city and that technology is the main driver to solve the complex societal challenges we face in contemporary cities. New technologies allow for rapid, distributed, contextual and personalized information exchange. It connects information from public organizations and becomes remixed, annotated and redistributed by the citizens (an informal network of people). These socio-technological evolutions fostered a strong belief in the possibilities for smart cities. The central position of technology is also present in all selected smart city projects except one (Future legends). Whereas technology certainly enables a lot of new opportunities, it is dangerous to believe that technology as such is sufficient to create a smarter city. This potential can only be harnessed if it is embedded in a social context. Technology can support city innovations, but to think of it as the main driver of social change is only a one-dimensional point of view. In order to overcome the short-term nature of smart city projects and have impact over a longer period of time, the social context should be central in smart city projects. Of our selected cases, Future Legends is the only project which became autonomous after the project ended. Not surprisingly, this project was the only one which used technology merely to serve social innovation.

6.7 Economic Value

In the end, smart city projects aim to generate economic and/or public value. While this is often part of the project legitimations when applying for funding, especially for the European Union, none of the selected cases was so far able to generate any substantial economic value. Although the value of Open Data and open collaborative innovation ecosystems is often put forward as a huge source of untapped potential, reality has not yet provided any substantial proof for this. This is one of the biggest challenges for smart city projects. If these projects are not able to boost economy or even be economically successful to be able to become autonomous, smart city projects will always have to rely on governmental support and funding.

6.8 Public Value

Besides monetary value, the generated value can have a public nature as well. Especially when supported by public resources, this might also be a valid project legitimation. Although the concept of public value is much harder to assess, the selected smart cities projects tend to generate at least some public value. For Zwerm, this value was validated trough academic research, confirming that the project had improved social cohesion in both neighborhoods [56]. The Future

Legends project resulted in both policy advise on the stimulation of culture participation for urban youngsters as well as a community driven crowd sourced radio station [57]. For the other smart city projects, the generated public value is more 'fuzzy' or still needs to be proven. Although the promises and project goals contain the creation of public value for all of the selected cases, it is unclear whether the creation of public value was actually achieved or not. In order to legitimize smart city projects, it is important for these projects to validate the creation of public value by measuring its impact.

6.9 Potential for Civic Engagement

Civic or community engagement is typically defined along a continuum of participation but it goes further than participation and involvement. It also involves capturing people's attention and focusing their efforts [58, p. 5]. Thus, one can distinguish many forms of community engagement, with varying levels of communication, such as providing knowledge to the public, consulting the public, involving the community, collaborating with the community or empowering the community to make decisions and to implement and manage change [59, p. 8]. Scearce [60] distinguishes five dimensions or processes in engagement:

- Listening to and consulting the crowds: e.g. online conversations and openly asking for advice
- Designing for serendipity: Creating collaborative environments, in person and online
- Bridging differences: Connecting people with different perspectives
- Catalyzing mutual support: Helping people help each other
- Providing handrails for collective action: Giving enough direction for individuals to take effective and coordinated action.

Based on these dimensions, she formulates best-practices related to the social potential of ICT on (a) a 'macro-level', creating a 'public sphere' that enables people in a society to communicate with each other about their positions as citizens and that helps them to act as a political entity; (b) an 'intermediate-level', creating more or less institutionalized and sustainable, but not necessarily formalized, interaction networks of individuals having the same or a similar social position, interests or desires; and (c) a 'micro-level' where ICT has become an important source for the development and acquisition of social capital [61–63].

Because smart cities aim to stimulate 'smart citizenship', they often focus on the empowerment of citizens and improving civic participation, interaction and engagement. All six smart city projects have a high potential for civic engagement. This proves that the above dimensions play a central role in smart cities and that these projects are fully incorporating the stimulation of civic engagement.

6.10 Knowledge Valorization

This dimension assesses the overall valorization of knowledge and surrounding affordances. Although the goals of smart city projects are often very promising, for most of the cases, evidence of solid, sustainable and meaningful valorization of knowledge and enabling processes within the smarty city ecosystem, is sparse. For the selected smart city projects, valorization is mostly of an academic nature (publishing) or serves the purpose of branding a city as an innovative city. Nevertheless, there certainly is ambition to overcome this problem and to stimulate an increased valorization of smart city projects. Through the European projects for example, local developments will be able to be applied in other European cities as well and the Gent Living Lab project aims at bringing together different smart city initiatives in order to optimally make use of the generated knowledge from different projects.

6.11 Sustainability

Sustainability is the main bottleneck of all selected smart city initiatives, with the exception of Future Legends. Smart city projects are often instigated and fuelled by (European) project funding. Once these projects finish, the generated technology, service and/or knowledge disappears. A second threat for the sustainability of smart city projects is technological-determinism. When technology has a central position in the project, the social dimension and the supporting context surrounding the technology are of often neglected. Therefore, most smart city projects have a hard time crossing the chasm from demonstrator towards an autonomous, sustainable product or service which can service without funding.

6.12 Potential for Economic Growth

In the analysis, a distinction is made between the actual generated economic value and the potential for economic growth. This assessment is hypothetical and analyses the potential value of the generated knowledge within the project over a longer period of time, if challenges such as sustainability would be overcome. This allows a comparison between the potential of the project and the actual valorization. In the selected smart city projects two distinct project goals can be distinguished: (a) projects aimed at the creation of public and economic value (Citadel, GLL, Apps For Ghent) and (b) projects aimed exclusively at the creation of public value (Zwerm, MDIVG, Future Legends). Notably, none of the cases merely has economic objectives. The potential for economic growth can be found in the use of Open Data for the development of innovative services (Apps For Ghent, Citadel) or the collaboration between different stakeholders to co-develop innovative services (GLL). The cases that do have potential for economic growth, however, still have to find a way to realize that potential.

6.13 Importance of Funding

All of the cases with the exception of Apps For Ghent, relied on funding for the kick-start of the project. For the European projects (Citadel and Zwerm) this dependency remains very strong even after the project launched. Without funding these projects (would) seize to exist. The local projects on the other hand rely less on European funding, but the downside of this is that this makes it hard for them to realize their full potential. These projects are governed by the city government, but the officials that are working on these projects have only little or no resources (especially time) to do so. In the case of Apps For Ghent, and especially GLL, promises and opportunities are very high but both projects lack the resources to harness these opportunities to their full potential. The Future Legends project is somehow exceptional in the sense that this project is fully supported by the community and no longer needs external support.

7 Conclusion and Discussion

The concept of a ‘smart city’ is a container of promises. It holds the belief that cities can and should act as smart collaborative ecosystems, enabled by state-of-the art technology. It envisions cities as laboratories and drivers for social change. In reality, however, a lot of the promises and the potential of a smart city still remain to be proven on multiple levels. In this chapter, a conceptual framework is proposed which enables the analysis of the architecture, collaboration and different dimensions of smart city projects. When this framework is applied to a set of smart city projects in one local ecosystem, different lessons concerning the current state of smart cities can be learned. By making an overarching analysis of six smart city projects in the city of Ghent, the analysis affords an assessment of the overall ‘smartness’ of a city.

Although smart cities claim to go beyond technology and to have a citizen centric nature, reality shows that a lot of smart city projects still have a rather techno-centric nature (e.g. placing sensors). While collaboration is central in smart cities, not all projects involve all the actors, policy, research, citizens and private partners, in the city. Especially the lack of involvement of private partners and possible business models forecloses the long-term sustainability and economic value creation of smart city projects. Smart cities do have the potential to enable multi-stakeholder collaborative value creation, but therefore they need central governance which stimulates this collaboration, serves as a container for the reuse of knowledge, potentially through Open Data and thus enhancing the sustainability of the generated knowledge. In this context, [64] put forward the concept of *knowledge retention* as an important process in the context of open innovation, indicating the storage and maintenance of knowledge over time. For the city of Ghent, the goals of GLL are most in line with this governance role. This platform, governed by the city government should act as a central actor in the smart city, allowing for an optimal valorization of public and economic value. But for this, sufficient resources are needed, which is the main bottleneck of current smart city initiatives. Most smart city projects rely heavily on public funding, but this funding only has a temporary nature and therefore forecloses long-term planning and strategies, beyond the projects themselves. So far, smart cities have not (or only little) been able to produce long-term creation of value. Most projects are showcases that prove what might be possible, without actual implementation or long-term integration in the everyday life of the city.

In order to move beyond promises and demonstrators, it is important to keep measuring the actual impact of smart city projects. The proposed framework in this chapter highlights and analyses some smart city dimensions, but actual impact measuring remains challenging. Nevertheless, lots of public funding is being consumed by smart city projects, so continuous monitoring and critical analysis is needed in order to force smart cities to prove their added value.

References

1. Eurostat. (2012). Eurostat regional yearbook 2012, Luxembourg.
2. Bloomberg, M. R. (2012). While nations talk, cities act. The @C40Cities mayors summit will advance urban solutions to combat climate change: http://bit.ly/ZGdC3x. March 8, 2012, 7:40 PM. Tweet (MikeBloomberg).
3. Caragliu, A., Del Bo, C., & Nijkamp, P. (2009) Smart cities in Europe. Paper presented at the 3rd Central European Conference on Regional Science (CERS), Košice, Slovak Republic.
4. Dolente, C., Galea, J., & Leporelli, C. (2010). Next generation access and digital divide: opposite sides of the same coin? Paper presented at the European Regional ITS Conference, Copenhagen, Denmark.
5. Verdegem, P. (2009). De digitale kloof en/in e-government: Uitdagingen voor de overheid in de informatiemaatschappij.
6. O'Reilly, T. (2007). What is Web 2.0: Design patterns and business models for the next generation of software. *Commun. Strateg., 65*, 17–37.
7. Van Audenhove, L., Lievens, B., & Cammaerts, B. (2005). E-democratie voor Vlaanderen: Stand van zaken., Brussels, Belgium.
8. Poiesz, T. B. C., & Van Raaij, W. F. (2002). *Synergetische Marketing. Een visie op oorzaken en gevolg van veranderend consumentengedrag*. Amsterdam: Prentice Hall.
9. Frissen, V., & van Lieshout, M. (2006). ICT and everyday life: The role of the user. In P. Verbeek & A. Slob (Eds.), *Technology, behavior and the environment, a multidisciplinary approach*. Dordrecht: Kluwer.
10. ITU. (2005). The internet of things: Executive summary, Geneva, Switzerland.
11. Kwon, O., & Kim, J. (2007). A methodology of identifying ubiquitous smart services for U-city development. In J. Indulska, L. T. Yang, J. Cao, J. Ma, E. Loukis, Y. Charalabidis, & J. Scholl (Eds.), *Ubiquitous intelligence and computing*. Berlin: Springer.

12. Shin, D. (2009). Ubiquitous city: Urban technologies, urban infrastructure and urban informatics. *Journal of Information Science, 35*, 515–526.
13. Choi, J. (2010). The city is connections: Seoul as an urban network. *Multimedia Systems, 16*, 75–84.
14. Loukis, E., Charalabidis, Y., & Scholl, J. (2011). Editorial of the special issue on digital cities. *Telematics and Informatics, 28*, 144–147.
15. Ergazakis, E., & Ergazakis, K. (2011). Digital cities: Towards an integrated decision support methodology. *Telematics and Informatics, 28*, 148–162.
16. Middleton, C., & Bryne, A. (2011). An exploration of user-generated wireless broadband infrastructures in digital cities. *Telematics and Informatics, 28*, 163–175.
17. Komninos, N. (2008). *Intelligent cities and globalisation of innovation networks*. London: Routledge.
18. Ojala, T., Valkama, V., Kukka, H., Heikkinen, T., Lindén, T., Jurmu, M., et al. (2010). UBI-hotspots: Sustainable ecosystem infrastructure for real world urban computing research and business. Presented at the 2nd International Conference on Management of Emergent Digital EcoSystems (MEDES 2010), Bangkok, Thailand.
19. Arnkil, R., Järvensivu, A., Koski, P., & Piirainen, T. (2010). Exploring the quadruple helix. Report of Quadruple Helix Research for the CLIQ Project., Tampere.
20. Etzkowitz, H. (2008). The triple helix: University–industry–government. Implications for Policy and Evaluation.
21. Schuurman, D., De Moor, K., De Marez, L., & Evens, T. (2011). A living lab research approach for mobile TV. *Telematics and Informatics, 28*, 271–282.
22. Westerlund, M., & Leminen, S. (2011). Managing the challenges of becoming an open innovation company: Experiences from Living Labs. *Technology Innovation Management Review, 15*, 223–231.
23. Feurstein, K., Hesmer, A., Hribernik, K., Thoben, T., & Schumacher, J. (2008). Living labs: A new development strategy. In J. Schumacher & V. P. Niitamo (Eds.), *European living labs—A new approach for human centric regional innovation*. Berlin: Wissenschaftlicher.
24. Buitendag, A. A. K., van der Walt, J. S., Malebane, T., & de Jager, L. (2012). Addressing knowledge support services as part of a living lab environment. *Issues in Informing Science and Information Technology, 9*, 221–241.
25. Cosgrave, E., Arbuthnot, K., & Tryfonas, T. (2013). Living labs, innovation districts and information marketplaces: A systems approach for smart cities. *Procedia Computer Science, 16*, 668–677.
26. Paskaleva, K. (2011). The smart city: A nexus for open innovation? *Intell. Buildings International, 3*, 153–171.
27. European Commission. (2011). Communication from the commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A Digital Agenda for Europe., Brussels, Belgium.
28. OECD. (2007). Guidelines for access to research data from public funding, Paris, France.
29. Swan, A. (2012). Policy guidelines for the development and promotion of open access. Paris: United Nations Educational, Scientific and Cultural Organization.
30. Von Hippel, E. (2005). *Democratizing innovation*. Cambridge: MIT Press.
31. Bogers, M., Afuah, A., & Bastian, B. (2010). Users as innovators: A review, critique, and future research directions. *Journal of Management, 36*, 857–875.
32. Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York: New York University Press.
33. Cosgrave, E., & Tryfonas, T. (2012). Exploring the relationship between smart city policy and implementation. In The First International Conference on Smart Systems, Devices and Technologies, pp. 79–82.
34. Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Boston: Harvard Business School Press.
35. Hargadon, A. B., & Bechky, B. A. (2006). When collections of creatives become creative collectives: A field study of problem solving at work. *Organization Science, 17*, 484–500.

36. Fransman, M. (2010). *The new ICT ecosystem: Implications for policy and regulation*. Cambridge: Cambridge University Press.
37. Stocker, A., Dösinger, G., Saed, A., & Wagner, C. (2007). The three pillars of "corporate web 2.0": A model for definition. In Proceedings of the I-MEDIA '07 and I-SEMANTICS '07 Conference, Graz, Austria.
38. Leminen, S., Westerlund, M., & Nyström, A. (2012). Living labs as open-innovation networks. TIM Rev. September, 6–11 2012.
39. Perkmann, M., & Walsh, K. (2007). University–industry relationships and open innovation: Towards a research agenda. *International Journal of Management Reviews, 9*, 259–280.
40. Brem, A., & Viardot, E. (2013). *Evolution of innovation management: Trends in an international context*. Hampshire: Palgrave Macmillan.
41. Bogers, M., & West, J. (2012). Managing distributed innovation: Strategic utilization of open and user innovation. *Creativity and Innovation Management, 21*, 61–75.
42. Hargadon, A. (1998). Firms as knowledge brokers: Lessons in pursuing continuous innovation. *California Management Review, 40*, 209–227.
43. Norman, D. (2002). *The design of everyday things*. New York: Basic Books.
44. Ballon, P. (2007). Business modelling revisited: The configuration of control and value. *Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media, 9*, 6–19.
45. Yin, R. (1984). *Case study research*. Beverly Hills: Sage Publications.
46. Eisenhardt, K. (1989). Building theories from case study research. *Academy of Management Review, 14*, 532–550.
47. Stad Gent. (2007). Strategische nota van het meerjarenplan 2007–2012. Available at http://www.gent.be
48. Schenk, E., & Guittard, C. (2009). Towards a characterization of crowdsourcing practices. Available online at: http://halshs.archives-ouvertes.fr/halshs-00439256/fr/
49. Basili, V. R., Briand, L. C., & Melo, W. L. (1996). How reuse influences productivity in object-oriented systems. *Communications of the ACM, 39*, 104–116.
50. Snijders, C., Matzat, U., & Reips, U. D. (2012). "Big data": Big gaps of knowledge in the field of internet science. *International Journal of Internet Science, 7*, 1–5.
51. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, Retrived May 21, 2014 from http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation
52. Anderson, J. Q., & Rainie, L. (2012). Big data: Experts say new forms of information analysis will help people be more nimble and adaptive, but worry over humans' capacity to understand and use these new tools well, Washington, DC.
53. Mantelero, A. (2012). Masters of big data: Concentration of power over digital information. Available SSRN http://ssrn.com/abstract=2048236
54. Davies, T. (2010). Open data, democracy and public sector reform. A look at open government data use from http://www.data.gov.uk. Retrived May 21, 2014 from http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf
55. Lippert. (2010). Public sector information reuse in Denmark: European public sector information platform, Topic Report No. 20. European Public Sector Information (PSI) Platform, funded by the European Commission.
56. Coenen, T., Mechant, P., Laureyssens, T., Claeys, L., & Criel, J. (2013). ZWERM: Stimulating urban neighborhood self-organization through gamification. Presented at the International Conference Using ICT, Social Media and Mobile Technologies to Foster Self-Organisation in Urban and Neighbourhood Governance, Delft, Netherlands.
57. All, A., Coorevits, L., & Schuurman, D. (2013). Bottom-up radio: Creating a new media format using living lab research. Paper presented at Breaking the Media Value Chain: VII International Conference on Communication and Reality, Barcelona, Spain.
58. Aslin, H. J. & Brown, V. A. (2004). Towards whole of community engagement: A practical toolkit. Canberra: Murray-Darling Basin Commission.

59. Thompson, L., Stenekes, N., Kruger, H. & Carr, A. (2009). Engaging in biosecurity: Literature review of community engagement approaches. Canberra: Bureau of Rural Sciences.
60. Scearce, D. (2011). Connected Citizens: The Power, Peril and Potential of Networks. Knight Foundation and Monitor Institute, Retrived May 21, 2014 from https://knight.app.box.com/shared/ng70lqn9hb
61. Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century*. New York: Farrar Straus and Giroux.
62. Stern, M., & Dillman, D. (2006). Community participation, social ties, and use of the internet. *City Community, 5*, 409–424.
63. Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Díaz, I. and Miyata, K. (2003). The Social Affordances of the Internet for Networked Individualism. *Journal of Computer-Mediated Communication, 8*, 0. doi:10.1111/j.1083-6101.2003.tb00216.x
64. Lichtenthaler, U., & Lichtenthaler, E. (2009). A capability-based framework for open innovation: Complementing absorptive capacity. *Journal of Management Studies, 48*, 1315–1338.
65. Mechant, P., Stevens, I., Evens, T., & Verdegem, P. (2012). E–deliberation 2.0 for smart cities: A critical assessment of two "idea generation" cases. *International Journal of Electronic Governance, 5*, 82–98.
66. Schuurman, D., Baccarne, B., Mechant, P., & De Marez, L. (2012). Smart ideas for smart cities: Investigating crowdsourcing for generating and selecting ideas for ICT innovation in a city context. *Journal of Theoretical and Applied Electronic Commerce Research, 7*, 11–12.

Environmental Sustainable Fleet Planning in B2C e-Commerce Urban Distribution Networks

Francesco Carrabs, Raffaele Cerulli and Anna Sciomachen

Abstract Sustainable distribution is one of the topics concerning the smart city concept. In this chapter we face the problem of delivering a given amount of goods in urban areas arising from e-channel department stores, with the aim of minimizing the overall distribution costs; costs take into account traveling components, loading and other operative aspects, and environmental issues. More precisely, in the present business to consumer distribution problem, we have to determine the fleet of not homogeneous vehicles (trucks, wagons, vans and picks-up) to be used for satisfying the demands of clients coming from e-channels, and their related itineraries, given the traveling limits imposed by the urban government; in particular, we have to respect the maximum route length constraints and use the appropriate vehicles for each kind of street. We propose a mathematical programming model to solve this computationally difficult problem, which is strategic for being able to implement sustainable distribution plans in a smart city context. Preliminary results of test bed cases related to different sized urban distribution networks are reported and analyzed.

Keywords City logistics • Sustainable distribution • e-Channel • Network models • Vehicle routing problem

F. Carrabs · R. Cerulli
Department of Mathematics, University of Salerno, Via Giovanni Paolo II, nr. 132, 84084 Fisciano, SA, Italy
e-mail: fcarrabs@unisa.it

R. Cerulli
e-mail: raffaele@unisa.it

A. Sciomachen (✉)
Department of Economics and Business Studies, University of Genoa, Via Vivaldi 5, 16125 Genoa, Italy
e-mail: sciomach@economia.unige.it

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS, 183
DOI: 10.1007/978-3-319-06160-3_9, © Springer International Publishing Switzerland 2014

1 Introduction

Nowadays, both large and small cities are proposing a new model, called “the smart city”, which represents high technological, sustainable, comfortable and secure living environment. Following this idea, a number of models have been developed and deployed with the help of technological advances in computer and communication, such as Information and Communication Technology (ICT) and Intelligent Transport Systems (ITS), which constitute precisely the basis of the smart city model [2, 12].

Sustainable distribution is one of the topics concerning the smart city concept. Recently, increasing attention has been particularly devoted to sustainable development of urban areas as well as mobility of goods for ensuring the wellbeing of community. The aim of a sustainable urban distribution network is to analyze how society intends to provide the means to properly meet economic, environmental and social needs efficiently and equitably, while minimizing negative impacts and their associated costs, including environmental issues, such as congestion, noise and air pollution. In this sense, the idea of city logistics has been proposed to establish efficient and environmentally friendly urban logistics systems [5, 11].

A difficulty in modeling city logistics comes from the complex interactions between private and public stakeholders involved in urban freight transport: shippers, freight carriers, administrators and residents (consumers). In fact, city logistics requires advanced optimization and simulation modeling approaches and tools to assist in the design, implementation and evaluation of schemes that satisfy the needs of all the above stakeholders, who hold different concerns and objectives. While the recent growth of research into urban distribution and city logistics is encouraging (see e.g. [6, 13]), only few works have been concerned with examining the likely impact of policy measures on distribution operations. A review of emerging techniques for enhancing the practical application of city logistics models is presented in [7, 12]; focuses on the evaluation of urban tours traveled by different types of commercial vehicles and their related costs. In Anderson et al. [1] a project is presented having the aim of investigating the ways in which alternative policy measures, such as weight and access time restrictions, can result in changes in the vehicle activities involved in urban distribution operations. New challenges have been observed for distribution systems designed within smart city frameworks. In particular, models of vehicle routing problems (VRP) are considered basic tools for implementing sustainable good distribution channels in urban areas. In this direction, a number of chapters on VRP have been published by operations researchers and practitioners (see, e.g. [9, 10]) with the aim of providing advances for the development of ITS within smart city models. In the present chapter we consider a particular case of VRP, originating from the need of delivering goods in an urban context arising from e-channel department stores. More precisely, in this urban business to consumer (B2C) distribution problem we have to determine the fleet of not homogeneous vehicles (trucks, wagons, vans and pick-up) to be used for the delivery of a given amount of goods in urban areas. Note that the management of the fleet and the global routing of vehicles in the urban network are key elements for sustainable goods distribution plans. Our problem is strongly connected to the design of city logistics systems for medium–large cities, where it provides the means to efficiently keep large trucks out of the city center, with small and environment-friendly vehicles providing the last leg of distribution activities [5]. Following this direction, in this chapter, each vehicle involved in the distribution process is characterized by two parameters: (1) the size (which allows it to cross only some types of roads) and (2) the maximum load capacity. Starting from a depot (to be determined) each vehicle must pass through the streets of the city (compatible with its size) to deliver the required goods along that road and go back to the chosen depot. The considered cost components, to be minimized, take into account traveling, loading and environmental issues.

We present a mathematical formulation of this novel urban B2C distribution problem for solving it. The referring urban B2C network problem (UB2CNP) is presented in more details in the next section. Section 3 reports the proposed network model and the related Mixed Integer Programming (MIP) formulation. Finally, some preliminary results and outlines for future works are given.

2 Problem Definition

The proposed urban logistic network problem (UB2CNP) can be seen as an extension of the classical vehicle routing problem VRP, encountered very frequently in making decisions about the distribution of goods and services. Given a number of customers with known demands and a fleet of not identical vehicles with known capacities, the problem consists in finding a set of routes originating and terminating at a central depot and serving all the customers exactly once. The routes cannot violate the capacity constraints on the vehicles. Differently from the classical VRP formulation, in addition, we must meet the size constraints on the streets, which specify which kind of vehicle can cross the street. All problem parameters, such as customer demand and typologies of streets, are assumed to be known with certainty. The standard objective of the UB2CNP problem consists of minimizing the total travel cost.

The UB2CNP is a basic distribution-management problem that can be used to model many real world problems. Some of the most useful applications of the UB2CNP include bank deliveries, postal deliveries, industrial refuse collection, national franchise restaurant services, school bus routing, security patrol services, and vendor deliveries for just-in-time manufacturing.

Here, the UB2CNP applies to deliver groceries ordered from e-channel department stores to customers who reside at their homes. The management of the department stores has hence to collect the orders and group them according to the allowable vehicles. Further, customers are identified according to their address with reference to the corresponding kind of street, for being able to define the routes necessary to satisfy the overall demand and choose the best vehicle to use for the delivery which minimizes costs and the environmental impact. The problem, as particular case of the classical VRP problem, is NP-complete [8] that is computational difficult to be solved, and instances involving more than 100 customers are very hard to solve optimally. For this reason it makes sense to focus on the development of efficient mixed integer programming formulation models, possibly accomplished by the creation of heuristics approach to solve the problem. For recent surveys on the state of the art in VRP research we recommend the survey by Cordeau et al. [4] that describes both exact and heuristic methods, and the survey by Bräysy and Gendreau [3] that focuses on metaheuristics.

3 The Urban Logistic Problem

Formally the UB2CNP is defined as follow. Let $G = (V \cup \{0\}, E, L)$ be a connected digraph where V is the set of locations, 0 is a special vertex representing the depot, L is the set of different typologies (label) of streets, A is a set of arcs to which two values are associated: (1) a nonnegative weight t_{ij}, denoting the travel time (or the edge length) and (2) a label indicating the edge (street) type. Let n, m and l be the cardinality of V, E and L, respectively. A service requirement q_i, which can be delivery from the depot, associated with each customer. Vehicles of different type and different capacity must be routed to serve all the customers. A feasible vehicle route $\rho = \{0, v_1, v_2, \ldots, v_{\ell-1}, v_\ell, 0\}$ of length l is an ordered sequence of different customers to be served such that the total capacity of the vehicle is not exceeded and, the streets constraints are satisfied. A feasible solution $S = \{\rho_1, \rho_2, \ldots, \rho_k\}$ of the problem is a collection of feasible routes. We denote by $c(\rho)$ the total length of route ρ and by $c(S) = \sum_{\rho_i \in S} c(\rho_i)$ the total length of the feasible solution S. The UB2CNP problem consists in computing the minimum cardinality set $S = \{\rho_1, \rho_2, \ldots, \rho_k\}$ of feasible routes such that all the customers are served and each customer is visited by a single vehicle. Note that this objective implies the minimization of the number of vehicles used for delivering the required goods, thus in turn reducing both the congestion and the pollution in the city tours as well as the final cost.

3.1 The Urban Logistic Network Model

To model this problem we use an edge labeled graph. The nodes represent intersections and the arcs the streets of the city. The nodes are classified as: depot nodes (where goods are stored), costumer nodes (where goods have to be delivered) and the transshipment nodes. We assign a different label to each type of street according to its width. Moreover, each label will be associated to a particular type of vehicle. Without loss of generality, we assume that the labels are ordered according to the width of the street. For example, if there are three types of roads there are three different labels: A, B and C. The vehicles associated to label A can travel along the streets of type A, B and C, labeled B vehicles can travel along streets B and C, while vehicles with label C are allowed to pass only through streets of type C. Each vehicle is characterized by two parameters: (1) the size (which allows it to cross only some types of roads) and (2) the maximum load capacity. Starting from a depot each

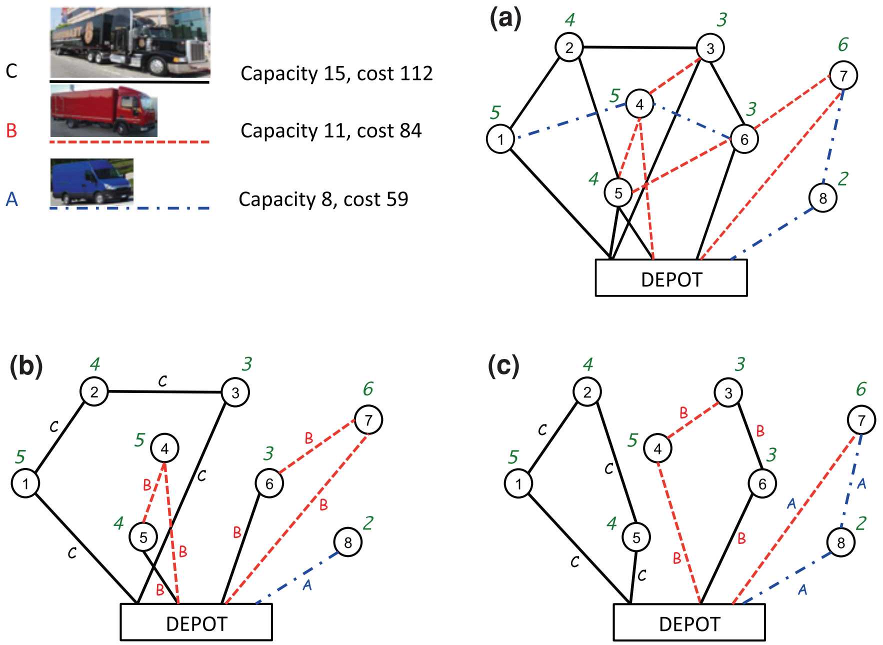


Fig. 1 A simple example of the problem where no cost are associated to the edge of the graph. **a** The labeled graph G. **b** A feasible solution with value 112 + 84 + 84 + 59 = 339 and **c** a better solution with value 112 + 84 + 59 = 255

vehicle must pass through the streets of the city (compatible with its size) to deliver the required goods along that road and go back to the depot. A cost c_k is associated to each type of vehicle. The length of each route cannot exceed a fixed value. A simple example of our referring urban B2C network model is reported in Fig. 1a, which shows a small urban center in which eight customers must be supplied from a single depot by using three type of vehicles. To each vehicle is associated a capacity 8, 11, 15 and a fixed cost 59, 84, 112, respectively. To each edge is associated a label representing the type of vehicle that can cross this edge. In this particular case, the cost of the edges is neglected. The numbers outside the nodes represent the associated goods' demand. In Fig. 1b and c are reported two feasible solutions with cost 339 and 255, respectively. Readers can easily note how the number of the used vehicles impacts on the final cost.

3.2 *Mixed Integer Programming Mathematical Formulation*

In this section we present a integer programming formulation for the UB2CNP. Before presenting the whole model let us summarize the required notations. Consider customers at various locations in the city which must be served by vehicles hosted at a central depot. Denote the central depot by 0 and the locations by $i = 1 \ldots n$. We can represent the input information using a directed network $G = (V \cup \{0\}, E, L)$, where V denotes the set of n vertices, E the set of m arcs (the streets) and L the set of labels associated to the arcs (the streets characteristic).

The following inputs are assumed to be available:

- T = number of vehicle types;
- Q_t = capacity of vehicle type $t(Q_1 < Q_2 < \cdots < Q_T)$;
- f_t = fixed activation cost of vehicle type $t(f_l < f_2 < \cdots < f_T)$;
- d_j = demand of customer j;
- c^t_{ij} cost to pay for each vehicle of type t that crosses the arc (i, j);
- a^t_{ij} that assumes value equal to 1 if the edge (i, j) can be traversed by the vehicles of type t;
- V_d set of demand nodes;
- V_p set of transhipment nodes ($V = V_d \cup V_p \cup \{0\}$ and $V' = V_d \cup V_p$).
- mk = number of vehicles of type k available

In addition, the following decision variables are used:

- binary variable x^k_{ij} that assumes value equal to 1 if a vehicle of type k travels from i to j, and 0 otherwise;
- continuous variable y_{ij} that represents the flow of goods from i to j.

Then, the (MIP) formulation of UB2CNP is the following:

$$\min \sum_{k \in T} f_k \sum_{j \in V'} x^k_{0j} + \sum_{k \in T} \sum_{\substack{i,j \in V \\ i \neq j}} c^k_{ij} x^k_{ij} \tag{1}$$

$$s.t. \sum_{k \in T} \sum_{i \in V} x^k_{ij} = 1 \qquad \forall j \in V_d \tag{2}$$

$$\sum_{i \in V} x^k_{ip} - \sum_{j \in V} x^k_{pj} = 0 \qquad \forall p \in V', \forall k \in T \tag{3}$$

$$x^k_{ij} \leq a^k_{ij} \qquad \forall i, j \in V,\ i \neq j,\ \forall k \in T \tag{4}$$

$$\sum_{j \in V'} x^k_{0j} \leq m_k \qquad \forall k \in T \tag{5}$$

$$\sum_{i \in V} y_{ij} - \sum_{i \in V} y_{ji} = q_j \qquad \forall j \in V_d \tag{6}$$

$$\sum_{i \in V} y_{ij} - \sum_{i \in V} y_{ji} = 0 \qquad \forall j \in V_p \tag{7}$$

$$y_{0j} \leq \sum_{k=1}^{T} (Q_k) x^k_{0j} \qquad j \in V' \tag{8}$$

$$y_{ij} \leq M \sum_{k=1}^{T} x_{ij}^{k} \qquad \forall (i,j) \in E \tag{9}$$

$$x_{ij}^{k} \in \{0,1\} \qquad \forall i,j \in V,\ i \neq j,\ \forall k \in T \tag{10}$$

$$y_{i,j} \geq 0 \qquad \forall (i,j) \in E \tag{11}$$

where M is chosen to be a large number so that (9) becomes redundant if $\sum_{k \in T} \sum_{i \in V} x_{ij}^{k} = 1$. For our problem is easy to see that a correct value for M is $\max_{k \in T}\{Q_k\}$. However, due to the constraints (4) we can associate different value M_{ij} to each arc (i,j) of the graph considering the maximum capacity of the vehicle among those that can traverse the arc (i,j).

In the above formulation, the objective function (1) requires the minimization of the total cost to serve all customers. Note that the cost coefficients depend on the type of the vehicles; in this way we are able to take into a proper account a sort of pollution charge depending on the environmental impact of the vehicle. Moreover we consider a fix cost f_k required to use the vehicle k. Constraints (2) and (3) impose that a customer is visited exactly once and that if a vehicle visits a customer, it must also depart from it. Constraints (4) guarantees that each vehicle can traverse only appropriate streets. The maximum number of vehicles available for each vehicle type is imposed by constraints (5). Constraints (6) and (7) are the commodity flow constraints: they specify that the difference between the quantity of goods a vehicle carries before and after visiting a customer is equal to the demand of that customer (this demand is equal to 0 for the transhipment nodes). The constraints (8) ensure that the vehicle capacity is never exceeded whenever the constraints (9) guarantee that the value y_{ij} can be greater than 0 only if exists at least a vehicle that crosses the arc (i,j). Finally, constraints (10) and (11) are the variables constraints.

4 Computational Tests

The model were coded in C++ and solved by CPLEX 12 on a 2.33 GHz Intel Core2 processor. We carried out the computational tests on a set of scenarios composed by three instances having the same number of vertices, edges and vehicles. In the randomly generated instances, the number of vertices ranges from 10 to 40 and the density ranges from 0.3 to 0.5. We used small instances because, how we will see in the following, the UB2CNP problem appears very hard to solve in particular when the density of the graph increases. Moreover, we generated instances with 2 and 3 different type of vehicles in order to evaluate also the impact of this parameter on the performance of the model; in particular, in our instances we consider two types of urban routes where vans and wagons, and vans, wagons and trucks are allowed, respectively.

Table 1 Test results carried out on the small instances with (a) 2 type of vehicles and (b) 3 type of vehicles

id	n	m	v	MIP		
				Obj	#Routes	Time
(*a*)						
1	10	13	2	483.33	1	0.02
2	10	18	2	417.66	1	0.02
3	10	21	2	1006.33	3.33	0.07
4	20	54	2	1880.66	5.66	1.06
5	20	73	2	1469.33	5	4.96
6	20	94	2	1983	5.33	20.33
7	30	127	2	3297	8.66	7.91
8	30	170	2	*2938.66**	*8.33**	434.35
9	30	216	2	2879.33	6	1593.98
10	40	233	2	*2735.33**	*8**	2480.98
11	40	307	2	*2545.33**	*7.66**	3142.33
12	40	385	2	*2229**	*7.33**	3748.88
(*b*)						
1	10	13	3	544.66	1	0.02
2	10	18	3	507.33	1	0.02
3	10	21	3	1789.33	4.66	0.18
4	20	54	3	4186	7.33	3.94
5	20	73	3	2909.33	5.33	10.74
6	20	94	3	3515.33	5.66	719.27
7	30	127	3	5324	8	324.07
8	30	170	3	*5078.66**	*8**	2627.9
9	30	216	3	*5174**	*7.33**	3763
10	40	233	3	*5122**	*9**	3971.91
11	40	307	3	*3461**	*7.33**	7210.03
12	40	385	3	N.D.	N.D.	N.D.

* is associated to the computational time if the optimal solution is not found within the fixed time limit.

In Table 1a are reported the results of the model with a number of vehicles equal to 2. The first four columns list the id (*id*), the number of vertices (n), the number of edges (m) and the number of different vehicles ($|T|$), respectively. The column MIP is divided in three subcolumns (*Obj*, *Routes* and *Time*) reporting the objective function value, the number of the routes found and the CPU time (in seconds) spent. A threshold of 2 h and of 3 GB of memory were imposed for the solution of each instance. The results reported in each line of the table are the average values computed on the three instances of the same scenario. Finally, if for at least an instance of a scenario the model finds a feasible (but not optimal) solution, within the time limit or the memory limit, the marker "*" is reported on the column *Obj* and *Routes* of that scenario. Moreover, if for at least an instance of a scenario the model does not find a feasible solution, within the thresholds, the term N.D. (Not Determined) is reported for this scenario.

From the results of Table 1a we can see that the model is able to solve all the instances up to 30 vertices except for the scenario n°8. On the scenarios up to 20 vertices, the model is very fast while on the instances with 30 vertices the computational time increases meaningfully. Obviously, as the density of the graph grows the computational time increases. However, it is interesting to notice that, in some cases, instances with more vertices and low density require less computational time than instances with less vertices but higher density (see scenarios 6 and 7). On the greatest instances with 40 vertices, the model never finds the optimal solution but, within the thresholds, a feasible solution is always found.

In Table 1b are shown the results of the model with a number of vehicles equal to 3. Comparing the Time columns of the two tables it is evident that the complexity of the instances meaningfully increases when the number of vehicles grows. Indeed, the model finds the optimal solution on the scenarios up to 7. On the remaining five scenarios, the model finds in four cases a feasible solution while on the scenario n°12 it fails to find a feasible solutions. Also in this table, the scenario n°6 required more computational time than the scenario n°7 and this enforces our conjecture that the performance of the model are more affected by the density than by the number of vertices of the graph.

It should be noted that the value of the solution is closely related to the environmental impact of the solution: the smaller the value of the solution (smaller the cost of the objective function), the lower is the congestion of city streets and, therefore, the lower is the emissions of greenhouse gases and air polluting compounds and noise congestion.

5 Conclusion and Outlines for Future Works

In this chapter we propose a variant of the classical vehicle routing problem (VRP). We called "Urban logistic network problem" (UB2CNP) this new variant. For this new problem we propose an integer mathematical formulation; the problem originated from the need of determining a sustainable fleet of vehicles to be used for delivering goods in a urban B2C distribution problem.

We execute some preliminary tests of our mathematical programming model on random generated graph instances, representing urban transportation networks.

In the future experimentation we will highlight the importance of the type of vehicles and how this type affects the optimal solution of the problem; in particular we deeply analyze the environmental impact in the objective function cost component. Moreover one of the aims we want to achieve is to study the relationship between the reduction of the emissions of greenhouse gases and the increased costs of the distribution service. To do this we will use the methodology of sensitivity analysis. From applicative point of view we strongly believe that the proposed novel variant of the classical VRP goes in the direction of the development of ITS which is one of the necessary tools for efficient smart city models.

References

1. Anderson, S., Allen, J., & Browne, M. (2005). Urban logistics—how can it meet policy makers—sustainability objectives? *Journal of Transport Geography, 13*, 71–81.
2. Anttiroiko, A. V., Valkama, P., & Bailey, S. J. (2013). Smart cities in the new service economy: Building platforms for smart services. *AI & Society,*. doi:10.1007/s00146-013-0464-0.
3. Bräysy, O., & Gendreau, M. (2005). Vehicle routing problem with time windows, part ii: Metaheuristics. *Transportation Science, 39*(1), 119–139.
4. Cordeau, J-F., Desaulniers, G., Desrosiers, J., Solomon, M. M., & Soumis, F. (2002). The vehicle routing problem. In P. Toth & D. Vigo (Eds.), volume 9 of SIAM monographs on discrete mathematics and applications, Chap. 7, 157193. SIAM, Philidelphia.
5. Crainic, T. G., Ricciardi, N., & Storchi, G. (2004). Advanced freight transportation systems for congested urban areas. *Transportation Research Part C, 12*, 119–137.
6. Crainic, T. G., Ricciardi, N., & Storchi, G. (2009). Models for evaluating and planning city logistics systems. *Transportation Science, 43*(4), 432–454.
7. Figliozzi, M. A. (2010). The impacts of congestion of commercial vehicle tour characteristics and costs. *Transportation Research Part E, 46*, 496–506.
8. Lenstra, J. K., & Rinnooy Kan, A. H. G. (1981). Complexity of vehicle routing and scheduling problems. *Networks, 11*, 221–227.
9. Li, J. Q., Borenstein, D., & Mirchandani, P. B. (2007). A decision support system for the single-depot vehicle rescheduling problem. *Computers & Operations Research, 34*, 1008–1032.
10. Qureshi, A. G., Taniguchi, E., & Yamada, T. (2009). An exact solution approach for vehicle routing and scheduling problems with soft time windows. *Transportation Research E, 45*(6), 960–977.
11. Taniguchi, E., Thompson, R. G., Yamada, T., & Van Duin, R. (2001). *City logistics—network modeling and Intelligent transport systems*. Oxford: Elsevier.
12. Taniguchi, E., Thompson, R. G., & Yamada, T. (2012). Emerging techniques for enhancing the practical application of city logistics models. *Procedia—Social and Behavioral Sciences, 39*, 3–18.
13. Teo, J. S. E., Taniguchi, E., & Qureshi, A. G. (2012). Evaluating city logistics measure in e-commerce with multiagent systems. *Procedia—Social and Behavioral Sciences, 39*, 349–359.

Smart Security: Integrated Systems for Security Policies in Urban Environments

Enrico di Bella, Francesca Odone, Matteo Corsi, Alberto Sillitti and Ruth Breu

Abstract Smart Security systems are applications of the Smart City paradigm for local crime prevention. Like most Smart City tools, they consist of informational and technological components that support decision-making processes. A prerequisite for such tools is that they are supposed to be means of ongoing management and policy innovations: we therefore review some of the crucial components of a Smart Security system from the viewpoint of a local government or a local branch of the public administration, in order to analyze the high-level requisites, characteristics and potentials of such a system. The objective is to help Public officials in identifying both what defines a useful technical tool but also what is required on the part of the public administration to actually make it useful. We therefore discuss the following problems. First, we address the issue of indicators, data and the use of statistical analysis to infer the likely determinants of crime and to define risk parameters for urban spaces. In doing that, we suggest innovative tools to introduce spatial information in crime count models. Second,

E. di Bella (✉) · M. Corsi
Department of Economics and Business Studies, University of Genoa, Genoa, Italy
e-mail: enrico.dibella@economia.unige.it

M. Corsi
e-mail: matteo.corsi@edu.unige.it

F. Odone
Department of Informatics Bioengineering Robotics and Systems Engineering, University of Genoa, Genoa, Italy
e-mail: francesca.odone@unige.it

A. Sillitti
Center for Applied Software Engineering, Free University of Bozen, Bolzano, Italy
e-mail: alberto.sillitti@unibz.it

R. Breu
Institut für Informatik, University of Innsbruck, Innsbruck, Austria
e-mail: ruth.breu@uibk.ac.at

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_10,

we discuss sensors and sensor output analysis, trying to define the circumstances that make it useful and the new possibilities offered by current technology. Then we discuss about integration of different information both from a conceptual and a technical point of view, stressing the importance of closing the gap between cold and hot data in order to realize an integrated early warning system. Finally, we discuss the problem of creating a scalable Smart Security system in a local government, indicating a list of significant international experiences.

Keywords Crime mapping • Urban security policies • Security dashboard • Smart security • Intelligent video surveillance

1 Introduction

Smartness for urban environments is supposed to imply a commitment to innovation in technology, management and policy, but the first element of this triad has been researched within the "smart" framework more extensively than the other two [1]. This is the case as well with the specific dimension of urban smartness that is security [2]. Systems for crime visualization, analysis and street surveillance have already been proposed and researched theoretically and applied in practice (e.g., [3, 4]). From an IT standpoint, the gradual innovation regarding these tools has been mostly confined to the integration of different technologies and the development of new technical tools. In a few cases, authentically smart projects have aimed at innovating the management and the policies of urban security "together with" instead of "as a consequence of" the technology of urban security, but they have been few and far between [5–7]. Our intent is, therefore, to illustrate the structure, the logic, the objectives and the requirements of a "Smart Security" system from a management and policy point of view. The issues that we will cover are, of course, just as technical, but each single technical tool or methodology is going to be discussed from a problem-solving point of view, with greater focus on directing public administrations towards promising fields and less on suggesting hardware or software solutions for IT experts.

The foundational assumption of Smart Security is that to improve quality of life, city governance and management should be based on an exhaustive amount of information on a wide range of activities occurring in public spaces [8]. When collected consistently and in the correct format, such information may constitute the input of analytical tools allowing local governments to anticipate and understand economic and social processes and to respond effectively to issues, crises and environmental changes (e.g., [9, 10]). In the specific field of crime prevention, local governments are not always and not only the main actors of public security (depending on national systems) but also decision-makers for a number of social, economic and urban planning policies that can have huge effects on crime. Because of this, a Smart City approach dealing with urban security should be focused on translating theoretical knowledge about crime and deviancy into indicators, early warning systems, models and analytical frameworks. Such a toolbox should then come into play when and where decision making takes place, supporting well informed, precisely targeted and correctly monitored policies.

In recent years, a number of large western cities have started massive investments aimed at innovating in the field of urban security and at building a better informational background to policy decisions about crime prevention, fear of crime and support to the more vulnerable components of the community (e.g., [7, 11]). There are, however, significant challenges to those efforts:

1. criminology offers a wide range of indicators concerning urban security, but most of them are disputed; different criminological theories suggest different ways of measuring crime, of measuring its determinants and defining the correct scale at which determinants should be identified;
2. behaviors and situations may be more accurate at defining crime then any indicator, but sensors meant to capture behaviors and situations either deliver information *post facto* or they are affected by a severe trade-off between accuracy and earliness;
3. indicators and sensors could theoretically work complementarily, both with the idea of extending the ability of a system to identify different and evolving threats and that of allowing triangulation [12]; however, integration of data sources of such different kinds is far from trivial and requires a consistent amount of planning and the cooperation of experts coming from different disciplines: criminologists, economists, statisticians, urban planners, video image analysts, and computer scientists;
4. even when information is available and reasonably accurate and timely, preventive action requires a lot on the part of the public administration; part of that is about technological innovation but a significant part is about management and policy innovation [1].

Smart Security should approach these challenges in two different ways: on one hand, it has to assume as relevant to its domain every technical solution that provides useful information and support to action for the Public Administration; on the other hand, it has to provide a constant evaluation of the consistency of each innovative tool with preexisting and preordained high level goals of innovative security policy and management. Conceptually, a Smart Security system consists of three logical units: the first one is the module for the analysis of “Place and Population”, where crime is analyzed in conjunction with its macro determinants; the second one is the “Individuals and Behavior” module where it is analyzed at micro level and where the actions and movements of the individuals are relevant; the last module is the “Integrated System” software infrastructure which coordinates all the information flows inside the system and includes the user’s frontend where most of the informative elements for the policy actions are shown. Compared to technologically-driven smart programs, Smart Security adds a virtual fourth element in the intense feedback between technological innovation and management and policy innovation. The new frontier in the field of Smart Security Systems consists of the integration of these partial elements in a single framework as the one described in Fig. 1.

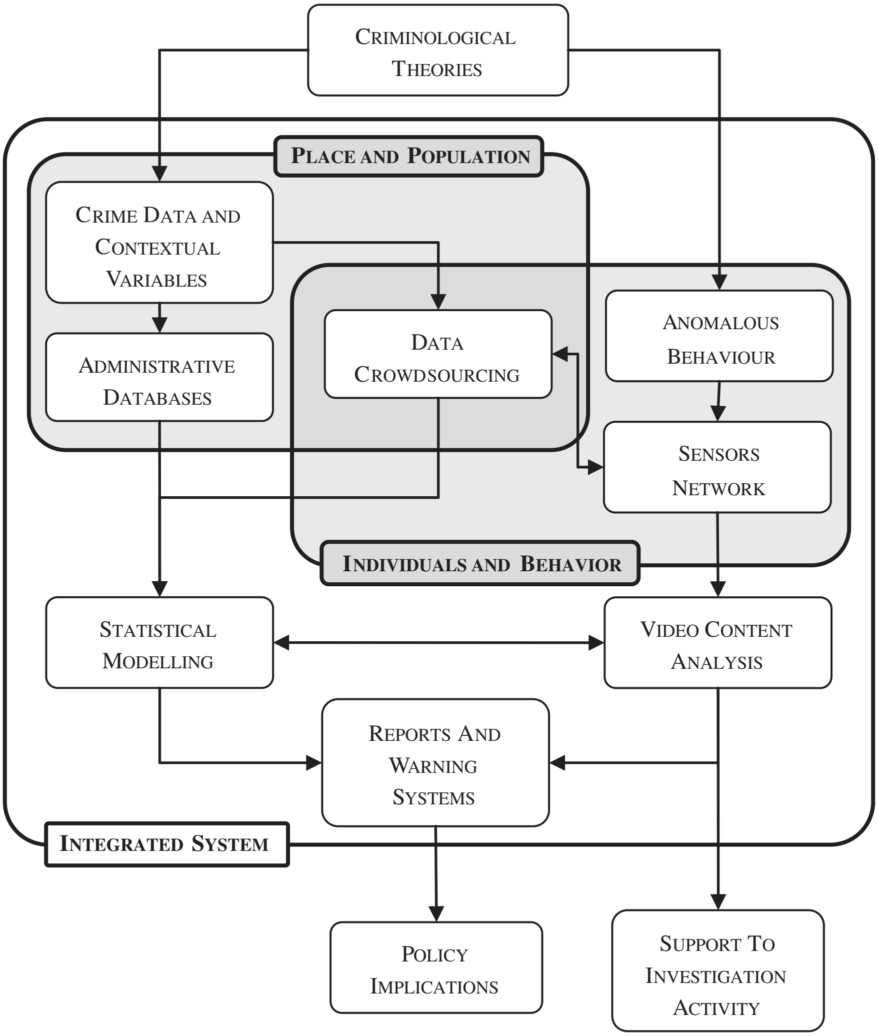


Fig. 1 Logical structure of a Smart Security Integrated System

2 Measuring Crime and Its Determinants in Urban Environments

Information concerning crime that is relevant to Smart Security includes measures of crime and measures of risk or mitigating factors. Such information may not be sufficient to create Smart Security systems, but it is all but necessary.

Like all measures, those concerning crime and its determinants are spatially and temporally located: they matter precisely because they provide intelligence about specific times and places. Since crime is not a constant over time and it is not distributed uniformly in space, it is common practice to draw crime trends and crime maps [13]; these are two relatively trivial building blocks of any informative system (including Smart Security systems) designed to support decision making on urban security and both have a history that's at least a century old. However, as obvious as crime trends and maps are these days, they imply a concept that should be key to any innovative Smart Security system. The concept is that temporal and spatial clusters of crime are the "footprints" of local risk factors and local mitigating factors. It goes without saying that local determinants may change not only in size/intensity, but also in quality. So a Smart system is increasingly informative the more it is capable of mapping crime and its determinants at high resolution.

Measuring crimes is a less trivial activity than one might think: a crime is a legal (abstract) entity consisting of complex behaviors and multiple acts which are hardly numerable in most cases; a simple count of crimes requires therefore, a first level abstraction/elaboration that consists in identifying a reasonable proxy indicator for crimes (like calls for service, police incident reports, victimization self-reports, complaints, sentences, etc.). The raw number of crimes is rarely of use in support of management and policy decisions, as it is inadequate for cross-sectional and inter-temporal comparisons [14]; other indicators have been used in criminology and for official data and statistics, usually as an elaboration of a raw count of crimes, like population-based rates, risk-based rates, densities and location quotients. However, decision makers and public officials should be advised that different indicators actually indicate different things, that is, each proxy and each elaboration of the simple count of crimes carries with itself more or less sophisticated assumptions and meaning differences [15–23]. As for the indicators of risk factors and mitigating factors, a long and intricate debate has been discussing the determinants of crime since the early years of the discipline of criminology. The Department of Sociology of the University of Chicago is the source of the Social Disorganization Theory [24]. By studying the vast growth of the city of Chicago between 1860 and 1910, they noticed that urban areas were more crime-prone than rural ones. Moreover, they identified a connection between crime and several urban issues like poverty, racial heterogeneity, and residential mobility, all leading to the weakening of social control and the disintegration of formal social organizations [13].

The interest in geographic criminology began during the 19th century in France and in Belgium after the publication of the first geographical map of crime. In 1829, Michel Andrè Guerry and Adriano Balbi [25] published a map representing the distribution of crime over the Departments of France between 1825 and 1827. This preliminary study was followed by that of the Belgian statistician and astronomer Quetelet in 1842 and by a number of studies in the Netherlands, England and Wales and in Italy [13].

Between the 60s and the 70s several authors (e.g., [26–28]) developed analytical frameworks of crime and insecurity in the urban environments focused on spatial and functional features of the built environment. Their work, which is globally labeled as the Ecological Theory of Crime, is the combination of very different approaches (Crime Prevention Through Environmental Design—CPTED, defensible space, eyes on the street, etc.).

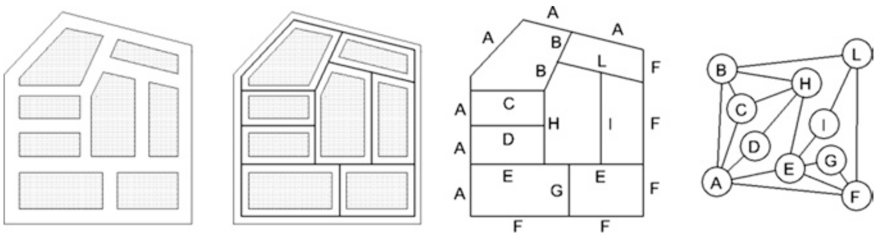


Fig. 2 Dual representation of the urban map. From the *left* to the *right* a simple urban space made of buildings (*shaded shapes*) and roads; the median lines network of the streets among the buildings; the roads network (letters stand for the street names); the graph corresponding to the original urban layout based on the street crossings

These ideas paved the way, during the 80s, to the development of situational crime prevention [29–32]. According to the followers of the situational crime prevention, to reduce the number of crimes, it is necessary to reduce the opportunities of committing a crime because "opportunity makes the thief" [33]. These ideas led crime analysts to increase the attention for urban design details (such as street furniture, street lighting, pedestrian pathways, housing design, visibility from the street and of the street) and to a deep study of the spatial configuration of the streets conducted through the Space Syntax Analysis (SSA) [34]. SSA was initially conceived as a theory to analyze small environments and their configurational features. This discipline studies the configurational properties of urban space [35] through quantitative measures. Thus, it allows the identification of patterns and structures which influence the development of activities in space, in particular movement and land use [36]. Figure 2 exemplifies for a simplified urban structure how it is possible to convert an urban layout (first figure on the left) into a graph (last figure on the right), a mathematical object whose characteristics can be measured in many different ways (e.g., [37]). Since movement and land use are thought to be linked to crime, SSA was used in the development of the CPTED proposed by Jeffery [28]. Thanks to the increasing number of measures used in the Space Syntax Analysis, it soon became possible to compute the relative degree of accessibility, connection, and integration of each street in its urban network, and to index numerically a large number of properties of the urban environment [36]. Among the others, [38] analyzed the street structure and its dependence with crime volumes: they found that streets with many twist and turns have higher crime rates.

During the last thirty years, a new theory on the spread of crime through urban spaces emerged. According to the Routine Activity Theory (e.g., [39]), the number of crimes increases if the number of opportunities for criminals rise and if society lacks an adequate surveillance against crime. Indeed, crimes are often committed in places where victims and offenders hold their routine activities, for example work, leisure, or social interaction, and where they satisfy their basic needs [40]. This theory focuses on space because it is considered an explicit determinant of human actions, including committing offences. Some empirical studies are in favor of this theory [39]. Used Routine Activity Theory to explain the increase in the number of crimes in American cities. For instance, they pointed out that, with more women working, a larger number of houses were empty during daytime and this fact led to the rise in the number of robberies increasing the vulnerability of suburbs [41]. Found out that, in Cleveland, streets with schools and bars are highly crime dense, while [42] identified the places near commercial stores as particularly risky. In this context, some studies on the relationship between crime and transports have been developed by [43]: they conclude that the structure of the public transport system can influence the number of crimes committed: higher numbers of crimes are recorded near stations and bus stops.

In recent years, a new interest for a combined study of socio-demographic and spatial factors in the analysis of crime has emerged. In fact, although crime mapping is certainly the most immediate way to obtain quick information on the criminal incidence in an area, it is interesting to study the relationship between urban crimes and the economic, socio-demographic and spatial features of the study region. Indeed, the study of crime in the context in which it happens could bring to the identification of both global and local risk factors, helping local governments in drawing up policies for Urban Security [44]. Provides empirical evidence for skepticism on the idea of "territoriality" and "defensible space" put forward by Newman [27]: he suggests that, other things being equal, property crimes tend to cluster in those globally or locally segregated areas. In detail, particularly risky areas can be found in cul-de-sac footpaths and rear dead end alleys, but also in those segregated short cul-de-sac carriageways which Newman considered to be the key places where local surveillance should be increased and casual intrusion by non-residents excluded. Hillier [45], discussing the work by Chih-Feng Shu [44], concludes that in Space Syntax Crime Analysis, spatial factors are relevant and that they operate both at a global and local level. More recently, [46] discuss the relationship between crime and urban planning presenting also the results of an empirical research conducted in the city of Vilnius: the aim of this study is to identify, with the use of ASA, the most vulnerable open public spaces of the city.

3 The Role of Statistical Analysis in Integrated Systems for Smart Security

Information of the kind presented in the previous paragraph becomes relevant to Smart Security systems when it allows local governments and the public administration to monitor the development of the situation, to infer plausible causal relationships between some theoretical determinant of crime and a certain measure of crime and when it allows either to identify promising actions that can be taken or situations that cannot be explained under the available information and require additional investigation. From a statistical standpoint, it means being able to produce basic descriptive statistics of crime and being able to produce statistical models. The basic statistics of crime are little more than the conceptualization of crime trends and crime maps: means, rates, standard deviations, spatial and temporal clusters, etc. These are the most commonly and widely used tools for the statistical analysis of crime and they let public administrator monitor the evolution of crime over time or compare crime rates in different areas, they do not include any interpretation of the counted events, neither suggest possible policies or actions that may be appropriate or useful. The availability of microdata that contain georeferenced information on relevant risk factors at street level of detail, allows a second, more effective level of analysis. The database of reported crimes, made of records containing a full set of the available information relating each crime (e.g., date and time of the event; gender, age, and nationality of the victim; place of the event; etc.), can be combined with all the other information that Municipalities possess for their administrative purposes.[1] Therefore, criminal events recorded by law enforcement agencies and risk factors suggested by criminological theory can be analyzed conjointly (e.g., [47]). Statistical models identify which contextual variables actually work as risk factors or mitigating factors and can be considered as explanatory variables with crime being the dependant variable. The interpretation of the model starts with the estimation of a set of coefficients, one per explanatory variable, which mediate their effect on the criminal occurrences over the whole city; the coefficients may be positive or negative depending on their role of increasing or decreasing crime risk. Thanks to the model, it is possible to compute for each spatial unit (street, street segment, block, etc.) a number of expected events based on the values of the contextual variables and to compare these expected events to the actual number of recorded crimes. In principle, even with important objections that, for the sake of simplicity, it is unnecessary to delve here, this difference among these quantities is a measure of goodness of fit of the model. As a general rule, if the criminological hypotheses fit well to the specific study area, most of the roads should have an expected number of criminal events that is close to the actual number of occurrences. On the contrary, high discrepancies among these values may identify situations with far fewer events, or too many events than the ones expected on the pure basis of the context variables values. The first situation suggests the presence of unspecified favorable conditions unaccounted for by the model: some relevant

[1] Possible examples are demographic elements (e.g., number of residents per age interval, gender, and nationality), socio-economic indicators (e.g., house values acquired from the Land Registry, aggregate tax return values, number of shops, number of gambling halls, number of bars and pubs, etc.), or configurational dimensions of urban spaces as they result, for instance, from the Space Syntax analysis (centrality of the street in the urban network, pedestrian movement, number of intersections of the street with other streets, etc.) or from the CCTV measurements.

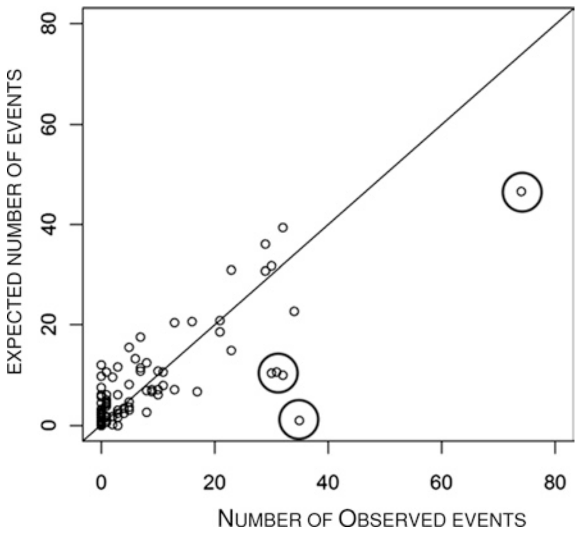


Fig. 3 Expected versus Observed arson and criminal damage occurrences on parked vehicles in a neighborhood of an Italian city in a 24 months period. Circled are the "soccer stadium effect" roads

factors that are omitted from it seem to actually mitigate crime. The area, in this case, is worth a specific investigation as its crime-mitigating characteristics might be be reproducible elsewhere in the city and used as positive experiences, as long as a later on-the-field analysis is able to identify the positive factors at play. In the second situation, we have a number of events that is much higher than expected and there are some elements, unspecified in the model, increasing the actual risk of the road. Figure 3 shows an example of a comparison between observed and expected values for the number of damage recorded for each street in a neighborhood of a city in northern Italy. The more the points (representing the single streets of the neighborhood) are positioned along the diagonal, the better the matching of the expected versus observed values. On the contrary, the more they move away from the diagonal, the stronger the effect of the contextual elements not included in the statistical model. In a Smart Security System, a statistical module with the characteristics herein described and whose skeleton structure is given in Fig. 4 allows local administrations to identify critical situations for which customized solutions are needed.

In addition to this, the literature of statistical methods for the analysis of crime is very vast and may be helpful to investigate the effect on crime of intervention policies, of new technologies, of social change or urban planning. For a recent review of these techniques, refer to [48]. The statistical tools can be properly customized to answer the needs of any municipality. However, it is relevant that when a Smart Security system is implemented, the key points that the Public Administrators need to check and monitor are clear and well defined in the system design phase: statistical analysis does not make sense by itself; it should be shaped around the information needs of law enforcement agencies.

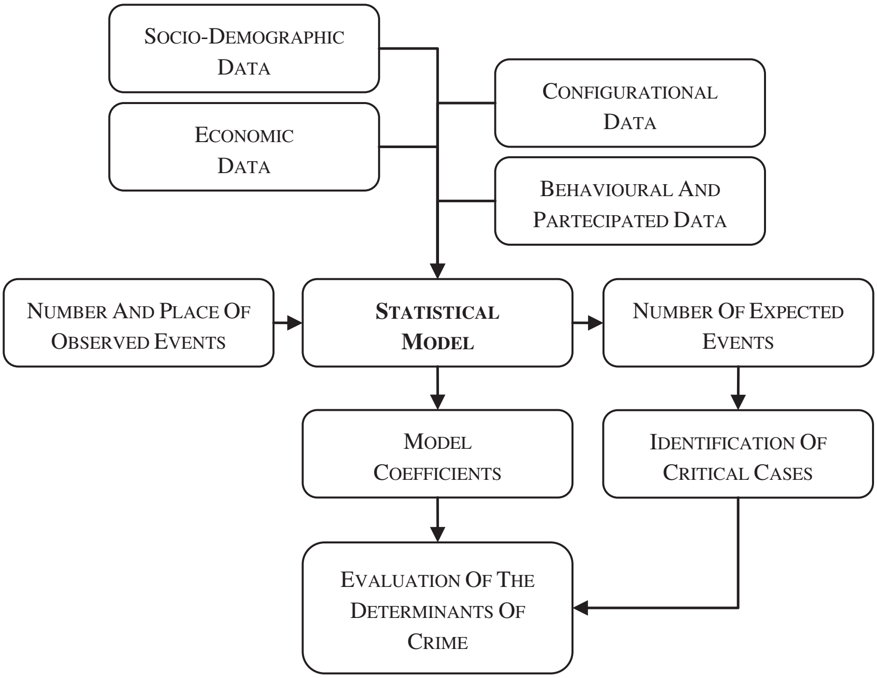


Fig. 4 The role of statistical modeling in the analysis of crime

4 Individual Behavior and Sensors

Sensors are supposed to have a crucial role in Smart City [49, 50] and the domain of Smart Security benefits from theories and practices concerning the use of sensors for crime control that predate the concept itself of Smart City by a few decades. Optical sensors are the most obvious example: the first experiments of video surveillance systems for crime prevention date back to the 80s. However, motion detectors, acoustic detectors (like gunfire locators) and even biological and chemical sensors all have been considered for their potential in crime prevention and repression strategies. The rationale behind the use of sensors in a crime prevention environment has usually been that of detecting individual behaviors, with the purpose of collecting evidence (in a forensic perspective), directing prevention or repression efforts against crime acts or deterring crime altogether by virtue of the mere possibility of collection of evidence and activation of preventive and repressive actions. In a Smart Security environment, the value of evidence collected through sensors is assumed as a given in the same way crime maps and trends are. Smart Security begins where the benefits of preventing crime, instead of repressing it, come into play. The evolution of video surveillance is paradigmatic with respect to the problem of deterrence, repression and prevention. For many years CCTV systems have been very controversial and their effectiveness for crime prevention has been questioned. While law enforcement agencies worldwide have been investing for years in CCTV as a crime-fighting technology and the technology behind CCTV rapidly developed and cameras proliferated, supporters of CCTV have typically argued that cameras make cities safer but recent studies have called into question this claim. According to some, their effectiveness might be limited and their impact on citizens' sense of security might be the opposite of what governments intend [51–56]. Surveillance systems have been welcomed by public administrations for monitoring purposes (parking lots, public transports), for access control (automatic car plate reading, etc.) or transport security [55]. In July 2005, during the attack to the London subway system the public video surveillance system installed allowed the authorities to identify the bombers and trace their paths. The system did not prevent the attacks but its help in subsequent investigations was priceless. This event encouraged public administrations to invest on video surveillance systems to prevent crimes and terrorist attacks. Since then, the scale of video surveillance networks has increased in scale [57] and today installations of 50,000 camera networks have been reported. The Singapore transport network is monitored by a 6,000 cameras network and, in general, most urban centres can count on camera networks of dozens of cameras.

These large systems are usually connected to centralized control centres, where a human operator interacts with dozens (or hundreds) of sensor sources using several separate monitors/windows for visualizing and analyzing the video or data streams. Although each separate source produces useful data, the human operator is easily overwhelmed with the task of integrating these varied forms of data into a complete global view and understanding of a scene.

This scenario will soon become obsolete thanks to the technological progress of intelligent systems and algorithms [57]. Indeed, the proliferation of surveillance cameras throughout public places stimulates the development of software able to monitor automatically the large amount of video footage produced. Human operators cannot monitor such a vast volume of data. This means that today most large installations have a limited effectiveness because of the lack of means to interrogate the content of the data generated. Once a camera network is installed, it is important to estimate the topology of the network to learn the relative positions of the cameras and the possible intersections between fields of view. This simplifies various tasks, among all an effective tracking of people within the space monitored by the network. The topology cannot be estimated manually if the network is large. Automatic procedures may also be applied to facilitate the design of the network: locate optimal positions of the camera for a maximum coverage [58].

Such networks are often heterogeneous as they often include cameras installed by the public administration specifically for the purpose of public security, plus private camera networks that may usefully complement the available information, such as cameras installed by ATMs, banks, stores, etc. This heterogeneity on the sensors, the transmission, and compression protocols, causes additional problems, producing asynchronous videos (e.g., [59]) and variable resolution signals.

Modern camera systems are able to control large areas, to zoom in (with optical zooms as in PTZ cameras or digitally, with mega or giga-pixel cameras), but also to detect moving objects and track them along the scene [60]. These systems perform real time analysis and, more importantly, record video footage for later use.

Information acquired by multiple cameras may be merged with the purpose of tracking moving objects across the views [61, 62]. If cameras have a field of view overlap one may associate corresponding view simultaneously. If the cameras have no field of view overlap moving objects may be associated along time based on an analysis of their similarity and on prior knowledge on the cameras mutual positions [63, 64].

In crime prevention, video surveillance is closely connected to biometry, since the ultimate goal is often to associate a face to the person who perpetrated the felony. Face biometry (i.e., the ability to associate automatically an identity to a face portrayed in an image or image frame) is particularly attractive, since it does not need any specific sensor but can be applied to the output of a high resolution video stream. The research community has been very active on this respect, addressing face recognition from different perspectives (e.g., [65]). Although the achievements on face biometry in the last decades are impressive, satisfactory results can be obtained mainly in constrained scenarios or with a relatively small set of enrolled identities and, for this reason, the use of face recognition in urban environments is still limited [66, 67].

Early intelligent video surveillance systems were able to detect the presence of people in forbidden zones. This was the extent of the forbidden/dangerous action taken into consideration. Nowadays, we are concentrating instead on dangerous behaviors of people and crowds [68–70]. More recently, the interest of the research community has been directed towards intelligent systems able to learn models of normal activities from long time observations and to apply them to detect anomalies in an adaptive way [71–74].

5 Where All That Is Observational Converges: Smart Security as a Preventive and Early Warning System

A Smart Security system should be designed to work at the point of convergence of multiple information sources. From what has been discussed so far, it is clear that some sources are "cold" data collected by various structures of the Public Administration; others are hot and consist of live raw or processed information coming from sensors situated in specific locations in the urban area.

An additional and very important source is a hybrid of the two: crowdsourcing[2] allows local governments to receive massive amounts of data, reports and contents generated via smartphones and the internet in general [76, 77]. Crowdsourcing can

[2] From a terminology standpoint, this entire field is still lacking consistency. We make use of the term "crowdsourcing" in its more general meaning of an organization outsourcing specific tasks (like producing goods, services or information) to vast crowds of unrelated individuals instead of using traditional employees or suppliers. As a matter of fact, the term is frequently associated with the generation of web contents because that was the first practical application of crowdsourcing [75], but a broader meaning should be acceptable as well. Specific forms of crowdsourcing that are particularly significant for Smart Security systems have specific names, like Crowdsensing or Smartsensing, that imply the use of ubiquitous sensors (mostly smartphones) to collect data.

integrate, in a vast number of fields, the traditional information used by the Public Administration [78–81] and it surely can mitigate the cost of building large networks of sensors while producing information that, being collected on the end user-side of the public services supply system, can be much more contextual (i.e., rich with information about what is being sensed, where and when, beyond what a single sensor is normally expected to capture).

A much debated issue concerning data from smart sensing tools is that of privacy. While this is a very serious and relevant problem, it is not substantially different from that of privacy with all the rest of geo-localized or remotely-sensed information that local governments already use (e.g., in G.I.S. systems). So, while the specificities of smart sensing have to be considered also under a privacy perspective and while privacy is obviously an issue when a smart system uses data concerning individual citizens, the hypothesis of using such systems seems to mostly require specifications and not innovations of existing privacy rules. Privacy and anonymity issues influenced the spread of public video surveillance systems [82, 83]. In most countries, current legislations do not prevent abuse or misuse of video footage. Misuse can be perpetrated by individuals with an access to the video stream or by organizations. While the debate is still open, to some extent, technology is offering different ways of protecting the privacy of citizens: face detection or text detection can be used to anonymize video footage [84], video encryption technologies allow us to protect video sources [85]. If these filters are implemented within the sensors, thanks to the use of embedded systems, then the video stream is protected from the source and can be transmitted safely.

Crowdsourcing is a significant addition not only because of the scope of its reach but also because it shows that a rigid distinction between hot and cold information limits the smartness of a system. Live sensors should be used to generate cold data as well [86]. Statistical analyses over time periods should help decoding the meaning of what a live sensor is capturing. In broader terms, in a Smart Security system there is relative continuity and exchange of information between the analytical environments of what has happened in weeks, months or even years and what is happening now or is going to happen in a few minutes. From the point of view of a Local Government or that of any local branch of the Public Administration, Smart Security is, in fact, an early warning system (or the premise of it) precisely by virtue of this integration of information relative to different timeframes. Early warning systems (EWS) are "The set of capacities needed to generate and disseminate timely and meaningful warning information to enable individuals, communities and organizations threatened by a hazard to prepare and to act appropriately and in sufficient time to reduce the possibility of harm or loss".[3] EWSs have been implemented in many fields, from disaster management and prevention to epidemiology, drug control, poverty reduction, drought and famine prevention, armed conflict prevention and so on. In the field of crime prevention, EWSs have been used to organize

[3] United Nations, Office for Disaster Reduction (UNISDR).

policing [87–90] and to predict individual behaviors [91]; while the concept is popular, however, its application in complex governance problems is only becoming feasible in current Smart City environments.[4]

In a Smart Security system, statistical tools, sensors and crowdsourcing information, integrated with each other, produce an output that consists of the synthetic results of the analysis performed by each, and of a system of flags that appear in front of the system managers when certain trigger conditions are met. For example, it may happen that the recent history of a place shows a particularly intense spatiotemporal concentration of crimes, or that the trend of its socio-economic and demographic characteristics that are likely determinants of crime may hint at a probable increase of the risk of crime. The objective of the smart tool is to communicate what the flag is about in simple, unambiguous, and exhaustive fashion, adopting output representations that can be easily interpreted by city officials that are responsible of the decision making process. More precisely, flags should be designed to be the first element of the decision making process at the end of which the Public Administration produces a policy change or an action of some sort to improve urban security conditions. Given these requirements, a smart tool for urban security adds to the units of analysis an interface for the management and the representation of data that is built around three distinct elements: a crime map, a dashboard, and a warning system. The crime map is the most basic level of the entire system; it is meant to allow the spatial representation of crime but can as easily be used to map relevant context variables, in particular when they show some correlation with the presence of crime or to illustrate composite indicators. Since one of the objectives of the unit performing the statistical analyses is modeling urban crime and then showing the difference between estimated and observed values, such estimated values and difference of values are two particularly significant examples of composite indicators. The crime map can have any sort of definition level, from that of large administrative subdivisions to that of a single street or street segment. Since the main objective of the map is to make apparent any geographic effect at play, it has to show how the concentration of each relevant variable changes from place to place, making the dislocation of high and low values more important than the values themselves. This usually means that the value of a variable in each geographical unit is synthesized through one out of a finite palette of colors (four to ten in most cases) and the overall chromatic patchwork created by the map should give, in a glance, the idea of dispersion and concentration. Dashboards represent the second level of the interface. They provide a different method to read the values synthesized on the map with less emphasis on the spatial effects and greater emphasis on ranking and prioritization, discrimination, and detailed comparison. Dashboards are intended to quantify the measure of significant variables, usually within a graphical representation that

[4] See [92] for a current commercial example. Similar examples can be found concerning predictive policing and disaster management.

helps interpreting the value, for example by adding a scale of colors ranging from green to red depending on the measured value. While apparently simple, dashboards imply some intricacies: the top and the bottom of the scale may be fixed or depend on historical longitudinal observations or on current cross-sectional values, with the average and the thresholds between low and average and average and high that change accordingly. Obviously, the difference is not only the different outcome but also the different meaning: measuring a value against its historical highs and lows is different than measuring it against the values of the same variable in different places. A dashboard can also help visualizing the difference between the current value and the value recorded in the previous time unit, from a few hours or days to months before, giving an immediate representation of change. The numerical values of a variable, of its change over time or its difference with the values in other places, allows decision makers to set priorities for their actions on one issue or to balance the effort between different issues. Dashboards allow to easily identify and list places where the value of a variable is above or beyond a certain level and to disentangle the effects of different explanatory variables on a dependent variable. This makes possible to understand which risk factor is high where actual crime is high or which risk factor is responsible for making expected crime high. Dashboards should be contextualized as much as possible: since they are an extremely synthetic tool, the user should be given as much information as possible on the characteristics of the place that the dashboard refers to, so that the reading is not left as an abstract and inexplicable value. Usually, maps and dashboards contain quantitative information on places and population. However, smart sensors and cameras, while primarily oriented to analyze individual behavior, are also a source of cumulated individual behaviors. Therefore, if a smart tool for urban security is built around them, a considerable number of variables that can be represented in maps and dashboards can actually come from a database of what was captured by smart sensors. The last element of a smart tool is a warning system. It can exist as a specific element of the tool or it can be integrated within the map and the dashboard. Its function is to help the user at noticing critical situations even when they are hidden in a large amount of information, indicating it with a flag, i.e., a specific and visible signal of some sorts. Flags may be the consequence of slow, gradual processes that progressively increase risk at a certain place beyond a given level. They may come from sudden increases, from cyclical peaks and they may as well depend on individual behaviors that are excessively distant from the average or from an accumulation of many concurrent and slightly anomalous behaviors of different people. Flags are not particularly sophisticated instruments; they are based on threshold values that trigger them when the reading goes above or below. The sophisticated part of a warning system is the balancing of the thresholds, of the sensitivity of the triggers, and the ability of the system to react to changes by updating its thresholds over time. Obviously, the objective is minimizing false positives as well as false negatives, keeping in mind that a smart tool is not a substitute of decision makers but just a support system and, consequently, whenever it is possible, flags and warnings should stimulate a cross-checking of results and an on-the-spot investigation before any actions are taken.

6 Handling Complex Systems: The Integrated Network System

Integrating information from different sources is a very complex activity, especially if the data sources are very different from each other (e.g., text, video, audio, etc.). To simplify the integration processes, sources other than text often need to be enhanced through the manual or automatic generation of meta-data that is a textual description of the content of the data source (e.g., the name of the people in a video, the date of the data collection, a transcript of the conversations from the audio, etc.) [93]. However, even in the simplest case in which we need to integrate only text-based information, the activity can present several challenges.

Moreover, there are several kind of information coming from different sources that can be integrated and used to improve the situation awareness (e.g., weather, air quality, light, etc.) that can provide a constant (and frequently large) stream of data. The amount and the heterogeneity of such data is extremely difficult to manage with the traditional approaches based on OLAP (On-Line Analytical Processing) and data warehouses [94]. To this end, new approaches have emerged and classified under the label *big data* and implemented through the so-called NoSQL databases [95].

On-line analysis of data is also required to ensure the reliability of sensors used for the data collection to identify immediately problems that may prevent useful subsequent analysis and integrations with other sources. Such analyses include simple statistical evaluations of the quality of the data and complex ad-hoc analyses based on the information coming from different sources that are related to each other and can be used to crosscheck their validity. The quality of the data collected is the starting point for implementing an effective integration reducing false positives and false negatives, therefore an alerting system based on on-line analysis can help in such activity.

To have complete information integration on which it is possible to develop reliable applications, it is required that the integration is implemented at different levels: communication, syntactic, and semantic [96]. The communication level deals with the technical aspects of the data transfer among the different systems involved in the integration; the syntactic level deals with the data formats and the transformations required to create a common representation of the information; finally, the semantic levels deals with the meaning of the different pieces of data and how to relate each other.

At each level, there are several challenges including the followings:

Communication level	size of the information and technical implementation
Syntactic level	kind of information and storage
Semantic level	organization of the information

The organization of the information deals with the ability to define high-level (and generally abstract) concepts and connect all pieces of information related to that concept. For instance, considering a robbery, there are several pieces of

```
<AGENDA>
  <PERSON>
    <NAME>JOHN</NAME>
    <SURNAME>SMITH</SURNAME>
    <EMAIL>SMITH@ENT.COM</EMAIL>
  </PERSON>
  <PERSON>
    <NAME>TOM</NAME>
    <SURNAME>BROWN</SURNAME>
    <EMAIL>TB@BROWN.COM</EMAIL>
  </PERSON>
</AGENDA>
```

```
<LIST>
  <CONTACT>
    <NAME>JOHN SMITH </NAME>
    <EMAIL>SMITH@ENT.COM</EMAIL>
  </CONTACT>
  <CONTACT>
    <NAME>TOM BROWN </NAME>
    <EMAIL>TB@BROWN.COM</EMAIL>
  </CONTACT>
</LIST>
```

Fig. 5 Incompatible structured information providing the same content through different structures

information from different sources that can be related including: the timeframe from the police reports, the suspected invited people from the investigation records of the police, the video of the surveillance cameras, etc. Such integration is very difficult to perform automatically and requires an extensive amount of research to be implemented in a general context even if in some very restricted domains it is feasible with the current technologies that are part of the so-called *semantic web* (even if the term includes the world “web”, the technology is not used only for the web but it is the domain where it comes from) [97].

The kind of information refers to its structure. We can classify information in two large sets: unstructured and structured. Unstructured information is any kind of text designed with human beings in mind (e.g., this book). On the contrary, structured information designed to be processed and stored easily by a machine through a database and exchanged using a semi-structured form that includes special markers (called *tags* in many languages used for this purpose) that makes processing possible. Languages like HTML (HyperText Markup Language) and XML (eXtensible Markup Language) are very popular in any kind of document-based representations (not just on the web for which they were conceived at first) and are based on such a concept to make the interpretation, the visualization, and the storage of information easier. However, even with structured and semi-structured information, the integration of different data sources is not straightforward since each source may use a different set of tags and organize the information in different ways. In Fig. 5 an example of two incompatible structures of the same information is given. Moreover, it happens very frequently that the differences among data sources are not just syntactic differences (Fig. 5) but also semantic ones (e.g., the same tags used to identify different content in different documents, different in information sets provided, etc.). Therefore, integrating different data sources requires a deep knowledge of the data representations and requires a considerable effort. However, given the importance of the applications that are based on information integration, there is an enormous amount of research in the area aiming at automating the integration as much as possible [98, 99]. One of the current trends in research about information integration is based on the development of ontologies that allow automatic conversion mechanisms and highlight incompatibilities [100].

How the information is stored and its size are additional aspects that need to be considered when dealing with different sources of information. Currently, every activity (human-based or machine-based) produces a large set of digital information stored in several databases. Such databases are huge, therefore transferring or copying the entire data sets to perform complex operations is often unfeasible. Beyond such problems, we also have to consider the sensitivity of some kind of data and/or the privacy aspects related to them. In such cases, a sanitization procedure is often required before allowing other kind of operations and analysis on such data removing the sensitive part of the data and/or aggregating at a higher level with no privacy or sensitivity concerns.

Therefore, it is required to develop on-line analysis techniques that are able to process and integrate information on the fly (whenever such information becomes available) and exchange only the relevant data without overloading the communication infrastructure. Moreover, relational databases that are often used to store information struggle in managing such large amount of data if there is not an adequate investment in the hardware infrastructure. As stated before, traditional approaches through data warehouses are not able to address properly this kind of problems, therefore NoSQL databases are emerging offering better performances and scalability at a much lower cost at expenses of some properties of the relational databases that can be relaxed in some application contexts. These technologies have been designed to address problems related to the storage of large data sets but their correct usage is linked to the specific problems the application has to address.The technical implementation is basically related to the usage of specific technologies. In the well-known world of relational databases, there are standards that are accepted by almost any implementation such as the SQL language to perform interrogations and insert/modify data. However, in the NoSQL world, there are no common standards for even basic operations and each implementation has its own approach producing two main effects: 1) it is difficult to switch from one technology to another and 2) every technology requires a complete set of new skills. For this reasons, the use of NoSQL technologies need to be considered only in specific cases since it may be difficult to fix some mistakes in the selection of the right technology to use.

There are plenty of open source technologies that can be used to implement such systems (databases, analysis and visualization tools, sensors, etc.) producing a set of advantages such as the absence of a license fee, the ability to adapt the tools to the specific needs, no vendor lock-in, etc. Moreover, when dealing with problems related to integration, security, and privacy, the usage of open data formats, protocols, and tools help in identifying bugs, assuring the absence of malicious code and enhance the overall interoperability and the level of integration of different systems.

From the architectural point of view, integrating several data sources at the same time is extremely complex due to the main problems described earlier. However, the technologies available today allow developers to split the problems in several smaller problems that are easier to address and integrate them only later on. In this way, it is possible to create a more scalable architecture able to integrate an arbitrary number of data sources limiting the complexity of their integration. In any case, even with just two data sources, the three level of integration (communication, syntactic, and semantic) should be taken into consideration to provide a meaningful integrated system.

A specific issue related to the integration of video sources requires a reference to interfaces, in particular when IP video cameras are concerned. Over the years the main producers developed various standards, currently the main one is ONVIF, founded by Axis, Bosch and Sony. ONVIF is about (1) standardization of communication between IP-based physical security devices and (2) interoperability between IP-based physical security products regardless of manufacturer. It is also worth mentioning HD-Serial Digital Interfaces (SDI), a family of digital video interfaces used for transmitting uncompressed, unencrypted digital video signals within analog television facilities. This technology has been conceived with the goal of bridging the gap between analog systems and digital installations over IP.

7 A Good Start: Roadmaps Towards a Smart Security

With all the different issues now on the table, we can conclude our work with an attempt at drawing guidelines for the implementation of a Smart Security system in a Public Administration context. Smart Security tools may have different levels of complexity, having to comply with different technical, administrative, and economic limitations (see Table 1 for a few documented examples), but some elements in their infrastructure and implementation are going to define if and how much they can actually be considered smart.

The first and crucial element that defines the smartness of a crime prevention system is that it should be built around the management and policy needs of the Public Administration and not as a retrofitting of them. "Technological performance is not to be taken for granted as a logical progression from technological advancement, but rather performance depends on effective management of technological systems and infrastructure" [1]. The bottom-up process of influential projects like COMPSTAT [101, 102], GeoArchive [87] and the general effort to introduce GIS as a crime prevention tool [4, 103] constitute very good examples.

Being an early warning approach, Smart Security requires an "early response" organizational framework as well. This means that, regardless of the source of the information that generates the warning (be it a statistical analysis, a live sensor, a crowdsensing tool or a triangulation combining any of them) the organizational goal must be that of having the resources required to prevent the issue and the determination and ability deploy them in a timely fashion.

Table 1 Existing projects and software containing significant smart elements

Notable examples	Smart elements
Compstat [101, 102, 104, 105]	Organizational focus, bottom-up development, data collection, mapping, statistical analysis, early warning philosophy, results evaluation
GeoArchive [4, 87]	Organizational focus, bottom-up development, data collection, mapping, statistical analysis, early warning philosophy
SACSI [106–108]	Data collection, mapping, statistical analysis, results evaluation
COMPASS [109]	Data collection and data sharing, mapping, statistical analysis, decision support, results evaluation
Operation virtual shield [100, 110]	CCTV, early warning philosophy
G.I.S.-based free and commercial software [103]	Mapping crime and context (with elements of statistical analysis)
Urban crime simulator [111]	Crime modeling based on criminological theory
Desurbs [7]	Organizational focus, data collection, urban planning and design focus, mapping, statistical analysis, decision support, integration
Commercial software (PredPol, IBM Spss and BlueCrush, Esri, …)	Mapping, statistical analysis, predictive policing

In terms of components, a Smart Security system is scalable according to evolving needs and consists of part or all of the following key elements:

1. Relevant administrative databases (e.g., crime records, socio demographic and economic data, urban graph);
2. Sensor network(s);
3. Crowdsourcing applications and websites;
4. Crime maps and trends visualization;
5. Security dashboard;
6. Intelligence module (Data integration and analysis).

The entire system, in order for the early warning mechanism to work, has to be integrated inside a single user interface with coordinated warning flags. However, of all these elements, some may already be in use in many local governments and just need to be integrated in the new system, while others require a greater deal of work. In spite of this, incremental developments are possible and, in many ways, superior to the "all-or-nothing" approach. Another crucial point is that Smart Systems are, by definition, tailored locally and, consequently, they do not necessarily need every element of this list. Some smaller settlements may never have enough data to justify a complex statistical tool. Some may have little need for having both a crime map and the dashboard. Sensor networks are a useful addition where and when their effectiveness is documented and sensors are worthwhile if there is a precise idea of what use to make of the data collected through them.

With respect to point 1, the main problem which may arise is technical, due to the existence of databases which are not normalized and which make difficult their querying or joining. Classical examples are different geographical boundaries of the statistical units, different levels of aggregation of data, different definitions of the same variable, coding errors caused by fields that are sensitive to spelling mistakes or different forms of abbreviation. Concerning point 2, most of the Municipalities interested into the implementation of a Smart Security system already have a CCTV system of cameras installed and it is usually reasonable to integrate an existing infrastructure into the framework of a smart system whenever possible. However, there is no guarantee that the technological standards and the aims of such an infrastructure are ultimately compatible with a smart system. Crowdsourcing and smartsensing projects (point 3) are currently being developed in some of the most advanced and innovative municipalities around the world, but compared to other elements of the system, here the emphasis should be on designing them with compatibility with a Smart Security environment in mind from the beginning. Whether they ask users to produce content, ideas, information or else, in a proactive creative process or they just ask them permission for capturing opportunistic information in a passive, "authorize and forget" manner, they make sense as an element of the system if they fill significant information areas with reliable data that can be confronted and integrated with data already available.

The Crime Mapping System (point 4), the Security Dashboard (point 5) use the administrative data and offer different kinds of graphical and numerical representation. They can occasionally be developed starting from existing municipal Geographic Information Systems (GIS) and/or linked to databases and other tools that already provide synthetic tables of information. There is a multiplicity of possible software combinations that answer the needs of each specific context, including open source solutions. The same holds for statistical software packages and, ultimately, decisions should be based on compatibility with pre-existing infrastructures and instruments and with the specific characteristics and requirements of each local government.

The Software interface (point 6) is a technical need for the setting up of the system. As a matter of fact, it can be intended into two ways. On one hand, it is the container inside which all the queries are executed, the datasets connected, and the computations done using the dedicated tools and packages. On the other hand, it is the tool which gives the output to the final user in an interactive and easy to use interface. In fact, the final goal of a Smart Security system is to assist the Public Administrators and law enforcement agencies to understand a fast changing world and to implement the most effective security policies. The software system is the environment inside which the automated procedures defined by the experts are repeated automatically without the need of the final users to possess advance competences of statistics, video analysis, or software engineering. Obviously, given the sensitive nature of the data, security and control over the system and the information in it is crucial.

Table 2 gives a general overview of what we discussed in this final section and outlines what is needed for each component of a Smart Security. Starting from what we indicate in Table 2, any Municipality or Law enforcement agency can find its own roadmap towards a Smart Security System. Note that the Warning

Table 2 Requirements table for the key elements of a Smart Security System

Element of the Smart Security System	Requirements						
	Administrative database	Sensor network	Crime mapping	Security dashboard	Intelligence module: statistical analysis	Intelligence module: sensors analysis	Intelligence module: data integration
Crime mapping	Yes	No		No	No	No	Yes
Security dashboard	Yes	No	No		No	No	Yes
Statistical analysis module	Yes	No	Yes	Yes		No	Yes
Sensor analysis module	No	Yes	No	No	No		Yes

System is not listed in Table 2, being a very advanced feature of the system has various requirement as it must be tailored on the specific needs of the users.

Finally, intelligent solutions are ways to optimize the capacity, efficiency, and sustainability of a system. Typically, by means of ICT-based information processing. Smart technology is not, in itself, enough for a smart solution if users and operators are not involved in a learning process and the institutions that will use the system need to be changed as well. The system design should not focus on the smart infrastructure alone and not only on the final goal, but rather the transition phase itself should be designed carefully, with much attention for intermediate and hybrid stages where sometimes the flexibility gained from the intelligent solution can already be put to use [11].

References

1. Nam, T., & Pardo, T. A. (2011). Smart city as urban innovation: Focusing on management, policy, and context. In *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance* (pp. 185–194).
2. Harrison, C., & Donnelly, I. (2011). A theory of smart cities. In: *Proceedings of the 55th Annual Meeting of the International Society for the Systems Sciences (ISSS)*, *Hull* (Vol. 55, pp. 1–15).
3. Van den Berg, L., Pol, P. M. J, Mingardo, G., & Speller, C. J. M. (Eds.) (2006). *The safe city: Safety and urban development in European cities*. Farnham: Ashgate Publishing.
4. Weisburd, D., & McEwen, T. (1998). *Crime mapping and crime prevention*. New York: Criminal Justice Press.
5. Weisburd, D., Mastrofski, S. D., Greenspan, R., & Willis, J. J. (2004). *Growth of Compstat in American policing*. US Department of Justice: National Institute of Justice.
6. Steden, R., Boutellier, H., Scholte, R. D., & Heijnen, M. (2012). Beyond crime statistics: The construction and application of a criminogenity monitor in Amsterdam. *European Journal on Criminal Policy and Research, 19*, 47–62.
7. Bonatsos, A., Middleton, L., Melas, P., & Sabeur, Z. (2013). Crime open data aggregation and management for the design of safer spaces in urban environments. In *Environmental Software Systems. Fotering information Sharing* (pp. 311–320). Berlin, Heidelberg: Springer.
8. Bettencourt, L. (2013). The uses of big data in cities. Santa Fe Institute Working Paper 29.
9. Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., et al. (2012). Smart cities of the future. *European Physical Journal: Special Topics, 214*(1), 481–518.
10. Paskaleva, K. A. (2011). The smart city: A nexus for open innovation? *Intelligent Buildings International, 3*(3), 153–171.
11. De Haan, J., Vrancken, J. L. M., & Lukszo, Z. (2011). Why is intelligent technology alone not an intelligent solution? *Futures, 43*(9), 970–978.
12. Denzin, N. K. (2006). *Sociological methods: A sourcebook*. Piscataway: Aldine Transactions.
13. Paulsen, D. J., & Robinson, M. B. (2009). *Crime mapping and spatial aspects of crime*. Boston: Allyn & Bacon.
14. Chamlin, M. B., & Cochran, J. K. (2004). An excursus on the population size-crime relationship. *Western Criminology Review, 5*(2), 119–130.
15. Andresen, M. A. (2007). Location quotients, ambient populations, and the spatial analysis of crime in Vancouver, Canada. *Environment and Planning A, 39*(10), 2423–2444.

16. Andresen, M. A. (2006). Crime measures and the spatial analysis of criminal activity. *British Journal of Criminology, 46*(2), 258–285.
17. Boggs, S. L. (1965). Urban crime patterns. *American Sociological Review, 30*(6), 899–908.
18. Cohen, L. E., Kaufman, R. L., & Gottfredson, M. R. (1985). Risk-based crime statistics: A forecasting comparison for burglary and auto theft. *Journal of Criminal Justice, 13*(5), 445–457.
19. Harries, K. D. (1991). Alternative denominators in conventional crime rates. In: P. Brantingham & P. Brantingham (Eds.), *Environmental criminology* (2nd ed., pp. 147–165). USA: Waveland Press.
20. Harries, K. D. (2006). Property crimes and violence in United States: an analysis of the influence of population density. *International Journal of Criminal Justice Sciences, 1*(2), 24–34.
21. Sparks, R. F. (1981). Measuring crime rates and opportunities for crime. In: R.G. Lehnen & W. G. Skogan (Eds.), *The national crime survey: Working papers volume I: current and historical perspectives* (Vol. 1, pp. 52–58). Washington, DC: U.S. Department of Justice.
22. Stipak, B. (1988). Alternatives to population-based crime rates. *International Journal of Comparative and Applied Criminal Justice, 12*(2), 247–260.
23. Zhang, H., & Peterson, M. (2007). A spatial analysis of neighborhood crime in Omaha, Nebraska using alternative measures of crime rates. Internet Journal of Criminology. http://www.internetjournalofcriminology.com/Zhang%20Peterson%20-%20A%20SPATIAL%20ANALYSIS%20OF%20NEIGHBOURHOOD%20CRIME.pdf.
24. Thomas, W. I. (1966). Social disorganization and social reorganization. In M. Janovitz (Ed.), *On social organization and social personality: Selected papers* (pp. 3–11). Chicago: The University of Chicago Press.
25. Guerry, A. M. (1833). *Essai sur la statistique morale de la France*. Cochard.
26. Jacobs, J. (1961). *The death and life of great american cities*. New York: Vintage.
27. Newman, O. (1972). *Defensible space: Crime prevention trough urban design*. New York: MacMillan.
28. Jeffery, C. (1971). *Crime prevention through environmental design*. Thousand Oaks: Sage Publishing.
29. Clarke, R. V. G. (1997). *Situational crime prevention*. New York: Criminal Justice Press.
30. Clarke, R. V. (1983). Situational crime prevention: Its theoretical basis and practical scope. In M. Tonry & N. Morris (Eds.), *Crime and justice: an annual review of research* (pp. Vol. 14, pp. 225–256). Chicago: The Chicago University Press.
31. Clarke, R. V. G. (1992). *Situational crime prevention: Successful case studies*. New York: Harrow and Heston.
32. Clarke, R. V. (1995). Situational crime prevention. In M. Tonry & D. Farrington (Eds.), *Building a safer society: Strategic approaches to crime prevention. Crime and justice: A review of research* (Vol. 19, pp. 91–150). Chicago: The Chicago University Press.
33. Felson, M., & Clarke, R. V. (1998). *Opportunity makes the thief: Practical theory for crime prevention*. London: Police and Reducing Crime Unit; Research, Development and Statistics Directorate. Police Research Series Paper 98.
34. Hillier, B. (1988). Against enclosure. In N. Teymur, T. A. Markus, & T. Woolley (Eds.), *Rehumanizing housing* (pp. 63–88). London: Butterworths.
35. Hillier, B., & Hanson, J. (1984). *The social logic of space*. Cambridge: Cambridge University Press.
36. Hillier, B., & Sahbaz, O. (2008). An evidence based approach to crime and urban design. In: R. Cooper, C. Boyko, G. Evans, & M. Adams (Eds.), *Designing sustainable cities: Decision-making tools and resources for design* (pp. 163–186). London: Wiley-Blackwell.
37. Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
38. Beavon, D. J. K., Brantingham, P. L., & Brantingham, P. J. (1994). The influence of street networks on the patterning of property offences. In R. V. Clarke (Ed.), *Crime prevention studies* (Vol. 2). New York: Willow Tree Press.
39. Cohen, L., & Felson, M. (1979). Social change and crime rates. *American Sociological Review, 44*, 588–608.

40. Eck, J. E., & Weisburd, D. (1995). Crime places in crime theory. In J. E. Eck & D. Weisburd (Eds.), *Crime and place*. Monsey.
41. Roncek, D. W. (1981). Dangerous places: Crime and residential environment. *Social Forces, 60*, 74–96.
42. Rice, K. J., & Smith, W. R. (2002). Socioecological models of automotive theft: Integrating routine activity and social disorganization approaches. *Journal of Research in Crime and Delinquency, 39*, 304–336.
43. Block, R., & Davis, S. (1996). The environs of rapid transit stations: A focus for street crime or just another risky place? In R. Clarke (Ed.), *Preventing mass transit crime*. New York: Criminal Justice Press.
44. Chih-Feng Shu, S. (2000). Housing layout and crime vulnerability. *Urban Design International, 5*(3–4), 177–188.
45. Hillier, B. (2004). Can streets be made safe? *Urban Design International, 9*(1), 31–45.
46. Sinkiene, J., Stankeviče, I., & Navickaite, K. (2012). Creating safer cities through urban planning and development. *Public Policy and Administration, 11*(3), 390–403.
47. di Bella, E., Persico, L., & Corsi, M. (2011). A multivariate analysis of the space syntax out-put for the definition of strata in street security surveys. In *DISEFIN Series of Economic Working Papers 5*.
48. Piquero, A. R., & Weisburd, D. (Eds.). (2010). *Handbook of quantitative criminology*. Berlin: Springer.
49. Schaffers, H., Komninos, N., & Pallot, M. (2012). Smart cities as innovation ecosystems sustained by the future internet. In *FIREBALL Project White Paper. EU* (pp. 1–65).
50. Yovanof, G. S., & Hazapis, G. N. (2009). An architectural framework and enabling wireless technologies for digital cities & intelligent urban environments. *Wireless Personal Communications, 49*, 445–463.
51. Farrington, D. P., Gill, M., Waples, S. J., & Argomaniz, J. (2007). The effects of closed-circuit television on crime: Meta-analysis of an English national quasi-experimental multi-site evaluation. *Journal of Experimental Criminology, 3*(1), 21–38.
52. Ratcliffe, J. H., Taniguchi, T., & Taylor, R. B. (2009). The crime reduction effects of public CCTV cameras: A multi-method spatial approach. *Justice Quarterly, 26*(4), 746–770.
53. Welsh, B. C., & Farrington, D. P. (2003). Effects of closed-circuit television on crime. *Annals of the American Academy of Political and Social Science, 587*, 110–135.
54. Welsh, B. C., & Farrington, D. P. (2005). Evidence-based crime prevention: Conclusions and directions for a safer society. *Canadian Journal of Criminology and Criminal Justice, 47*(2), 337–354.
55. Welsh, B. C., & Farrington, D. P. (2009). Public area CCTV and crime prevention: An updated systematic review and meta-analysis. *Justice Quarterly, 26*(4), 716–745.
56. Welsh, B. C., Mudge, M. E., & Farrington, D. P. (2010). Reconceptualizing public area surveillance and crime prevention: Security guards, place managers and defensible space. *Security Journal, 23*(4), 299–319.
57. Remagnino, P., Monekosso, D. N., & Jain, L. C. (2011). *Innovations in defence support systems—3: Intelligent paradigms in security*. Berlin: Springer.
58. van den Hengel, A., Hill, R., Wart, B., Cichowski, A., Detmold, H., Madden, C., Dick, A., & Bastian, J. (2009). Automatic camera placement for large scale surveillance networks. In *Workshop on Applications of Computer Vision*.
59. Zini, L., Cavallaro, A., & Odone, F. (2013). Action-based multi-camera synchronization. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 3*(2), 165–174.
60. Haritaoglu, I., Harwood, D., & Davis, L. (2000). W4: real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(8), 809–830.
61. Khan, B. S., & Shah, M. (2003). Consistent labeling of tracked objects in multiple cameras with overlapping fields of view. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(10), 1355–1360.
62. Taj, M., & Cavallaro, A. (2011). Distributed and decentralized multi-camera tracking: A survey. *IEEE Signal Processing Magazine, 28*(3), 46–58.

63. Gheissari, N., Sebastian, T., & Hartley, R. (2006). Person re-identification using spatiotemporal appearance. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1528–1535).
64. Zheng, W., Gong, S., & Xiang, T. (2011). Person re-identification by probabilistic relative distance comparison. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 649–656).
65. Stan, L., & Jain, A. (2011). *Handbook of face recognition*. Berlin: Springer.
66. Shachtman, N. (2006, January 25). The new security: cameras that never forget your face. *The New York Times*, Published.
67. Viola, P., & Jones, M. (2004). Robust real-time face detection. *International Journal of Computer Vision, 57*(2), 137–154.
68. Poppe, R. (2010). A survey on vision-based action recognition. *Image and Vision Computing, 28*(6), 976–990.
69. Bird, N., Masoud, O., Papanikolopoulos, N., & Isaacs, A. (2005). Detection of loitering individuals in public transport areas. *IEEE Transactions on Intelligent Transportation Systems, 6*(2), 167–177.
70. Krausz, B., & Bauckhage, C. (2011). Automatic detection of dangerous motion behavior in human crowds. In *IEEE International Conference on Advanced Video and Signal-based Surveillance, AVSS*.
71. Stauffer, C., Eric, W., & Grimson, L. (2000). Learning patterns of activity using real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(8), 747–757.
72. Atev, S., Masoud, O., & Papanikolopoulos, N. (2006). Learning traffic patterns at intersections by spectral clustering of motion trajectories. In *IROS Intelligent Robots and Systems* (pp. 4851–4856).
73. Morris, B. T., & Trivedi, M. M. (2008). A survey of vision-based trajectory learning and analysis for surveillance. *IEEE Transactions on Circuits and Systems for Video Technology, 18*(8), 1114–1127.
74. Noceti, N., & Odone, F. (2012). Learning common behaviors from large sets of unlabeled temporal series. *Image and Vision Computing, 30*(11), 875–895.
75. Estellés-Arolas, E., & González-Ladrón-De-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science, 38*, 189–200.
76. Cuff, D., Hansen, M., & Kang, J. (2008). Urban sensing: Out of the woods. In *Communications of the ACM*, Vol. 51.
77. Ganti, R. K., Ye, F., & Lei, H. (2011). Mobile crowdsensing: Current state and future challenges. In *IEEE Communications Magazine* (pp. 32–39).
78. Cardone, G., & Foschini, L. (2013). Fostering participation in smart cities: A geo-social crowdsensing platform. *IEEE Communications Magazine, 51*(6), 112–119.
79. Coric, V. & Gruteser, M. (2013). Crowdsensing maps of on-street parking spaces. In *2013 IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)* (pp. 115–122).
80. Ghose, A., Bhaumik, C., & Chakravarty, T. (2013). BlueEye: A system for proximity detection using bluetooth on mobile phones. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication* (pp. 1135–1142).
81. Heipke, C. (2010). Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 550–557.
82. Kumagai, J., & Cherry, S. (2004). Sensors and sensibility. *IEEE Spectrum, 41*(7), 22–28.
83. The Constitution Project. (2006). Guidelines for public video surveillance, a guide to protecting communities and preserving civil liberties.
84. Newton, E. M., Sweeney, L., & Malin, B. (2005). Preserving privacy by de-identifying face images. *IEEE Transactions on Knowledge and Data Engineering, 17*(2), 232–243.
85. Cavallaro, A. (2004). Adding privacy constraints to video-based applications. In *European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology*.
86. Dufour, J. Y. (Ed.) (2012). *Intelligent video surveillance systems*. London: Wiley.
87. Block, C. (1998). The geoArchive: An information foundation for community policing. In *Crime Mapping and Crime Prevention* (pp. 27–81).

88. Buslik, M., & Maltz, M. (1998). Power to the people: Mapping and information sharing in the Chicago Police Department. In *Crime Mapping and Crime Prevention, Crime Prevention Studies* (Vol. 8).
89. Cohen, J., Gorr, W., & Olligschlaeger, A. (2007). Leading indicators and spatial interactions: A crime forecasting model for proactive police deployment. *Geographical Analysis, 39*(1), 105–127.
90. Cohen, J., & Gorr, W. (2005). *Development of crime forecasting and mapping systems for use by police*. Pittsburgh: H. John Heinz III School of Public Policy and Management, Carnegie Mellon University.
91. Cocx, T. K., Kosters, W. A., & Laros, J. F. J. (2008). *An Early Warning System for the Prediction of Criminal Careers. MICAI 2008: Advances in Artificial Intelligence* (pp. 77–89). Berlin, Heidelberg: Springer.
92. Andersen, J. J. (2013). Assess the urban surveillance infrastructure: Develop a framework. In *IBM Development* (pp. 1–11).
93. Rezaei, A., Rossi, B., Sillitti, A., & Succi, G. (2012). Knowledge extraction from events flows. In G. Anastasi, E. Bellini, E. Di Nitto, C. Ghezzi, L. Tanca, & E. Zimeo (Eds.), *Methodologies and technologies for networked enterprises*. Berlin: Springer.
94. Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling*. London: Wiley.
95. NoSQL Archive. http://nosql-database.org/.
96. Predonzani, P., Sillitti, A., & Vernazza, T. (2001). Components and data-flow applied to the integration of web services. In *IEEE Conference on Industrial Electronics Society* (Vol. 3, pp. 2204–2207).
97. Semantic Web. http://semanticweb.org/.
98. Scotto, M., Sillitti, A., Vernazza, T., & Succi, G. (2001). Managing web-based information. In *5th International Conference on Enterprise Information Systems* (Vol. 1, pp. 575–578).
99. Sillitti, A., Scotto, M., Succi, G., & Vernazza, T. (2003). News miner: A tool for information retrieval. In *7th International Conference on Intelligent Engineering Systems*.
100. W3C Standards Semantic Web. http://www.w3.org/standards/semanticweb/ontology.
101. Moore, M. H., & Aa, Braga. (2003). Measuring and improving police performance: The lessons of Compstat and its progeny. *Policing: An International Journal of Police Strategies & Management, 26*, 439–453.
102. Weisburd, D., Mastrofski, S. D., Greenspan, R., & Willis, J. J. (2004). *The growth of compstat in American policing*. Washington, DC: Police Foundation.
103. Chainey, S., & Ratcliffe, J. (2008). *GIS and crime mapping*. London: Wiley.
104. Harris, K. D. (1999). *Mapping crime: principles and practice*. Washington, DC: U.S. Department of Justice.
105. Kennedy, L. W., Caplan, J. M., & Piza, E. (2010). Risk clusters, hotspots, and spatial intelligence: Risk terrain modeling as an algorithm for police resource allocation strategies. *Journal of Quantitate Criminology, 27*, 339–362.
106. Groff, E., Fleury, J., & Stoe, D. (2001). *Strategic approaches to community safety initiative (SACSI): Enhancing the analytic capacity of a local problem-solving effort*. Washington, DC: National Institute of Justice.
107. Pattavina, A., Pierce, G., & Saiz, A. (2002). Urban neighborhood information systems: crime prevention and control applications. *Journal of Urban Technology, 9*, 37–41.
108. Roehl, J., Rosenbaum, D. P., Costello, S. K., Coldren, J. R, Jr, Schuck, A. M., Kunard, L., et al. (2008). *Brief paving the way for project safe neighborhoods: SACSI in 10 US cities*. Washington, DC: U.S Department of Justice.
109. Boba, R. (2009). *Crime analysis with crime mapping*. Beverley Hills: Sage Publications.
110. Monahan, T., & Mokos, T. J. (2013). Crowdsourcing urban surveillance: The development of homeland security markets for environmental sensor networks. *Geoforum, 49*, 279–288.
111. Lee, J., & Yu, C. (2010). The development of Urban Crime Simulator. In *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application—COM.Geo'10* (p. 1). New York, USA: ACM Press.

The Co-production of Social Innovation: The Case of Living Lab

Anna Cossetta and Mauro Palumbo

Abstract Our article aims to reflect on some key concepts that have emerged in the recent literature on innovation. In particular, it will seek convergence between social and open innovation within the framework of Smart Cities. The Smart cities are embedded in the last 20 years processes of change that have altered conditions and modalities of innovation and knowledge generation. The city is still, like Robert Park in 1915, the "social laboratory" par excellence for the study of human behavior in a modern urban environment. If we consider recent debate on Smart city definition, we can find that ICT can be a powerful tool for building the collaborative digital environment that enhances the intelligent capacity of localities [30]. In that sense we can consider use the most used definition: "a city may be called smart when investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic development and a high quality of life, with a wise management of natural resources, through participatory governance". Early as at this definition we can find the pillars of our reflection: the innovation as social innovation, the new role of the 2.0 citizen–public, the issue of governance.

Keywords Innovation • Open innovation • Triple helix • Quadruple helix • Living labs

A. Cossetta (✉) · M. Palumbo
Department of Education, University of Genoa, Genoa, Italy
e-mail: anna.cossetta@gmail.com

M. Palumbo
e-mail: palumbo@unige.it

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3_11,

1 Innovation

Innovation seems to be the most urgent need of our society. Innovation, said recently Edmund Phelps is the only antidote to the crisis but also to inequality. A true elixir to satisfy the changing needs of an ever more personalized (and wounded) world.

The post Fordism, the rise of knowledge and creativity economy, the radical change in factors of production (raw material, labor, capital) support an additional power of knowledge creativity.

"In an essential sense, innovation concerns the search for, and the discovery, experimentation, development, imitation, and adoption of new products, new production processes and new organizational set-ups" [16]. This neoschumpeterian definition suggests us to consider innovation as a result of productivity efficiency and adaptive efficiency. Innovation is a social fact driven by individuals as well as large institutions, associations, online or offline community and so on.

Innovation, in polanyian word, is embedded in society: this is the starting point of the large recent literature on social innovation.

2 Social Innovation

If we read some definition of social innovation, we can find, for example: A novel solution to a social problem that is more effective, efficient, sustainable, or just than existing solutions and for which the value created accrues primarily to society as a whole rather than private individuals [29]. We define social innovations as new ideas (products, services and models) that simultaneously meet social needs and create new relationships or collaborations. In other words, they are innovations that are both good for society and enhance society's capacity to act [27]. Social innovation can be defined as the development and implementation of new ideas (products, services and models) to meet social needs and create new social relationships or collaborations. It represents new responses to pressing social demand, which affect the process of social interactions. It is aimed at improving human well-being. Social innovations are innovations that are social in both their ends and their means. They are innovations that are not only good for society but also enhance individuals' capacity to act [21].

These are dense definitions, that should be analyzed word for word, but here it is important that we underline the connection between the concept of social innovation and the stakeholder ecosystem [17]. The social dimension of innovation engages local systems, close-knit territorial networks full of tacit, atypical knowledge and hence of particular relevance. From this point of view every economic, institutional and social actor is able to innovate: the crucial element is that we have to recognize the role of hybridization and the meeting of diverse realities and organizational culture. On the contrary, the incapability to innovate is tied to an ineffectiveness to adopt different perspective when analyzing problems or to risk collectively not kindly recognized by the majority. If we consider the social innovation as one of the pillars of our theoretical approach, we have to underline that innovation starts when there is a social request that a social need is met. The first step is the recognition of an unmet need, and then the necessity to find one or more solutions. Social innovation approach leads us to a emphatic model: institutions and companies can no longer study the user's profile, but they have to enter in the user's world, sharing opinions and narrations, sharing, we can say, all the process of innovation.

3 Helix of Innovation

We can also say that the contemporary social construction of innovation is very comparable to Triple and Quadruple Helix approach proposed by Etzkowitz and Leydesorff [19] and by Carayannis and Campbell [6]. Starting from a "Mode 1", characterized by a "linear model of innovation", according to which university and research centres are the starting points of innovation process and the role of end users was confined to the "passive consumer one", 20 years ago we passed to a "Mode 2" [22, pp. 3–4], characterized by five principles: (1) knowledge produced in context of application; (2) transdisciplinarity; (3) heterogeneity and organizational diversity; (4) social accountability and reflexivity; (5) quality control. As Carayannis et al. [8, pp. 3–4] pointed out, this Mode paved the way to the Triple Helix model, that stressed on the importance for innovation of university-industry-government relations [19]. But at the same time stimulated the passage to the Mode 3, "that is more inclined to emphasize the coexistence and coevolution of different knowledge and innovation modes (…) accentuates pluralism and diversity of innovation modes as being necessary for advancing societies and economies" [8]. This is an important step because it stressed the importance of cross-fertilization that in any case seems to let into the circle of researchers and firms although with the help of the government. Citizens, consumers, end users, do not come into play yet, in this model, except through the guarantee that the government should ensure their interests. A substantial change occurs with Quadruple Helix model [6, p. 218, 206], that adds a fourth helix: the public, defined bay these authors as "the media based and culture based public" and "the civil society" and associated with the "creative class". In the meantime, social studies about science and innovation proposed the Social Construction of Technology [31] and the Actor Network Theory [24], to underline that innovation is social context dependent and can't be limited to the closed network university-industry, also if this circle is heterogeneous and transdisciplinary.

These theories consider not only the social character of innovation, but also the necessity for a new territorialization. Innovation need to a place-based strategy linked to territorial specificities. Governance must be responsive to a self-potential discovery: the legitimacy, however, requires the involvement of end users. Social innovation, we can say, can be possible, only if we move from triple helix to

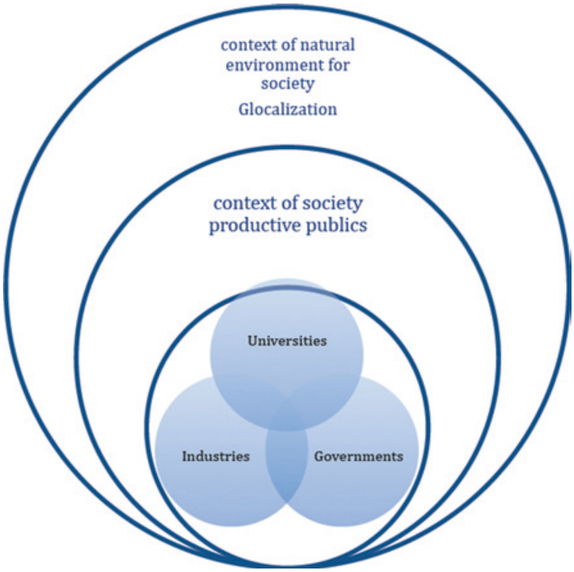


Fig. 1 Our elaboration from Carayannis et al. [8]

quadruple helix, adding “the General public” to the “classical” three actors, University, Industries and Government. More recently Carayannis [7] introduce quintuple helix, adding context of natural environments for society (Fig. 1).

4 Open Innovation

During last two decades companies have realized the progressive loss of importance about control of innovation according to close traditional model. In traditional closed innovation, a company generates, develops and commercialized its own ideas. The approach of self-reliance dominated the R&D operations of many industrial corporation for most of the 20th century. Chesbrough [10] coined the term “open innovation” a concept based on the observed fact that useful knowledge today is widely distributed, and no company, no matter how capable or how big, could innovate effectively on its own [12]. The official definition (2006) said that Open innovation is the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and expand the markets for external use of innovation, respectively. [This paradigm] assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as they look to advance their technology [11].

Open innovation, we can say with Joel West, is using the market rather internal hierarchies to source and commercialize innovation. Firms start with capturing ideas from a larger group (and often from web community), using the new forms of collaborating with external actors, creating the environment and the trust, then managing ideas and interactions and turns ideas into innovation. In open innovation model there are also institutional ways to support the role of the network (business services, company, government), as well as bottom up channels (company, individuals, clients), and together they create an interconnected system. R&D, production, financing, creation, business incubators, marketing, consumption, enduser platform, services providers and customer care become the gears of a complex mechanism.

If we considered public policy it is clear that many measures have their roots in the closed innovation era. They shoot from a logic focused on developing large national or regional markets, defending local firms, restricting foreign workers and students, and subsiding large local firms to keep them innovating. If we consider open innovation approach policy have to change into a strong support to knowledge diffusion: government have to facilitate mobility of workforce and the educational system must systematically create highly qualified labor and new intellectual property norms.

5 Changing Actors in a Changing System

In order to explain the relationship between social and open innovation and the Helix system, we need to focus directly on the dynamic relationships that underlie contemporary innovation systems. Not enough of the traditional dyadic relationships, impromptu and extemporized, between an individual researcher and an individual entrepreneur, or even formal agreements university-government, government-enterprises, universities and enterprises. Today the paradigm of innovation demands joint learning between the three actors in the chain: the activism of a pivot organization is important, but the structural nature of innovation processes assumed synergies and strategic shifts, changes and adjustments for each other. More than specific knowledge, distinguishing the individual actors, it is central the potential collective and place-based knowledge.

In that sense it becomes very important to focus on the players of the fourth helix: the end users, public production, smart cities consumers-citizens who actively participate on the innovation process.

Until a few decades ago, the world of production was describable by actors and roles defined. Economic sociology had its certainties, its patterns. A complex frame inhabited by recognizable subjects: the entrepreneurs, the workers, the employees, the managers, the supervisors, etc. The relationship between public and private was complex, but with recognizable and often governed boundaries.

Economic production was always been a private matter, in private place, often fenced, sometimes secret, mostly closed (cf. [3]).

In 1990s, but its possible to recognize also earlier warning signs, technological innovations and diversification of capitalisms, led to radical changes. From the birth of the Web, in particular, it was possible to put into practice, many desires of the hacker culture, as well as theorization of Prosumer Movement according to “The third wave” of Alvin Toffler.

The collaborative production was possible to a few, but become a reality. In last decade this process was stepped up and economical, and sociological literature coined terms as “co-production” and “co-creation” [23], “the public productive” [3], “societing” [4], “wikinomics” [35], etc.

On its turn, the *digitalization or the democratization of production* not only allowed automatization of existing manufacturing techniques but also brought in life new manufacturing processes such as the additive manufacturing process, well known as 3d printing, F/OSS systems, Wiki platforms and so on. In his seminal 2006 work ‘The Wealth of Networks’, Yochai Benkler presents a new era in the production of information, the ‘networked information economy’, facilitating action by decentralized individual users, and in particular ‘commons-based peer production’ initiatives which provided a feasible, nonproprietary alternative to information production by corporate (or State) entities. Important and revolutionary features of this new kind of production compared to previous forms were the non-hierarchical decentralized organization of the initiatives, their ‘non-market’ nature i.e. the fact that production took place altruistically and communally without remuneration or proprietary rights for participants and the fact that the information produced could be disseminated worldwide for very little cost.

There is no doubt that some of these systems are contributing to the development of the sharing economy or even of the gift economy, (we talked about it in [2]) but it is necessary to avoid falling in naive optimism. Recent history of New Economy and ICT Giants as Amazon, Apple and Google show how crowd collaboration can be exploited to make profits.

It is useful to recall here that the possibilities of web platforms, and in particular the activation of collaborative processes of participation and co production, are the result of changes, which occurred in the post Fordism. Commenting on the literature on the end of the standard enterprise and mass production, Gary B. Herrigel in 2000 argued that the various forms of vertical disintegration, the flexible specialization, the production of diversified quality, systematic rationalization, were waiting for a new model of practice. Actually it was not a unique model, but rather a set of places and platforms habitable in the Web.

The fragmentation of mass markets had taken place for years. The so-called Original Equipment Manufacturers (OEM) began to realize more and more customized goods [1] to meet in increasingly sophisticated consumers [9], but the personalized production need new technologies. Global markets open up new opportunities, but at the same time, new risks and difficulties to understand the needs of customers. OEMs are under increasing pressure to a strong outsourcing not only to control costs, but also to a request for specialization which fail to support.[1]

The post-Fordism had shown that it was increasingly difficult and expensive to predict and anticipate consumer tastes.

Consumer trends and modalities of the individualistic consumer stressed enterprises facing a situation of increasing complexity: not only is there more of a

[1] There is a vast literature on this phenomenon: see for example Sako and Helper [33].

consumer-typology to which standardize the production, but the same segmentation appears more difficult. As a result, the similarities between consumers being increasingly temporary and not affecting the totality of the individual's behavior but just specific consumption activities: firms therefore, instead of focusing on the product or consumers are increasingly brought to prioritize the criteria that guide, from time to time, consumers' choices.

Consumers are becoming more eclectic, they make contradictory choices, they move away from old hierarchical prestige symbols by goods acquisition. At the same time they are enchanting [32] by an hypertrophic development of the possibility to choose so many goods on the market, which increases both the variety and the renewal rate and also the growth of communicative and expressive factors of the products. Markets, in that sense, are no longer defined by a set of products that perform the same use function, but by everything that can compete symbolically to satisfy its intangible needs.

Firms are enchanted as their consumers; they are disoriented in a cloud of outsourcing and in a sea of no longer understood consumers. In this climate of uncertainty, companies have started to change their strategies. They began to open up, to look for new ways to interact with suppliers, consumers, consultants, other firms, etc. The new production framework has profoundly converted industrial relations up to risemantization of the notion of competition.

The meaning of competition—from the Latin *cum-pete*—do not in fact refer to a kind of natural selection that rewards the strongest at the expense of others, but rather to the ability to converge towards a common goal while moving from different starting points. If this is the goal that drives cross-cutting subjects of the production of social value through business initiative, the principle of co-ordination that is shown most effective is that of cooperation.

According to Richard Sennett [34] cooperate is very different from simple collaboration: in the first case, in addition to the objectives, it has to be shared also the means and the goals of the action. Networks of relationships increase their importance not only for connection among people and organizations, but through a variety of methods and forms of regulation. Relational systems are not only the output, the result of initiatives that aim to increase the level of coordination, but the input to create complex and more effective systems, in order to generate social value. A value that, to be true, it needs to be shared and so you need to rest on a network able to give voice to the needs and attract resources and availability of a wide area.

A parallel process intervened in the domain of public participation to public choices. For a long period the representative democracy and its decision making process was based on three pillars: the public institution, to which people delegated decisional power by means of the electoral competition; the experts (or the technicians-bureaucrats), whose power was due to scientific or organizational knowledge and, finally, the representatives of the main social economic people's interests, i.e. social or economic organizations (e.g. Trade Unions). For a long time, these three kind of actors represented citizens enough to discourage the direct commitment in decision making processes, according also to the *free rider* Olson's model. But in the last decades, and particularly thanks to the mobilizing power of ICT, people asked for a direct participation to public choices, especially to the micro or meso level ones. The evolution of ICT interacts with the raising of an "adulthood of citizens" [13] and allows forms of direct and real time involvement of citizens that both integrate or (try to) substitute the "traditional" forms of democratic decision making.

An Italian scholar [25], points out that the social production of knowledge is maximized when you add to the enhancement of knowledge (tacit or explicit) the creation OD organizational structures that allow relationships and cooperation among social actors.

There notations paving the way to further developments, detected by contemporary scholars, that we will find in Living Labs in the next section. The connection between processes of development of individual skills (micro level), and the creation of meso structures where skills can grow with continuity and stability lead to the territorialization of the triple or Quadruple helix. In Smart Cities this is a key phenomenon: what are optimum conditions for innovation? What are the essential networks and nodes for Smart Cities?

Socialization and cultural guide to innovation become crucial: a highly socialized innovation means an innovation perceived and experienced as a collective target priority and a vibrant part of the society (organized actors and individuals). The socialization process, in that sense, transforms innovations, from technical issue into a widespread social action object, while the political and cultural leadership refers to all regulations, policies and initiatives of different institutional levels (International, European, national and local) in order to the same, clear and sharing elements that come into play in the innovation process [25].

For some author [15], the key elements of orchestration are quality of Research, socialization of innovation and governance, while others [14] say that are knowledge mobility, innovation appropriability and network stability. Relations among the player are characterized, necessarily, by *coopetition*, combining competition and cooperation in the value net [28] which is represented as a diamond shape, with four defined player designations at the corners: customers, suppliers, competitors and *complementors*. E.G. Carayannis insists on *coopetition:* already in 1999, in an article written with Jeffrey Alexander, put the attention to the relationships linking the firm to its environment at the market, political and ecosystems levels. The introduction of the ecosystem level, that Carayannis proposes in a number of works, paves the way to the direct intervention of end user in innovation processes, that is one of the key features of Living Labs.

6 The Living Labs

A Living Lab, according to a EC [20, p. 7] document, is "a user driven open innovation ecosystem based on business-citizens-government partnership which enables users to take active part in the research, development and innovation process" or "a user driven, open innovation environment in real-life settings in which users test and experiment new products or services, in a framework integrating companies, people, research and innovation actors and public sector (the so called Public-Private-People Partnership, PPPP)." Recently also Wikipedia proposes a similar definition: "A living lab is a user-centred, open-innovation ecosystem, often operating in a territorial context (e.g. city, agglomeration, region), integrating concurrent research and innovation processes a public-private-people partnership. The concept is based on a systematic user co-creation approach integrating research and innovation processes. These are integrated through the co-creation, exploration, experimentation and evaluation of innovative ideas, scenarios, concepts and related technological artifacts in real life use cases. Such use cases involve user communities, not only as observed subjects but also as a source of creation."

The concept of Living Labs was born in Boston, where professor William Mitchell was used to observe the living patterns of users in smart homes. The idea was to involving city dwellers more actively in urban planning and city design [26], but suddenly Living Lab was traduced in Europe in wider use to "Enhance innovation, inclusion, usefulness and usability of ICT and its application in society" [18].

The main and more innovative dimensions of a LL are:

(a) The first one is for sure the involvement of end users at the early stages of innovation process. This involvement, however, has different motivations which co-exist in different types of LL submitting different logic. These motivations can be arranged along a continuum, which has at one end a "corporate oriented", in which the early involvement of end users ensures a better compliance of the products to the need of the consumer, reducing time from conception to commercialization: this means a better competitiveness of enterprises. At the other extreme the "need oriented", that is the attempt of putting before the need and the problems instead of solutions and products.
In this case the role of end users is not only limited to an active part in a process driven by firms (or research institutions), but is a guiding role, allowed (we think) by the public governance (and sometimes by public funding) of the Living Lab. Is matter of fact that the first experiences of LLs derives from "enlighten" firms or research centers, that opened their doors to end users, but during their evolution LLs recognized to end users a role of growing importance. Some Authors underline that the methodology and the methods used to build and to conduct the Living Lab play an important role in its future and that the concept design phase is crucial for its success [5, p. 1]. The crucial role of end user is obtained if "we can shift the perspective from problems to opportunities and from requirements to needs" (Ibid.); in this way, from the point of view of users, we have the best insurance that their needs will be put on the center (and on the beginning) of the innovation process and, from the point of view of firms, they'll be sure about the success of the products that will derive from the LL. So empowerment of citizens walk hand in hand with competitiveness of economic sector and an "user driven innovation" will really took place.

(b) Open and social innovation. Living Labs has the function of open innovation intermediaries that aims to provide structure and governance to user involvement. In this sense, Living Lab is home of user contribution, identifying and codifying tacit and practice based knowledge and diffusing into ad hoc innovation network. Living Lab is also a place of social innovation, because it is a real life environment, where is possible to generate new socially negotiated meanings for products and services. If we consider the methodological point of view, the "social" aspect of the innovation process derives not only by the end user's involvement, but also from the "social" character of the process by which a Living Lab works: real or virtual meetings, direct involvement of end users in ideation phase, use of methodologies that can maximize the participation and the interaction (also with other actors, not only end users, but also public, experts, researchers and social representatives). From the substantial point of view, first of all innovative can be the process or the product and the "social" aspect derives from the shared benefits in a bigger community; this is also linked to the main area in which LLs usually works: although quite every matter can be the subject of a LL, a great part of its refers to [20] e-Wellbeing, e-Services in Rural or Developing Areas, e-Democracy and e-Governance, ICT for Energy Efficiency. Also in the Ligurian case of Alcotra the concerned sectors are closely linked to primary needs of citizens (health, energy, mobility). Social innovation is in our opinion closely linked also to the principle of co-creation, that means that all stakeholders must cooperate to the final outcome of the Living Lab and that cooperative way of work are a key feature of a Living Lab.

(c) The (public) governance of the Living Lab. This aspect is not usually quoted, because Living labs can also arise "from the scratch", or thanks to the solicitation of firms or research centers (less probably, by end users' associations). Anyway, we think that a minimum set of rules warrantied by a public body are necessary to give to the participants the starting trust to share knowledge (and to devote time) with other people and to commit in a common effort for common objectives. There is no doubt that in contemporary society there is an increasingly availability to cooperate and collaborate, in particular through new technologies. Phenomenon as Wikipedia, but also open source communities, crowdfunding, peer to peer networks and so on demonstrate [2] that people tend to participate to imagined community sharing knowledge and intelligence: a new and old way to exchange and build relationships. In case of Living Lab the aim is not giving economy but competitiveness: it is very likely and desirable that LL products provide profits to firms and development for territories. The presence of local public organizations should ensure that LL innovation stimulates both companies and research system putting at the center social needs. Local government support the idea that "needs are opportunities waiting to be exploited" [5: 3] albeit in a logic of competitiveness. Public institution, as we see in Alcotra Innovation experience, is an irreplaceable actor because of ensuring costs and organizations of the startup phase of LL with a methodological and monitoring and evaluation system that is functional to the inclusion of all primary and secondary stakeholders. For a good governance is required an ICT infrastructure, that allows shared participation, immediate feedback, direct democracy in the governance. ICT infrastructure is closely linked to the governance of the LL. Although public governance is an important requirement, a LL works well if it's spontaneous and if all stakeholders play a role in decision making process and cooperate (as suggested above) to final results. So LL must be democratic and participative, not only to be coherent with its philosophy, but also to give room to all competencies and availabilities.

(d) A real life setting and the goal to produce new goods or services, or to improve in an innovative way actual good or services of public interest. This means that a LL is not an arena in which people only debates new ideas, but an ecosystem in which innovation take place and produce something of new and useful for people, firms and the involved communities. Of course in some case the real life experimentation will be most important than in other cases, in which crowdsourcing of ideas will be privileged, but in any case something of new and useful must be the outcome of LLs. This must be also an important part of evaluation, that can't be limited to the process, but must include results and impacts. The way in which profit oriented actors and socially oriented actors can cooperate is linked also to the way in which each of them can have a gain, because a LL must be a win–win game to be seriously played.

7 Alcotra Innovation Living Lab

The Alcotra Innovation strategic project, funded by the Alcotra Italy–France 2007–2013 territorial cross-border cooperation program, had as partners Rhône-Alpes and Provence-Alpes-Cote d'Azur Regions, in France, and those of Piedmont (acting as Coordinator), Liguria and Aosta Valley, in Italy, as well as the Province of Turin. The project, launched in September 2010 and lasting for 3 years, aimed at experimentally introducing the Living Lab approach into the respective innovation policies and practices, according to a transnational perspective, namely through the building up and operation of cross-border Living Labs in the five participant regions. It was therefore quite natural for the Alcotra Innovation partners during the project design phase (in the year 2009), to be attracted by the potential contribution of the Living Lab approach to existing, and upcoming, regional innovation policies and practices. There was already evidence in that sense in the three Regions: Piedmont—being a member of the ENoLL (European Network of Living Labs) since 2008—PACA—with the success story of "PACALabs", one of the earliest examples of user driven and territorially oriented innovation policy promoted by the public hand—and Rhône-Alpes with 7 Living Labs (most of them created in 2009) is particularly active in the domain of media, design and uses innovation. However, a new perception was emerging that the full potential of Living Labs for innovation policy should be grasped in the broader framework of the Alps-Mediterranean EuroRegion—including Liguria, Piedmont, Provence-Alpes-Côte d'Azur, Rhône-Alpes and Aosta Valley—rather than at single regional level. Therefore, the cross-border dimension was added to the picture.

Four thematic domains were selected for the purpose of Living Lab experimentation, namely:

- Intelligent Mobility, coordinated by Piedmont Region and Liguria Region;
- Smart Energies, alternative sources of power and energy efficiency, coordinated by Aosta Valley Region;
- e-Health, coordinated by PACA Region;
- Creative Industries, coordinated by Rhône-Alpes Region.

In a first phase of the project, each Region organized local workshops with the purpose of raising awareness of the Quadruple Helix stakeholders on the Alcotra Innovation objectives, the cross-border Living Lab's idea and its possible advantages compared to other approaches. With the main exception of PACA Region, where the PACALabs Initiative had been in place since 2008, most regional stakeholders did not know much about user driven open innovation and therefore had to learn about previous successful experiences. In a second phase, having formed the cross-border working groups, which were animated and facilitated by both thematic and methodology experts, participants started to become familiar with the concept and to think about the design of possible pilot actions involving the Living Labs' operational principles in a meaningful and useful way.

The experimentation of Intelligent Mobility was characterized by several starting meetings with interactive methodologies in online and offline contest. The aim of the groups was to develop innovative solutions for tourists and open air travellers and the output of the laboratory was the prototype of an application for mobile device with two different interface. The Living Lab Creative Industries has used a mixed user centered methodology: the aim was create and test innovative solution in museum fruition. After several cross border workshops participants decided to develop two experiments, one in Rhone-Alpes and one in Piedmont, during which artists, software developers, designers, contractors, visitors and museum curators, worked together to prototype interactive museum design. Participants were immersed in the context of the museum: the laboratory where ideas were born, tested, changed, imagined and co created in augmented reality.

8 Conclusions

The experience of cross-border Living Lab Alcotra Innovation show both the potential of this model of open innovation and the need/opportunity to adjust it into local context or in specific issue.

In none of Alcotra Living Lab there is a leadership role of companies, important and articulated are the role of end users, while public institutions were protagonist. Public institutions has in fact some features that, especially in a cross border dimension, can hardly encompassed:

(a) they establish the basic rules of the game, and solve the functions of trust intermediaries;
(b) they tied most important social need with priority for action planning or regional and transnational programming;
(c) cross border dimension show opportunity and difficulties of coopetition among territories. In particular in touristic field, cross border show how territories are complementary but at the same time, they are in a strong competition. Public institutions, in that case, have to select policies to maximize synergies and minimize replacement effects;
(d) they provide some basic services (animation, sharing platform, administration and payment of pocket costs as travel, hospitality etc.),
(e) they are able to steer, especially in the field of public services, both the demand and the offer. This it was evident in e-health, but in tourism and info mobility too.

It was found in all experiment the role of public institution especially at the end of the project, when the need to results perpetuation, found companies and end user associations unable to build a business plan containing living lab costs. The weak point might be considered that an excess of public intervention could alleviate overly the entrepreneurial component of Living Lab. In a Schumpeterian way we can say that innovation arises out of new combinations of existing capabilities and openness is crucial: the famous NIH (Not Invented Here) syndrome is always around the corner, especially in public institutions. Smart Cities need a profound change: from NIH to TFE (Thankfully Founded Elsewhere): it means not reinvent hot water again, but to “use” what has already been invented elsewhere, restarting from there to some new frontiers.

References

1. Addis, M., & Holbrook, M. B. (2001). On the conceptual link between mass customisation and experiential consumption: an explosion of subjectivity. *Journal of Consumer Behaviour, 1*(1), 50–66.
2. Aime, M., & Cossetta, A. (2010). *Il dono al tempo di Internet*. Torino: Einaudi.
3. Arvidsson, A. (2005). Brands a critical perspective. *Journal of Consumer Culture, 5*(2), 235–258.
4. Arvidsson, A., & Giordano, A. (2013). *Societing reloaded*. Milano: Egea.
5. Bergvall-Kåreborn, B., Holst, M., & Ståhlbröst, A. (2009). Concept design with a living lab approach. In Proceedings of the 42nd Hawaii international conference on system sciences. http://originwww.computer.org/portal/web/search/advanced?p_p_id=searchadvanced_WAR_pluginssearch_INSTANCE_eO7R&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1.
6. Carayannis, E. G., & Campbell, D. F. J. (2009). ‘Mode 3’and’Quadruple Helix’: toward a 21st century fractal innovation ecosystem. *International Journal of Technology Management, 46*(3), 201–234.

7. Carayannis, E. G., & Campbell, D. F. J. (2011). Open innovation diplomacy and a 21st century fractal research, education and innovation (FREIE) ecosystem: building on the Quadruple and Quintuple Helix innovation concepts and the Mode 3 knowledge production system. *Journal of the Knowledge Economy, 2*(3), 327–372.
8. Carayannis, E. G., Barth, T. D., & Campbell, D. F. J. (2012). The Quintuple Helix innovation model: global warming as a challenge and driver for innovation. *Journal of Innovation and Entrepreneurship, 1*(2), 1–12.
9. Carù, A., & Cova, B. (Eds.). (2007). *Consuming experience*. London: Routledge.
10. Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Boston: Harvard Business Press.
11. Chesbrough, H., Vanhaverbeke, W., & West, J. (Eds.). (2006). *Open innovation: Researching a new paradigm*. Oxford university press.
12. Chesbrough, H. (2011). Open services innovation. Rethinking your business to growth.
13. Crozier, M. (1987). *Etat modeste, état moderne, stratégie pour un autre changement*. Paris: Fayard.
14. Dhanaraj, C., & Parkhe, A. (2006). Orchestrating innovation networks. *Academy of Management Review, 31*(3), 659–669.
15. D'Andrea, L. (2006). L'innovazione come processo sociale. Conoscenza & Innovazione. http://conoscenzaeinnovazione.org.
16. Dosi, G. (1988). Sources, procedures, and microeconomic effects of innovation. *Journal of economic literature,* 1120−1171.
17. Fung, A., & Wright, E. O. (2001). Deepening democracy: Innovations in empowered participatory governance. *Politics and Society, 29*(1), 5–42.
18. Eriksson, M., Niitamo, V. P., & Kulkki, S. (2005). *State-of-the-art in utilizing Living Labs approach to user-centric ICT innovation-a European approach*. Lulea: Center for Distance-spanning Technology. Lulea University of Technology Sweden: Lulea. Online under: http://www.cdt.ltu.se/main.php/SOA_LivingLabs.pdf.
19. Etkowitz, H., & Leyedesdorf, f L. (2000). The dynamics of innovation; from National Systems and Mode 2 to a Triple Elix of university-industry-government relations. *Research Policy, 29*, 109–123.
20. European Commission (2009). *Living Labs for user-driven open innovation*. Directorate General for Information Society and Media, Bruxelles.
21. European Commission (2013). *Guide to social innovation*, DG Regional and Urban policy and DG Employment, Social Affairs and Inclusion.
22. Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). The new production of knowledge: The dynamics of science and research in contemporary societies. Sage.
23. Humphreys, A., & Grayson, K. (2008). The intersecting roles of consumer and producer: A critical perspective on co-production, co-creation and prosumption. *Sociology Compass, 2*(3), 963–980.
24. Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, Mass: Cambridge University Press.
25. Martini, E. (2011). *Socializzare per innovare*. Loffredo Napoli: Il modello della Tripla Elica.
26. Mitchell, W. J. (2005). Constructing complexity. In *Computer aided architectural design futures 2005* (pp. 41–50). Netherlands: Springer.
27. Murray, R., Caulier-Grice, J., & Mulgan, G. (2010). *The open book of social innovation*. National Endowment for Science, Technology and the Art.
28. Nalebuff, B. J., & Brandenburger, A. (1996). *Co-opetition*. London: HarperCollinsBusiness.
29. Phills, J. A., Deiglmeier, K., & Miller, D. T. (2008). Rediscovering social innovation. *Stanford Social Innovation Review, 6*(4), 34–43.
30. Paskaleva K., E-governance ad an enabler of the smart city, in Deakin, M. (2013) *Smart cities: Governing, modelling and analysing the transition*.
31. Pinch, T. J., & Bijker, W. E. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science, 14*, 388–441.

32. Ritzer, G., & Jurgenson, N. (2010). Production, consumption, prosumption the nature of capitalism in the age of the digital prosumer. *Journal of Consumer Culture, 10*(1), 13–36.
33. Sako, M., & Helper, S. (1998). Determinants of trust in supplier relations: Evidence from the automotive industry in Japan and the United States. *Journal of Economic Behavior & Organization, 34*(3), 387–417.
34. Sennett, R. (2012). Together: the rituals, pleasures and politics of cooperation. Yale University Press.
35. Tapscott, D., & Williams, A. D. (2008). Wikinomics: How mass collaboration changes everything. Penguin.com.

Index

R. P. Dameri and C. Rosenthal-Sabroux (eds.), *Smart City*, Progress in IS,
DOI: 10.1007/978-3-319-06160-3, © Springer International Publishing Switzerland 2014

Eur. Phys. J. Special Topics **214**, 481–518 (2012)

DOI: 10.1140/epjst/e2012-01703-3

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Regular Article

Smart cities of the future

M. Batty[1,a], K.W. Axhausen[2], F. Giannotti[3], A. Pozdnoukhov[4], A. Bazzani[5], M. Wachowicz[6], G. Ouzounis[7], and Y. Portugali[8]

[1] Centre for Advanced Spatial Analysis (CASA), University College London, 90, Tottenham Court Road, London W1N 6TR, UK
[2] IVT, ETH, Wolfgang-Pauli-Strasse 15, HIL F31.1, 8093 Zrich, Switzerland
[3] KDD Lab, Istituto ISTI, Area della Ricerca CNR di Pisa, Universit di Pisa, via G., Moruzzi 1, 56124 Pisa, Italy
[4] National Centre for Geocomputation, Iontas Building, NUI Maynooth, Co., Kildare, Ireland
[5] Department of Physics and National Institute of Nuclear Physics, University of Bologna, via Irnerio 46, 40126 Bologna, Italy
[6] University of New Brunswick, Department of Geodesy and Geomatics Engineering, 15 Dineen Drive, PO Box 4400, Fredericton, NB, Canada, E3B 5A
[7] Geo-Spatial Information Analysis for Global Security and Stability, Institute for the Security and Protection of the Citizen (IPSC), Joint Research Centre, European Commission, T.P., 267, via E. Fermi 2749, 21027 Ispra, Italy
[8] Tel-Aviv University, Department of Geography and the Human Environment, Tel Aviv 69978, Israel

Received in final form 9 October 2012
Published online 5 December 2012

Abstract. Here we sketch the rudiments of what constitutes a smart city which we define as a city in which ICT is merged with traditional infrastructures, coordinated and integrated using new digital technologies. We first sketch our vision defining seven goals which concern: developing a new understanding of urban problems; effective and feasible ways to coordinate urban technologies; models and methods for using urban data across spatial and temporal scales; developing new technologies for communication and dissemination; developing new forms of urban governance and organisation; defining critical problems relating to cities, transport, and energy; and identifying risk, uncertainty, and hazards in the smart city. To this, we add six research challenges: to relate the infrastructure of smart cities to their operational functioning and planning through management, control and optimisation; to explore the notion of the city as a laboratory for innovation; to provide portfolios of urban simulation which inform future designs; to develop technologies that ensure equity, fairness and realise a better quality of city life; to develop technologies that ensure informed participation and create shared knowledge for democratic city governance; and to ensure greater and more effective mobility and access to opportunities for

[a] e-mail: m.batty@ucl.ac.uk

urban populations. We begin by defining the state of the art, explaining the science of smart cities. We define six scenarios based on new cities badging themselves as smart, older cities regenerating themselves as smart, the development of science parks, tech cities, and technopoles focused on high technologies, the development of urban services using contemporary ICT, the use of ICT to develop new urban intelligence functions, and the development of online and mobile forms of participation. Seven project areas are then proposed: Integrated Databases for the Smart City, Sensing, Networking and the Impact of New Social Media, Modelling Network Performance, Mobility and Travel Behaviour, Modelling Urban Land Use, Transport and Economic Interactions, Modelling Urban Transactional Activities in Labour and Housing Markets, Decision Support as Urban Intelligence, Participatory Governance and Planning Structures for the Smart City. Finally we anticipate the paradigm shifts that will occur in this research and define a series of key demonstrators which we believe are important to progressing a science of smart cities.

1 Our visionary approach

For much of the 20th century, the idea that a city could be smart was a science fiction that was pictured in the popular media but quite suddenly with the massive proliferation of computable devices across many scales and with a modicum of intelligence being embedded into such devices, the prospect that a city might become smart, sentient even, is fast becoming the new reality. The convergence of information and communication technologies is producing urban environments that are quite different from anything that we have experienced hitherto. Cities are becoming smart not only in terms of the way we can automate routine functions serving individual persons, buildings, traffic systems but in ways that enable us to monitor, understand, analyse and plan the city to improve the efficiency, equity and quality of life for its citizens in real time. This is changing the way we are able to plan across multiple time scales, raising the prospect that cities can be made smarter in the long term by continuous reflection in the short term.

Smart cities are often pictured as constellations of instruments across many scales that are connected through multiple networks which provide continuous data regarding the movements of people and materials in terms of the flow of decisions about the physical and social form of the city. Cities however can only be smart if there are intelligence functions that are able to integrate and synthesise this data to some purpose, ways of improving the efficiency, equity, sustainability and quality of life in cities. In *FuturICT*, we will research smart cities not simply in terms of their instrumentation which is the domain of both large and small ICT companies who are providing the detailed hardware and software to provide what some have called the operating system for the smart city, but in terms of the way this instrumentation is opening up dramatically different forms of social organisation.

We will focus directly on ways in which this infrastructure can be integrated, how the data that are being collected can be mined, how services delivered by traditional means can be organised and delivered much more efficiently using these new technologies, all part of the idea of the Planetary Nervous System that is central to our proposal. This is our first goal but in parallel and embedded within this, we are interested in standing back from the nuts and bolts of the smart city, and devising much more effective models and simulations that will address problems of efficiency, equity and quality of life, set within a new context where a much wider group of citizens can engage in the science of smart cities through new ways of participating in the future design of their cities and neighbourhoods. These embrace our ideas of Living Earth

Simulators and the Participatory Platforms that are key to the *FuturICT* approach. In short, the smart city will be the boost to new forms of policy analysis and planning in the information age, and the greatest impacts of new technologies will be on the way we organise ourselves in cities and the way we plan this organisation. The main goal of *FuturICT* is to provide intelligence functions that will make this possible in the most effective and equitable ways.

The concept of the smart city emerged during the last decade as a fusion of ideas about how information and communications technologies might improve the functioning of cities, enhancing their efficiency, improving their competitiveness, and providing new ways in which problems of poverty, social deprivation, and poor environment might be addressed [26]. The essence of the idea revolves around the need to coordinate and integrate technologies that have hitherto been developed separately from one another but have clear synergies in their operation and need to be coupled so that many new opportunities which will improve the quality of life can be realized. The term smart city in fact has many faces [40]. Intelligent cities, virtual cities, digital cities, information cities are all perspectives on the idea that ICT is central to the operation of the future city [1]. Our research will embrace this challenge in the belief that coupling, coordination and integration are required so that future and emerging technologies can best be exploited in the interests of the community at large. An essential strand in our approach is to use ICT to engage the community through diverse instruments and initiatives that build upon online engagement in solving the key problems of cities, using the kinds of computer-based tools, techniques, methods and organisational structures that we will research here. To focus our research, we define seven goals.

1.1 Goals of research

A New Understanding of Urban Problems. Cities are complex systems par excellence, more than the sum of their parts and developed through a multitude of individual and collective decisions from the bottom up to the top down. The complexity sciences are integral to their understanding which is a moving target in that cities themselves are becoming more complex through the very technologies that we are using to understand them. We will not only fashion a programme for Europe to grow our understanding as a prelude to action and decision, but embed this as part of a wider international effort.

Effective and Feasible Ways to Coordinate Urban Technologies. Rapid advances in building information technologies into the very fabric of the city while at the same time using these technologies to integrate and add value to the provision of urban services provide the mandate for the sustained development of new methods. This will involve integrating data, software and organisational forms that best improve the efficiency and competitiveness of the environment in which cities operate.

Models and Methods for Using Urban Data across Spatial and Temporal Scales. Much data which is being generated in real time in cities needs to be merged with more traditional cross sectional sources but built on simulations that link real time, more routine problems to longer term strategic planning and action. Multilevel integrated modelling is thus key to this effort.

Developing New Technologies for Communication and Dissemination. New sources of urban data, the articulation of urban problems, plans and policies, and all the apparatus used in engaging the community in developing smart cities require new forms of online participation making use of the latest ICT in terms of distributed computation and state of the art human computer interaction (HCI).

New Forms of Urban Governance and Organisation. New ways of re-engineering cities to make them smart, responsive, competitive and equitable will require new forms of governance for an online world. Issues of privacy and access are key to this vision.

Defining Critical Problems Relating to Cities, Transport, and Energy. FuturICT is focussed on defining critical problems that emerge rapidly and unexpectedly in human society, some of which reveal critical infrastructures. The analysis of such problems and their identification is crucial to the sustainability and resilience of smart cities. Models and simulations associated with rapid changes in cities from housing market bubbles to regeneration to ethnic segregation will be explored using new approaches. The idea that cities are far-from-equilibrium, dominated by fast and slow dynamics in short and long cycles is central to our approach.

Risk, Uncertainty and Hazard in the Smart City. A much more informed understanding of risks in urban society is required which involves new data, new technologies, and new collective approaches to decision-making. Our notion of cities as strongly coupled systems that generate unexpected and surprising dynamics needs to be understood and the introduction of new technologies into cities is changing the nature of this dynamics, not necessarily for the better. Our quest in one sense is to develop technologies that will outsmart the smart city, anticipating this dynamics. Future and emerging technologies will be key to developing such approaches.

1.2 Research opportunities

We stand at a threshold in beginning to make sense of new information and communication technologies that will be deeply interwoven with conventional material technologies within the next decade. Already 22 percent of all people in the UK have smart phones in which they are able to access online services and this is growing at around 38 percent each year. At this rate, there will be total penetration by smart phones of the market (see `http://en.wikipedia.org/wiki/Smartphone/`) for mobile devices by 2015. Other estimates suggest that by 2015, there will be more than 2 billion smart phone users globally which suggests that access to online services will be the dominant mode of accessing information by the end of the *FuturICT* project in 2022.

The opportunities for the development of smart cities using mobile and other platforms at every spatial scale and over very different time spans will be enormous but the real challenge is to put in place new technologies that will integrate individualised and local technologies that are fast proliferating. For there to be real synergy, the idea of the smart city with all its economic and social benefits will only become a reality if such coordination is specifically addressed. This will involve a blend of sciences and arts which *FuturICT* is extremely well placed to initiate. It will involve a clear synthesis of hardware, software, database, and organisational technologies that are able to relate to the key problems of society and will require entirely new methods and models for synthesising diverse data and ideas that are currently not being addressed. This in turn implies a massive paradigm shift in how we address social and economic problems using ICT. The smart city will be at the forefront of this revolution.

1.3 Research challenges

The major intellectual challenge that we and the rest of society face, is embracing the idea that as we develop new digital technologies, we use those same technologies to study the processes of their application, implementation and impact on society.

We shape our tools and thereafter our tools shape us said Marshall McLuhan in 1964 his seminal book [32], and this is the challenge that we need to resolve in developing truly smart cities that will benefit the quality of life of all our citizens. In this, it is likely that participation in formulating policies might be very different from the past when futures were dictated by the elite, primarily because of its access to information. Already it is clear that a citizenry which is informed through the power of the net is beginning to make a difference as new forms of data and advice are being implemented using crowd-sourcing. New forms of preference elicitation are being generated using mobile and other applications, while the economy is essentially moving online with the disappearance of material tokens (cash). These are profound changes that we need to mobilize using the equally powerful science that *FuturICT* will unleash.

To this end, we can identify five key scientific challenges that relate to the seven goals outlined above.

To Relate the Infrastructure of Smart Cities to their Operational Functioning and Planning Through Management, Control and Optimisation. ICT is being fast embedded into the very fabric of the city in terms of its materials and infrastructure while wireless solutions are proliferating in ways that are hard to understand. At the same time, developments in computation and data to inform the planning of such cities are using these same technologies. There needs to be a major effort at showing how such developments can be integrated so that cities can become truly smart in the way their planners and citizenry can use such technologies to improve the quality of life. This style of multimodality constitutes a real challenge.

To Explore the Notion of the City as a Laboratory for Innovation. ICT is being developed to increase the efficiency of energy systems, the delivery of services ranging from utilities to retailing in cities, and to improve communications and transportation. The prospect of building models of cities functioning in real time from routinely sensed data is now a clear prospect and smart cities should evolve intelligence functions in the form of laboratories – that enable their monitoring and design. The competitive edge that a city offers is crucial to such intelligence.

To Provide Portfolios of Urban Simulation which Inform Future Designs. As the real time city and its sensing gets closer to providing information about longer term changes, there will be a new immediacy in the construction of urban simulation models Aggregate models will be replaced with disaggregate and our project will explore many different kinds of models building on and extending the sciences of complexity. We consider it important to build many different models of the same situation in the belief that a pluralistic approach is central to improved understanding of this complexity.

To Develop Technologies that Ensure Equity, Fairness and Realise a Better Quality of City Life. Efficiency must be balanced with equity. New technologies have a tendency to polarise and divide at many levels and we need to explore how new forms of regulation at the level of urban and transport planning, and economic and community development can be improved using future and emerging technologies. The smart city of course offers the prospect of ending the digital divide but it will also open up different divides and our challenge is to anticipate and plan for these.

To Develop Technologies that Ensure Widespread Participation. New ICT is essentially network-based and enables extensive interactions across many domains and scales. Part of the process of coordination and integration using state of the art data systems and distributed computing must involve ways in which the citizenry is able to participate and to blend their personal knowledge with that of experts who are developing these technologies. Privacy concerns as well as security are key to this challenge.

To Ensure and Enhance Mobility for Urban Populations. New ICT is able to improve mobility on many levels thereby increasing spatial and aspatial accessibilities to jobs, leisure, social opportunities and so on, thereby enabling the citizenry to increase their levels of life satisfaction.

2 The state of the art

2.1 Key themes

2.1.1 Explanation: The science of smart cities

Our proposal is predicated on a robust and consistent definition of the smart city. The term smart is peculiarly American in that it is widely used in everyday speech to refer to ideas and people that provide clever insights but it has been adopted more recently in city planning through the cliche smart growth. Rather than letting the market dictate the way cities grow and sprawl, smart growth is a movement that implies we can achieve greater efficiencies through coordinating the forces that lead to laissez faire growth: transportation, land speculation, conservation, and economic development. We adopt here the definition that is coined by Caragliu, Del Bo, and Nijkamp (2009) [12] which is summarised in Wikipedia: a smart city is a synthesis of hard infrastructure (or physical capital) with the availability and quality of knowledge communication and social infrastructure. The latter form of capital is decisive for urban competitiveness. (`http://en.wikipedia.org/wiki/Smart_city`). Digital cities tend to focus on the hard infrastructure whereas intelligent cities on the way such infrastructure is used [3–5]. Earlier conceptions included the idea of the wired city [13] which originally came from James Martins conception of the wired society.

To this nexus, we add the notion that smart cities are also instruments for improving competitiveness in such a way that community and quality of life are enhanced. Cities that are smart only with respect to their economy are not smart at all if they disregard the social conditions of their citizenry. In fact the term smart city has become shorthand for the way companies that are developing global ICT from infrastructure such as networks to software as services large companies such as IBM, CISCO, Microsoft, Oracle, SAP are beginning to generalise their products as they see markets in cities representing the next wave of product development in the globally distributed world that now exists. In the sequel, we will sketch the main themes that are developing around the smart city idea from hardware through to services and from data through to simulation and thence prediction, participation, and design.

Many of the worlds most successful ICT companies are extending their focus in software to area-wide applications that involve developing online service delivery systems for cities and they are badging their products as part of the move to make cities smart. For example, IBM are in the vanguard of these services and although the focus is still from the perspective of routine services involving utilities and traffic in the first instance, very much in the tradition of urban operations research, their remit is expanding to deal with more strategic functions and intelligence. This is for directing and guiding the city as well as improving its long term quality of life. IBM under its Smart Planet initiatives has key centres working on Smart Cities as their site (`http://www.ibm.com/smarterplanet/us/en/smarter_cities/overview/index.html`) shows. Their business development division has produced a useful summary from which we take Fig. 1 that illustrates the range of services that this particular

Today...	What if a city could...	Already, cities are...
City services • Service delivery in silos with one size fits all	• Tailor services to the needs of individual citizens	• Using technology to integrate the information systems of different service delivery agencies to enable better services for citizens
Citizens • Cities have difficulty using all the information at their disposal • Citizens face limited access to information about their healthcare, education and housing needs.	• Reduce crime and react faster to public safety threats, by analyzing information in realtime? • Use better connections and advanced analytics to interpret vast amounts of data collected to improve health outcomes?	• Putting in place a new public safety system in Chicago, allowing realtime video surveillance and faster more effective response to emergencies • Giving doctors in Copenhagen instant access to patients' health records, achieving the highest satisfaction and lowest error rates in the world.[38]
Transport • Transporting people and goods is dogged by congestion, wasted hours and wasted fuel.	• Eliminate congestion and generate sustainable new revenues, while integrating all transport modes with each other and the wider economy?	• Bringing in a dynamically priced congestion charge for cars to enter Stockholm, reducing inner-city traffic by 25 percent and emissions by 14 percent, while boosting inner-city retail by 6 percent and generating new revenue streams.[39]
Communication • Many cities have yet to provide connectivity for citizens • "Going online" typically means at slow speeds and at a fixed location.	• Connect up all businesses, citizens and systems with universal affordable high-speed connectivity?	• Merging medical, business, residential and government data systems into a so-called ubiquitous city in Songdo, Korea, giving citizens and business a range of new services, from automated recycling to universal smartcards for paying bills and accessing medical records.
Water • Half of all water generated is wasted, while water quality is uncertain.	• Analyze entire water ecosystems, from rivers and reservoirs to the pumps and pipes in our homes? • Give individuals and businesses timely insight into their own water use, raising awareness, locating inefficiencies and decreasing unnecessary demand?	• Monitoring, managing and forecasting water-based challenges, in Galway, Ireland, through an advanced sensor network and realtime data analysis, giving all stakeholders – from scientists to commercial fishing – up-to-date information.
Business • Businesses must deal with unnecessary administrative burdens in some areas, while regulation lags behind in others.	• Impose the highest standards on business activities, while improving business efficiency?	• Boosting public sector productivity, while simplifying processes for business in Dubai through a Single Window System that simplifies and integrates delivery and procedures across a range of almost 100 public services.[40]
Energy • Insecure and unsustainable energy sources.	• Allow consumers to send price signals – and energy – back to the market, smoothing consumption and lowering usage?	• Giving households access to live energy prices and adjust their use accordingly, as in a Seattle-based trial, reducing stress on the grid by up to 15 percent and energy bills by 10 percent on average.[41]

Source: IBM Center for Economic Development analysis.

Fig. 1. IBMs typology of urban issues for the smart city, now and in the future.

company sees as part of their Smart City initiative [43]. This is not dissimilar to the sorts of problems and policies that have been defined for cities in more traditional terms over the last 50 or more years but where the focus of the solutions is now on ICT. In fact what *FuturICT* will offer is a much deeper and considered inquiry into the development of smart cities, linking these technologies to social questions, co-evolving these new technologies along with initiatives of the many other ICT companies who have similar missions to that of IBM.

2.1.2 Representation: Measuring and mining urban data

Traffic flows were the first data to be automatically sensed in cities but databases of various spatial data go back centuries. Indeed digital computing emerged as much from a concern over collecting census data as it did from a concern for scientific computation. Herman Hollerith introduced punched card technology for the 1890 US Census from which sprang the company that ultimately became IBM. Since the late 1990s, such data has been routinely collected and displayed using GIS (geographic information systems) technology, and the first visual systems to be widely available on the web were maps for navigation. This is the backcloth against which many different initiatives in collecting data from new varieties of digital access are being fashioned such as GPS in vehicles and on the person, from electronic messaging in the form of social media sites, from traces left through purchase of goods and related demand-supply situations, and from access to many kinds of web site. Satellite remote-sensing data is also now widely deployed, more local scale sensing from LIDAR is proliferating, and a variety of scanning technologies that range from the region to the person and to very fine scale tagging as in the focus associated with the internet of things, are becoming significant. One of the most extensive crowd-sourcing applications, next to Wikipedia, is Open Street Map built from a community of some 20000 active users who continually update the map using GPS. Indeed new models of scientific discovery are emerging from developments in rather focussed crowd-sourcing and these are applicable to how we might figure out good designs for efficient and equitable cities [33].

Within the next twenty years, most of the data that we will use to understand cities will come from digital sensors of our transactions and will be available in various forms, with temporal tags as well as geotags in many instances. To interpret such data, we need to exploit and extend a variety of data mining techniques through which the visualisation of correlations and patterns in such data will be essential. The open data movement is gaining momentum (e.g. `http://data.gov.uk/`) and we see *FuturICT* as sharply focussed on how we might integrate such data using new forms of database design adapted to and distributed at the city-wide scale. Moreover we also see the idea of crowd-sourcing as key to many new data sets that will be useful to smart cities while noting that these types of interactive technologies can also be used to elicit preferences and to engage in social experimentation with respect to what we know and think about key urban problems. New applications which elicit quite subjective preferences concerning happiness and associating these with places are currently being developed (`http://www.mappiness.org.uk/`).

The basic ingredient for the new wave of city analytics that has emerged during the last decade is big data sets concerning human mobility, fostered by the widespread diffusion of wireless technologies, such as the satellite-enabled Global Positioning System (GPS) and mobile phone networks. These network infrastructures, as a by-product of their normal operations, allow for sensing and collecting massive repositories of spatio-temporal data, such as the call detail records from mobile phones and the GPS tracks from navigation devices, which represent society-wide proxies for human mobile activities. These big mobility data provide a powerful social microscope, which may help us understand human mobility, and discover the hidden patterns and models that characterize the trajectories humans follow during their daily activity [8,21]. In such research, privacy is always considered but is rarely in danger for the data is anonymised through several levels of scrutiny and confidentiality.

The direction of this research has recently attracted scientists from diverse disciplines, being not only a major intellectual challenge, but also given its importance in domains such as urban planning, sustainable mobility, transportation engineering, public health, and economic forecasting. The European FET project GeoPKDD (Geographic Privacy-aware Knowledge Discovery and Delivery, `www.geopkdd.eu`, 2005–2009) is a precursor in mining human mobility data, which has developed various analytical and mining methods for spatio-temporal data. This and other projects, in Europe and internationally, have shown how to support the complex knowledge discovery process from the raw data of individual trajectories up to high-level collective mobility knowledge, capable of supporting the decisions of mobility and transportation managers, thus revealing the striking analytical power of big mobility data. Analysts reason about these high-level concepts, such as systematic versus occasional movement behaviour, purpose of a trip, and home-work commuting patterns. Accordingly, the mainstream analytical tools of transportation engineering, such as origin/destination matrices, are based on semantically rich data collected by means of field surveys and interviews. It is therefore not obvious that big, yet raw, mobility data can be used to overcome the limits of surveys, namely their high cost, infrequent periodicity, quick obsolescence, incompleteness, and inaccuracy.

On the other hand, automatically sensed mobility data are ground truthed: real mobile activities are directly and continuously sampled as they occur in real time, but clearly they do not have any semantic annotation or context. Much research has begun to show that the semantic deficiency of big mobility data can be bridged by their size and precision [23]. Large-scale experiments have shown how it is possible to find answers to many challenging analytical questions about mobility behaviour, such as: What are the most popular itineraries followed by individual travel and what is the spatio-temporal distribution of such travel? How do people behave when approaching a key attractor, such as a big station or airport? How do people reach and leave the site of an extraordinary event, such as an important football match? How can we predict areas of dense traffic in the near future? How can we characterize traffic jams and congestion? More than just examples, these questions are paradigmatic representatives of the analysts need to disentangle the huge diversity of individual locations and tracks to discover the subgroups of travel characterised by common behaviour and purpose [22]. It is no surprise, therefore, that finding answers to these questions is still beyond the limits of the current generation of available systems. In *FuturICT* we will push back these limits.

Explanations of why traffic significantly varies from one day to another even if demand profiles are similar is extremely weak and we consider that new ICT will provide us with dramatically new data sets that will inform this problem. It is not also clear how one could predict spatiotemporally the development and propagation of congestion with small errors. The severity of these effects is even stronger in case of non-recurrent events (e.g. accidents, road constructions), which can affect the resilience and productivity of transportation systems. How all of the above are related to the network topology and how small or large perturbations in demand profile and network characteristics affect choices of people (in terms of route, departure time and mode) is one of the challenges we need to address.

For long periods choices of people in transportation networks are based on equilibrium conditions with small variations. New data can help us understand whether or not the real urban traffic can be considered an equilibrium system with respect to a cost function, how people really make choices and how these choices affect the development and spreading of congestion in the networks. The large number of trajectories and disaggregated traffic data from cities of different sizes and in different locations globally will provide a unique way to identify the macroscopic observables and control parameters that affect individual decisions and integrate them in agent-based models. Current day-to-day traffic assignment models are not suitable for modelling the traffic evolution under strong changes of network topology (e.g. a heavy disruption or a new mode of transport such as that occurring in many cities in developing countries because such models assume that drivers build on their experiences from past days. But when significant network changes occur, lack of observation measurements do not allow for the realistic modelling of pattern evolution and identification of equilibrium or non-equilibrium. Our research programme in smart cities will attempt to address all these challenges.

The concept of reality mining [14] which concerns highly pervasive sensing of complex social systems using ubiquity of mobile phones is a key determinant of how citizens interacting with each other in cities are becoming smarter. Reality mining depends on software to log location, communication activity, as well as usage of other services and applications. The analysis of derived datasets helps extracting eigenbehaviours to identify structure in everyday routine activities [15]. This data have enormous potential to gain new insights into urban dynamic processes at high temporal and spatial resolution at the scale of cities, given enough effort is put into resolving privacy concerns. Recent studies of datasets of mobile phone usage and geo-referenced online social networks and micro-blogging platforms have identified promising directions for this work. The main areas of interest are mobility patterns, spatial aspects in social network structure, group behaviour identification and data-driven characterization of city functioning.

From data mining and network analysis, we are able to create an urban mobility atlas, i.e., a comprehensive catalogue of the mobility behaviours in a city, an atlas that can be browsed (by the hours of the day, the days of the week, geographic area, meteorological conditions, and so on) in order to explore the pulse of a city in varying circumstances, while also observing emerging deviations from normal [7]. To fully realise the idea of an urban mobility atlas for the smart city, there is the need to integrate increasingly richer sources of mobility data, including the data from public transportation, road sensors, surveys and official statistics, social media and participatory sensing, into coherently integrated databases, as well as connecting mobility with socio-economic networks. This integration will be extremely relevant to understanding how public energy saving transportation systems could satisfy the demand for individual mobility.

2.1.3 Connection: The idea of coupling networks

The social fabric of a city is the result of many intertwined, multi-faceted networks of relations between persons, institutions, places, and more: beyond mobility, we need to take into account social and economic networks. We believe that the key insight for understanding the city is in understanding the structure of these coupled networks, and how this structure evolves. One important example of coupling/connecting the different facets of urban life is between mobility, consumer exchange involving retailing, and social networks. Despite the recent explosion of research in social networks pushed by big data, the bulk of work has focused on the social space only, leaving important questions involving to what extent individual mobility patterns shape and impact social networks, barely explored to date, but there is some preliminary work on transport behaviour and its psychology [42]. Indeed, social links are often driven by spatial proximity, from job- and family-imposed programs to joint involvement in various social activities. These shared social foci and face-to-face interactions, represented as overlap in individuals trajectories, are expected to have significant impacts on the structure of social networks, from the maintenance of long-lasting friendships to the formation of new links.

Our knowledge of the interplay between individual mobility and social networks is limited, partly due to the difficulty in collecting large-scale data that simultaneously record dynamical traces of individual movements and social interactions. This situation is changing rapidly, however, thanks to the pervasive use of mobile phones. Indeed, the records of mobile communications collected by telecommunications carriers provide extensive proxies of individual trajectories and social relationships, by keeping track of each phone call between any two parties and the localisation in space and time of the party that initiates the call. The high penetration of mobile phones implies that such data captures a large fraction of the population of an entire country. The availability of these massive CDRs (Call Detail Records) has made possible, for instance, the empirical validation in a large-scale setting of traditional social network hypotheses such as Granovetter's strength of weak ties [25], and the development of a first generation of realistic models of human mobility and their predictability [24].

Indeed, despite the heterogeneity of spatial resolution (the uneven reception area of mobile phone towers) and sampling rates (the timing of calls), the large volume of

CDR data allows us to reconstruct many salient aspects of individual daily routines, such as the most frequently visited cells, and the time and periodicity of such stays. Therefore, these data help us scrutinize the spatial patterns together with social structure and the intensity of social interactions. Recent studies [2] show empirical evidence that these three facets, co-location, network proximity and tie strength, are correlated with each other: the higher the likelihood that two persons co-locate (i.e. are observed in the same location/cell), the higher the chance that they are strongly connected in the social network, and that they have intense direct interactions [31]. The emergence of such surprising three-fold correlation hints that it is conceivable to explain (and predict) how new social ties will form as a function of mobile behaviour, and vice versa. In perspective, by coupling the social and mobility networks with further information, it may be possible to edge towards exploring the evolutionary dynamics of the urban social sphere, predict the spreading of sentiments, opinions and diseases, and thus to understand in real time the evolving borders of the community structure of a city.

To make sense of this great proliferation of data, we need to establish standards for integration of this data, for ensuring that quality standards are met, for assessing the accuracy and error in such data, and for providing ways of filling in missing data using models of the very systems that this data pertains to. Much of this data is networked and we consider that coupling such networks of data bases will be key to making sense of this material. New methods of coupling in terms of hardware and software will be needed and this will be central to the sort of collective intelligence functions that we see the smart city developing. Currently control centres in cities tend to exist only for the most routine and constrained systems such as traffic but to realise the vision of the smart city such networked-based coupling will be essential. This is an enormous challenge.

2.1.4 Coordination: The need for joined-up planning

Urban planning was first institutionalised in the late 19th century in several western industrialised nations and since then, it has diffused as a function of government in most countries worldwide. Its functions are exercised at national, regional, metropolitan, city, district and neighbourhood levels with rural, environmental and transport planning representing more specialised foci. However, planning is much wider in its import than these institutional frames for it is exercised as a function of many businesses and community groups, indeed across every activity that has a distinct interest in the city. In short, when we call for joined-up planning, we mean integration across the board, selectively and sensitively of course, but integration that enables system-wide effects to be tracked, understood and built into the very responses and designs that characterise the operations and functions of the city. This relates very strongly to the previous theme involving connection, networks and data integration. The sort of intelligence functions that we envisage for the smart city would be woven into the fabric of existing city institutions whose mandate is producing a better quality of life for its citizenry.

We envisage that the smart city would focus on the usual components that make the city function as a competitive entity as well as a social organism. The idea of ICT penetrating wherever it can improve performance and generate a better quality of life is central to this quest [18]. In this project, we will focus on ICT in buildings and the built environment, urban design, transport planning, local planning, metropolitan planning, regional planning, and upwards to the European level where, for example, the ESPON (European Spatial Planning Observation Network) project is already mobilising resources to examine smart city ideas across Europe through traditional urban and regional planning instruments.

2.1.5 Participation: Citizen science

Public participation is a long-standing tradition in institutionalised planning but the emergence of the digital world has turned the activity on its head. The ability for all citizens to communicate with one another and with agencies and groups that represent them, has provided a new sense of urgency and possibility to the idea that smart cities are based on smart communities whose citizens can play an active part in their operation and design. There are many such initiatives at the present time and we will focus on ways in which citizens can first access information about what is happening in their communities and cities but also explore ways in which a wide range of different groups can become actively involved in the design and planning process, both remotely and in face-to-face situations using data, models and scenarios all informed by contemporary ICT [27].

Current forms of participation are responding to new ICT but still remain inert and somewhat passive. New media and the web are increasing the liquidity of this type of interaction as both data and plans are being shared [16]. Participation is becoming more bottom-up than top-down, more in the spirit of the way complex systems actually evolve. In *FuturICT*, we envisage that we would pioneer a number of demonstrators as to how an informed citizenry might engage with experts from many domains in generating scenarios for improving the quality of urban life and urban performance, in ways that hitherto have not been possible. This will require a huge mobilisation of resources which draw on many aspects of the *FuturICT* proposals and imply serious progress in data, model, and policy integration. Already key data sources are being opened up such as mapping data, crime and policing, house price data, health data and so on and this will provide the momentum for the various demonstrators that we will initiate.

In our vision, participation and self-organisation are the cornerstones to building a global knowledge resource that, by design, will represent a public good, accessible to every citizen, institution or business. On the one hand, people should be fully aware of the kind of public knowledge infrastructure they are contributing to, and of the potential benefits they will be able to get from it. On the other, people should be in full control of their contributed data/profiles: how their data are being acquired, managed, analysed and used, when and for how long. Only a public system capable of delivering high-quality information within a trusted framework has the potential for raising a high degree of participation, and only large, democratic participation can ensure the creation of reliable, timely and trustworthy information about collective phenomena. This view is at the basis of a citizens science, where sentiment and opinion mining from trusted information can detect shifts in collective mood in a timely manner, detect the weak signal of important changes, and detect the structure and evolution of social communities.

2.2 Understanding smart cities

2.2.1 Sensing and measuring

In the past, cities have been understood and planned across several levels. Currently there is real momentum in sensing urban change from the ground up, so-to-speak, using new sensing technologies that depend on hand-held and remote devices through to assembling transactional data from online transactions processing which measure how individuals and groups expend energy, use information, and interact with respect to money. Network technologies make possible extensive data collection and coordination of such data in terms of the data itself and in terms of how that data is stored and made accessible. New forms of data base organisation and mining offer the prospect for adding value to the data in terms of massive data integration. The list of components and sectors in the city is almost too long to catalogue but the key sectors which currently are being heavily networked involve: transport systems of all modes in terms of operation, coordination, timetabling, utilities networks which are being enabled using smart metering, local weather, pollution levels and waste disposal, land and planning applications, building technologies in terms of energy and materials, health information systems in terms of access to facilities by patients the list is endless. The point is that we urgently need a map of this terrain so that we can connect up these diverse activities.

There is a strong evidence of spatial patterns in social structures. Distance decay in spatial interaction processes is a well-known phenomenon, and many recent studies on mobile phone data sets have confirmed distance decay relationships: the likelihood that two individuals are connected decreases with distance between them [30]. Generally, social structures of communities detected from phone data analysis show strong spatial regularity in regional [20] and city scale levels [45], revealing finer patterns concerned with cultural and socio-economic heterogeneity of cities. More work is required to extend these findings to reveal temporal dynamics as the potential for optimising urban transportation systems through exploiting social structure is enormous. Spatial interaction also manifests itself in strong spatial consistency in the function of the city, both defined by infrastructure and underlying urban planning, so as in terms of the ways people tend to use cities [38] and the temporal evolution of these patterns. The eigenplaces approach does not ascribe any semantics to these regimes but it is possible to provide an interpretation by examining changes over space and time, taking into account land use and services distribution in the city. The analysis of content-rich data from micro-blogging adds novel dimensions to these studies. Content analysis allows semantically-rich analysis of land use [41], temporal variability of current-moment interests within cities and related travel behaviours [37], and inter-city comparisons which reveal common dependencies in human mobility across countries and continents [34]. Notwithstanding these developments, integrating diverse real time data from different sources is problematic not only from a computational perspective but also in terms of building integrated models.

Linking GPS, satellite remote-sensing, online interactive data systems focussed on crowd-sourcing, all with the automation of standard secondary sources of data, and then meshing this with more unconventional data elicited from social media provides a very rich nexus of possibilities in terms of providing new and open sources of data essential to a better understanding of how smart cities will function.

2.2.2 Movement and networks

Travel is the minimum price we pay to participate in out-of-home activities as individuals. Networks are the infrastructure we provide collectively to reduce these costs. Increasing urbanisation in conjunction with the growing depth of networks in terms of technologies (water, travel, energy, communication) put an emphasis on our ability to grow these networks in a way which strikes a compromise between their current and future costs, reliability and resilience. The ever-growing density and volume of information about the state of the networks as built artefacts and about the state of the networks in terms of user experience provides an enormous opportunity to improve both their planning and operation. This growing complexity should be approached with local solutions and local rules (grammars) to reduce the planning costs for all participants, especially in parts of the world where the planning system does not have the strength to impose or bribe large scale and long range solutions. These grammars will have to be defined at both large and small spatial scales to address both the local and regional/national scale of the demands of society. These grammars will be the focus of the research work planned.

2.2.3 Travel behaviour

The density and timeliness of information flows allow travellers to respond immediately with an ever more complete picture of the situation in mind, if they are confident enough in the accuracy of the information and in their ability to judge the impact of the information used on the response of all other travellers. The difficulty for the traveller is to learn how much he or she can trust her judgement in a particular style of situation and related to this, how much weight should she give to the particular source (radio station A, government advice B, friend C, the last minutes of travel flow/movement observed). It is also part of the learning process to learn when the effort of changing the route or mode is not worth the effort where maintaining the existing plan is the least-effort choice.

This empowerment of the traveller is a challenge for any advice system where it aims for self-consistency. The self-consistency of any advice system in a group of advice systems is an unexplored problem, especially for a mixture of public and private systems with different assumed objective functions for the travellers. The *FuturICT* project will address this issue through both mathematical and simulation-based work, both to gain understanding of the interactions between the travellers and different advice systems, as much as to guide the systems to the best achievable system configuration given the number and types of players involved. The development of multimodal trip planners and advice systems are in their infancy and we expect *FuturICT* to spur the development of such applications.

2.2.4 Land use and transport

The increasing richness in social terms of the telecommunication experience raises the question of whether or not cities are still needed in their central function as places, which enable innovation and scale through the enforced concentration of many actors: suppliers and buyers; designers and engineers; businessmen and retailers; scientists and craftsmen. The increasing physical range of travellers (and logistic-chains) reduce the importance of any one place, while not discounting the symbol importance of certain locations for different actors groups and milieu. The net of networks, both social and physical, needs to be explored using a wealth of traces which our computing and telecommunication systems can now generate. The assessment of their structure and impact needs to cover the whole range of impacts from the economic to the environmental; both in steady state and in crisis, and both with regards to the possibilities of governmental control and its ability for emergent social control.

2.2.5 Urban markets and exchange

Cities are essentially sets of markets where individuals and groups come together to exchange. This is their traditional rationale to overcome the friction of distance which enables people to compete for scarce resources in the highly favourable conditions generated by agglomerations of labour with quite specialised tasks and skills. It is well known that ICT changes the effects of distance (and cost) quite radically and it would be surprising if this were not critical to the structure of cities and the way they function. So far markets in cities have worked in traditional ways but with globalisation, network economies, and the substitution and complementation of information for energy and materials, the local economy is moving increasingly online and this is influencing locational decisions in ways that we are largely unaware of. It is clear for example that bookshops and some other retail outlets are changing in terms of their presence and location as much of this activity moves to the net but in the case of housing markets, labour markets, economic development, the demand and supply of transport facilities and access to education and health care, the impact of ICT on markets is complex and largely unexplained. Moreover the development of many new systems for the movement of goods and freight is likely to revolutionise the logistic industry and new systems are emerging from the impact of ICT on transport.

We need to understand these markets in the network economy much more clearly and part of our proposal will be to initiate new approaches to such understanding, linked strongly to other projects in *FuturICT* that involve the simulation of financial markets and behaviours [11]. Moreover, new networks and markets are emerging such as the market for energy [17] and this is yet another example where the city is increasing in complexity as human behaviour is enriched by access to new ways of deciding how to utilise spatial resources.

2.2.6 Firms and organisations

The current modelling approaches used in urban planning have a very reduced understanding of firms and their spatial behaviour. Based on the possibilities of tracing choices over time and space on the web, it will be necessary to develop agent-representations of the firms by size, type and sector. These agents will be crucial in the new models of land use, transport and the economy. Without them, agent-based modelling of urban markets would be incomplete and potentially misleading. Traditionally aggregate economic forecasting models based on input-output analysis have dominated this area hitherto but there is an urgent need for new models which reflect stronger behavioural rules that are clearly relevant to such decision-making.

2.2.7 Communities and networks

Traditionally there has been a strong focus in cities on questions of community which 50 years ago were largely based on replacing worn out infrastructure in the form of housing built during the 19th century. The public sector was dominant in these activities in many western industrialised cities but from then on, the role of this sector has declined. The focus has also shifted from problems manifested in the built environment to problems of social deprivation and the lack of economic opportunities. Policies involving welfare and social conditions combined with positive attempts to steer labour markets have begun to replace more infrastructure-based instruments. ICT is making obvious inroads into the delivery of information about these issues to the affected populations while at the same time providing interactive advice concerning how such opportunities can be realised. Smart technologies are central to this as they are to related services such as health and education. There is however still a massive effort required to enable such services to connect to one another and to gain real value through such connectivity.

The role of ICT in terms of collecting and modelling new ideas about community through much deeper analysis of social networks is an obvious extension to our analytical capabilities in dealing with questions of community. *FuturICT* will explore

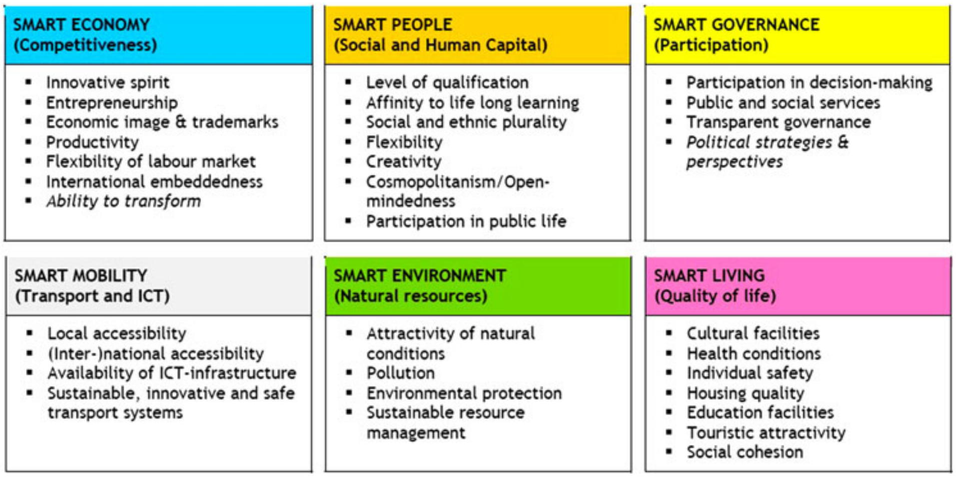


Fig. 2. A typology of smart city functions.

the way community networks can be generated using new social media and related connectivities that can be mined from mobile device data bases and web sites, and explore how these can be linked to data on housing and labour markets. Moreover, in planning smarter communities, there is a design dimension to social networks, which needs to be factored into new ways in which we can generate plans, involving the community itself in the analysis of its own data.

2.3 Planning smart cities

2.3.1 The need for coordination and coupling

Coordination, communication, coupling and integration are different perspectives in developing the smart city which we see as a programme of connecting up infrastructures and services so that the city can function more effectively. This will require new forms of database, new methods of mining and pattern analysis, new software for integrating diverse and hitherto unconnected components and sectors in urban functioning, and new forms of organisation and governance, which will enable such connectivity to become effective and fair. The smart city balances efficiency against equity with a focus on improving the ability of its citizenry to innovate through a balance of cooperation with competition. There are strong links here to uniform markets for generating greater liquidity and mobility in the use and provision of urban services and these goals lie behind the mission statements produced by governments and businesses for the smart city. There are many examples of how various sectors might connect up presented in these papers and mission statements. Here we show one of these as an exemplar: this is reproduced below in Fig. 2 from two sources: from `www.networks-etp.eu` [18] and from `www.smart-cities.eu` [19].

2.3.2 New data systems and integration

In our quest to master the complexity of the knowledge discovery process for the smart city, we need to build an entirely new holistic system for integrated data acquisition, querying and mining. The entire analytical process able to create the knowledge services should be expressible within systems which support the following:

- The acquisition of data from multiple distributed sources, including services for participatory sensing and online communities
- The management of data streams
- The integration of heterogeneous data into a coherent database
- Data transformations and preparations
- Definition of new observables to extract relevant information
- Methods for distributed data mining and network analytics
- The management of extracted models and patterns and the seamless composition of patterns, models and data with further analyses and mining
- Tools for evaluating the quality of the extracted models and patterns
- Visual analytics for the exploration of behavioural patterns and models
- Simulation and prediction methods built on top of the mined patterns and models
- Incremental and distributed mining strategies needed to overcome the scalability issues that emerge when dealing with big data.

Although preliminary, partial examples of this line of research already exist, in specific domains, such as mobility, which is fast becoming an exemplar for encompassing all the domains of data, patterns and models for the smart city, a continuing research challenge for *FuturICT*.

2.3.3 Governance in smart cities

We have already argued that a much stronger intelligence function is required for coordinating the many different components that comprise the smart city. These will depend on some sort of structure that brings together traditional functions of government and business. Business has the expertise in providing hardware, software and data solutions enabling cities to be smarter while government is engaging users of services, community and citizen interests whose traditional focus in on the quality of life of their communities. Such governance reaches out to higher level NGOs whose function is to set the community in a wider context extending to extra territorial agencies such as the EU and of course national governments.

There can be no one agency responsible for all of this but just as cities are increasingly being seen as constellations of active agencies and groups, so their overall governance and coordination can be constituted in the same way, from the bottom up as well as the top down. Again, the idea of governance that extends in this way to the many functions that we envisage and will be coordinated in the smart city, is a relatively new prospect and is part of the wider debate about decentralisation of governance in the information age [28]. It relates strongly to privacy, security as well as economic performance, social inclusivity and a host of issues that are being changed by new ICT. We envisage that this theme will be part of our programme of research related to new organisational infrastructure for smart cities that are built around new developments in ICT. We consider that the traditional role of planning which now includes regeneration, traffic, economic development, and housing, will involve questions of the operation of utilities, the access of citizens to services, health, education, indeed any functions which have a spatial effect on the city in terms of location and movement.

2.3.4 New methods for design and planning

Over the last 50 years, various simulation models operating at different spatial scales and over different temporal intervals have been developed for understanding how cities function. Most of these models have been focussed on understanding as a prelude to their use to inform the planning and design process. These models have been focussed largely on simulating the location of physical activities, albeit through an economic and demographic lens which enables material transport and the location of land uses to be predicted using computer models of various sorts. Some extensions of such models into optimisation have been made as in urban operations research but generally such models are used in a somewhat intuitive dialogue with policy makers in the context of what are now called planning support systems [16].

The emergence of the smart city poses enormous challenges for these styles of modelling for many reasons. First, the city itself is being transformed from a place dominated by physical actions to one in which such action is complemented by extensive use of information technologies. Second, many routine functions in cities are being replaced by computer control and various forms of automation are increasingly being blended with human actions. Third, the provision of data from these new electronic functions in the city offers the prospect of a world in which the implications of how the city is functioning is continuously available and such immediacy is compressing time scales in such a way that longer term planning itself faces the prospect of becoming continuous as data is updated in real time. We also face the prospect of developing intelligence and planning functions at the same time as the very object that we are concerned with the city is changing its nature due to similar if not the same functions being used in its operation. This kind of space-time convergence in cities implies a level of complexity that only the new and powerful science of the kind that we will pioneer in *FuturICT* can address.

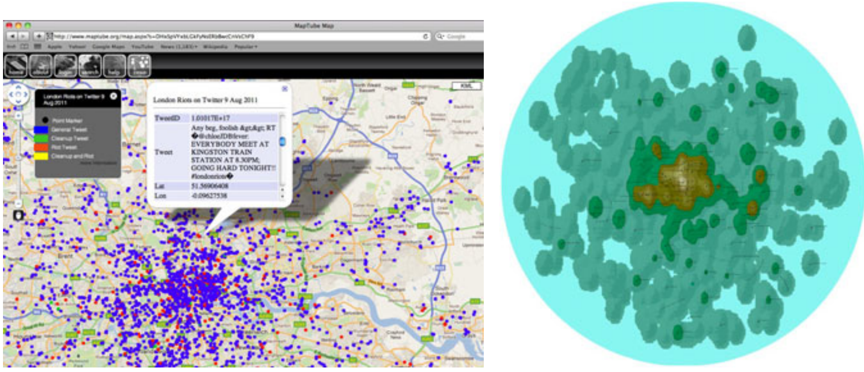
Much of our focus on smart cities will be in evolving new models of the city in its various sectors that pertain to new kinds of data and movements and actions that are largely operated over digital networks while at the same time, relating these to traditional movements and locational activity. Very clear conceptions of how these models might be used to inform planning at different scales and very different time periods are critical to this focus. We envisage that quite new forms of integrated and coordinated decision support systems will be forthcoming from this research.

2.3.5 Participation and online communications

We have already implied that new forms of participation in developing the smart city need to be generated from new forms of ICT. To date, web-based participation is largely passive and only quite recently have Web 2.0 technologies which presume true interactivity come on stream. In fact for purposes of participation in designing the city, there is an enormous overhead of time and interest required and therefore processes have to be devised which enable interested citizens to make an impact in terms of their participation. The HCI issues in such developments are crucial as are issues of privacy and confidentiality. Online communication comes in many forms and in terms of mobilising the wider citizenry, then we can define at least four key modes of interactivity: first, portals and other access points to useful information about any aspect of routine living and working in cities, second ways in which citizens can interact with software that enables them to learn more about the city by engaging with other users online and actually creatively manipulating information, third citizens engaging with crowd-sourced systems in which they are responding to queries and uploading information, and fourth, fully fledged decision support systems which enable citizens to engage in actual design and planning itself in term of the future city.

2.4 A sample of contemporary exemplars

There are hundreds of examples now that pertain to how ICT is being embedded into cities yielding new data for understanding key urban problems, enabling better urban



(a) (b)

Fig. 3. Mapping Twitter feeds: (a) the London riots, Summer 2011, and (b) a map of Twitter feeds over 24 hours in Paris.

functioning and generating new solutions that improve urban performance and the quality of urban life. We will identify seven and simply outline their salient points to give some sense of how ICT is being successfully used and how it might be used in the short and medium term future.

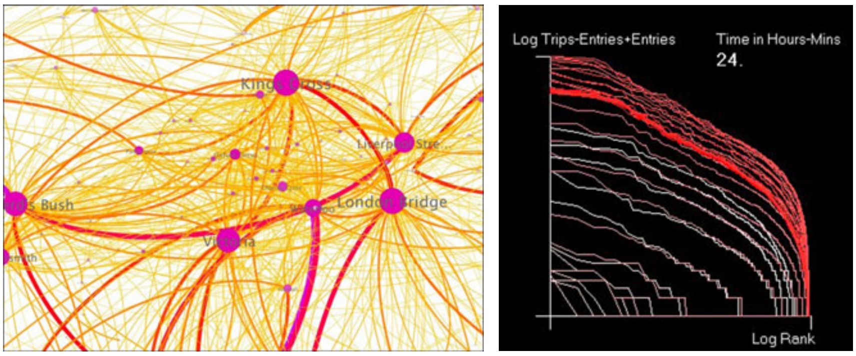
2.4.1 Real time sensing: Crowd-sourcing and mapping social media

The most high profile social media currently is short text messaging in which any kind of information can be transmitted in less than 140 characters to anyone signed up to particular network systems, the best example being Twitter. Any smart device or computer that is able to access the Twitter domain can enable a geo-coding facility that locates where the message is sent from and there is considerable research into how social networks as well as spatial networks might be fashioned from such data. Geo-locating the sources of such messages is in fact the main application to date. We are only just beginning to see the potential of such media and the extraction of content is a major issue. However it is likely that there will be considerable advances in content to extract useful information and this might be linked to new schemes for mobility management; even if this is not the case with Twitter data, the point is that similar data will come on stream that is likely to be more focussed as this kind of technology becomes widespread. Below are maps of Twitter feeds showing the London riots of August 7–9, 2011 in Fig. 3(a) and a map of Twitter densities in Paris produced during 24 hours in June 21, 2010 in Fig. 3(b).

The largest data at present of a more professional usage is produced by crowd-sourcing and this is Open Street Map which is produced by armies of volunteers recording positional information using GPS like technologies which deployed in the field. We will not show this here but it has high accuracy in many areas of the world and it can be used as a base-map for other social media data as in the map above of Twitter feeds. The map above is Google Maps but OSM is an obvious alternative.

2.4.2 Multiple networks: The london oyster card data

Movement on public transport in large cities is increasingly made by passengers using cards that are loaded with cash to facilitate more than one journey. In London, all buses, overland (heavy rail), and underground (tube) can be used with an integrated



(a) (b)

Fig. 4. Daily flows (a) and volumes (b) between and at hubs on the London rail network from qyster card data.

card and there is a detailed data base of almost one billion events corresponding to entries, exits, transfers, and refunds over a 6 month date ranging from November 2010 June 2011 which we have been mining and visualising. So far we have examined data on the overground and tube systems and we are intent on examining patterns in the flows between hubs as in Fig. 4(a) and volumes at hubs as in Fig. 4(b).

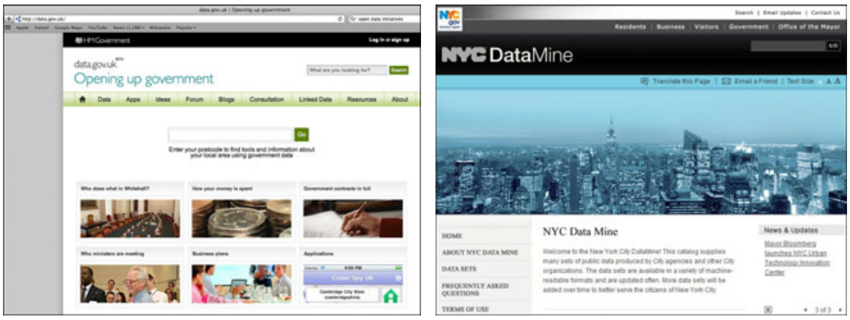
Algorithms are being developed for constructing multimodal trips associated with this data that will require various assumptions about movement between modes due to the fact that it is only on rail that the card is used for swiping in and out, thus giving origins and destinations. The map on the left in 4(a) shows flows between various hubs while that on the right in 4(b) is a series of rank size plots of the total traffic volumes of all nodes in the network at every 20 minute interval over the 24 hour day. Already as part of such projects multimodal trip advisors are being created.

2.4.3 New urban data systems: Open data

A number of national governments have developed initiatives in opening up public data to a wide audience of interested publics and professionals. So far, following the lead of the US, many other countries such as the UK and many city governments have taken up the challenge and are making public data available in many different formats. This is also part of the transparency agenda in contemporary government, which is founded on accountability but it also relates to questions of confidentiality and privacy. The EU is also heavily involved in these issues with respect to PSI directives (`http://ec.europa.eu/information_society/`). The portal for the UK is shown below left in Fig. 5(a) and the New York City equivalent right in Fig. 5(b). In terms of smart cities, we envisage that many such portals will emerge in the next decade.

2.4.4 New models of movement and location: MATSim and simulacra

The range of interactions between the different agents in the transport and land use system requires agent-based approaches to capture their impacts: travellers, network operators, transport service providers (taxi, car pooling, car sharing, bus, trains, etc.) information system providers, activity providers (retailers, bar operators, restaurant chains, cinema owners, etc.), developers and property owners, policy makers in competing jurisdictions and others. Open-source agent-based micro-simulation such as



(a) (b)

Fig. 5. Portals for open data: (a) the UK and (b) New York city.

(a) (b)

Fig. 6. Land use transport models: an ABM for Tel-Aviv at different scales (a) and an aggregative model for the London metropolitan region (b).

MATSim has an excellent track record for computational speed and the size of the problems it can address. It will be the basis for a richer version of the model which will enable the inclusion of these different agent classes in both equilibrium and non-equilibrium (path oriented) versions. We show a typical application for Tel Aviv (in Fig. 6(a) where the model is being used to simulate traffic at different spatial scales where interaction with activities at these scales is modelled. (See `www.matsim.org` for software download, tutorials, papers and result animations). Simulation using aggregative models of a more traditional kind are illustrated in Fig. 6(b) for the London region. *FuturICT* will also improve the speed and quality of these models in terms of computational infrastructure being developed and of course the range of issues that such modelling systems are able to address.

2.4.5 Risk analysis of development paths

Brute force risk assessment of a system as complex as a metropolitan area is beyond the current computing abilities of even the highest speed clusters: the number and types of agents are too large and the time to enable such computation too long. The aim has to be to develop intelligent systems to (a) capture the range and correlation structure of the many driving factors of urban development and (b) search mechanisms which are able to identify the range and form of the joint distribution of the central outcomes in terms of quality of life, urban success, resilience and robustness. And of course extreme events such as terrorist attacks, natural disasters, criminality and so on.

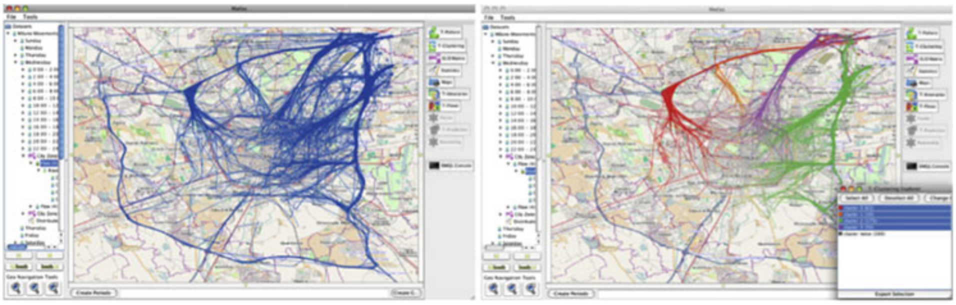


Fig. 7. GPS tracks in metro Milan (a) city centre to the North East and (b) clusters of like trajectories.

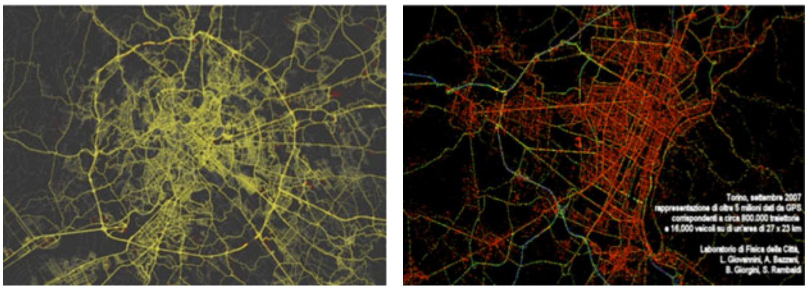
2.4.6 New models and systems for mobility behaviour discovery: M-Atlas

The M-Atlas system was conceived as a framework to master the complexity of the mobility knowledge discovery process. M-Atlas is an integrated querying and mining system, centred onto the concept of a trajectory, i.e., a sequence of time-stamped locations, sampled from the itinerary of a moving object. Trajectories pertaining to human travel or movement can be reconstructed from data sensed in various contexts, including GPS tracks from vehicular or hand-held navigation devices, call detail records from mobile phone (GSM) carriers and providers, time-stamped location records from online services or social networks, such as Flickr, Foursquare, Google Latitude, and so on. M-Atlas supports the complex knowledge discovery process from raw data of individual trajectories up to high-level collective mobility knowledge, by means of methods for:

- Trajectory reconstruction from the raw location data
- Trajectory database management and querying
- Trajectory mining: pattern extraction, clustering and prediction/classification
- Trajectory visual analytics and model presentation/exploration.

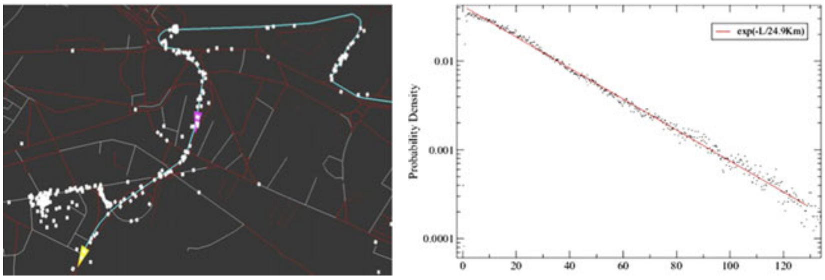
As an example, Figs. 7(a) and (b) above show a possible analysis performed on dataset of GPS tracks from approximately 20,000 private cars sensed over a period of one week in the metropolitan area of Milan, Italy. In the example, the trajectory clustering method reveals typical profiles of commuting behaviour from two selected areas in the city. In the left-hand Fig. 7(a), the M-Atlas visual interface shows the data selected as input for the clustering algorithm, namely the trajectories from the city centre to the North-East suburbs. Figure 7(b) to the right shows the clusters obtained, where the trajectories in the same cluster are visualized by M-Atlas with the same colour. Once the most popular commuting profiles and routes have been identified, the analyst can zoom on the different clusters, studying, e.g., the hourly distribution of travels, etc.

Example projects that are being pursued by various stakeholders using M-Atlas as enabling technology include the characterisation of tourist profiles in Paris, France, on the basis of GSM roaming call detail records, the simulation of a participatory car pooling and car sharing service in Tuscany, and the preparation (based on integrated GPS, GSM, sensors, participatory mobility data) of a detailed mobility atlas for the urban area of coastal Tuscany, as a tool of policy definition for the public administrations at urban and regional scale. We will begin with this M-Atlas as a map of the smart city and then extend it to other kinds of Atlas that relate to the integrated databases that we will progress.



(a) (b)

Fig. 8. (a) GPS volumes in Rome and (b) velocities (red to blue) in Turin.



(a) (b)

Fig. 9. Trajectory reconstruction (a) and statistical behaviour mobility (b).

2.4.7 New tools for the governance of mobility demand

GPS technologies allow us to record individual mobility data across an entire urban network. In Italy, a sample of 3 percent of the whole vehicle population is monitored for insurance purposes, providing information on single trajectories with a spatial scale of 2 km and a time scale 30 seconds [9]. Moreover, one datum is always recorded when the vehicle engine starts or stops. Each datum includes position, speed, motion direction, GPS quality. In Fig. 8, we show the data georeferencing for whole metropolitan area of Rome (40000 vehicles) along the entire month of May 2010 in Fig. 8(a) and the metropolitan area of Turin (16000 vehicles) during the month of September 2007 in Fig. 8(b). The colour scale (from red to blue) gives information on the travelling velocity thus illustrating the structure of the network.

Despite the relatively poor spatial resolution of such GPS data, it is possible to perform a real time reconstruction of the individual trajectory dynamics on the road network. This result is achieved by applying specific algorithms that select the paths consistent with experimental observations, the individual habits of travellers and different road usage. In the Fig. 9(a) we show an example of trajectory reconstruction where we also plot the historical data to point out the relevance of habit behaviours of individuals. Moreover the data analysis suggests the existence of general individual behaviours related to the use of urban space-time. In Fig. 9(b) we show the spatial distribution of the daily individual mobility derived from GPS private car data: the straight line in a logarithmic scale implies an exponential decay with the length that reminds us of the statistical Boltzmanns distribution. The results shed lights on the average traffic properties in a city, but future microscopic data will allow us to study the evolution of transient states from a microscopic (individual) point of view [10].

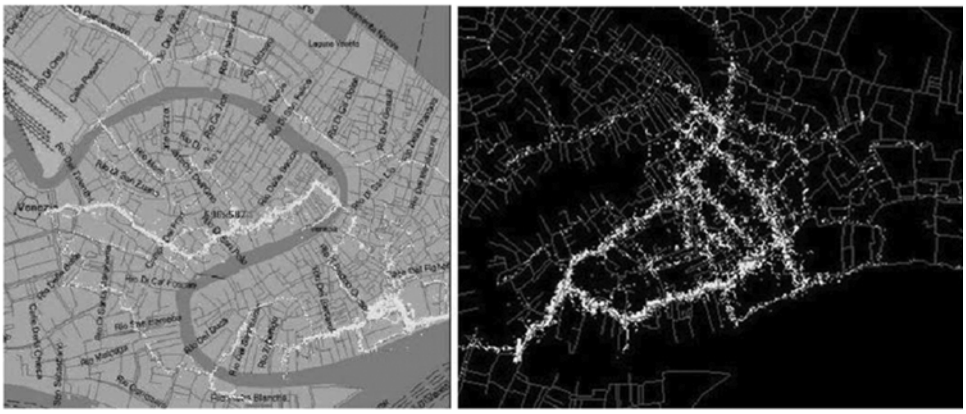


Fig. 10. (left) and (right) Reconstruction of movement dynamics of pedestrian movements at different scales in the Venice carnival 2007.

This also opens new research opportunities that use microscopic mobility data as a paradigm to study the human decision mechanisms and the information based interactions. These macroscopic laws will be the starting point of a new generation of microscopic models based on individual mobility demand, and will enable us to perform a real-time reconstruction of the traffic state across the whole urban network (nowcasting), to integrate the private mobility with the public mobility realizing low-energy sustainable transportation policies, and to predict future scenarios simulating emerging crisis events. This activity enters in a road-map toward a safe-city within the *FuturICT* project. Our aim is to generate an entire research dimension with respect to the role of failsafe mechanisms which pertain to crises that are generated by problems of mobility.

Moreover in order to understand human mobility behaviour, one takes in account the cognitive issues which originate the intentional dynamics of travellers. ICT technologies allow us to investigate the crowd dynamics using a microscopic approach. For example, Figs. 10 (a) and (b) shows pedestrian flow reconstruction via GPS measures at the Venice Carnival 2007.

Using this input data base, we can model and simulate the pedestrian mobility on the network and furthermore crowding in key places and nearby critical bottlenecks. Figure 11(a) shows a snapshot of a pedestrian flows simulation nearby Punta della Dogana (left) and of similar flows in 3D form as a snapshot of the crowding mobility in San Marco square Fig. 11(b). Extreme events will be certainly more and more frequent in the big future cities and their impact on citizenry safety and security is a key problem.

These data base types and modelling functions are fundamental in an ICT framework to project and to build up e-governance tools, and to connect individual information with the cooperative participation for fluid, safe and low-energy mobility.

2.5 Scenarios for the smart city

As we have noted there are many smart city initiatives as the term has become extremely popular for initiatives associated with the development of ICT and competition, with cities, particularly large cities, at the forefront of these proposals. We need to position ourselves on this spectrum of opportunities because there are several

(a) (b)

Fig. 11. Pedestrian movement dynamics in 2D (a) and 3D (b).

elements of the smart city movement that we will strongly relate to but others that will be more peripheral. We can identify at least seven types of initiative:

The development of new cities badging themselves as smart. These are proliferating in rapidly growing countries. Masdar outside of Abu Dhabi is being developed by GE as the worlds first carbon neutral city, Paredes in Portugal is where Microsoft are wiring an energy efficient city, Dongtan in the Yangtze Delta is being developed by Arup as a smart green eco-town, and Songdo in South Korea is where Cisco are building a town wired at all levels.

The development of older cities regenerating themselves as smart. In much more bottom-up fashion, which include many cities who are embedding new ICT as a matter of course. Examples of best practice are to be found in world cities where spontaneous developments of new technologies are emerging in places such as Silicon Alley (New York City), Silicon Roundabout (London) and Akihabara (Tokyo).

The development of science parks, tech cities, and technopoles focused on high technologies. Silicon Valley and Route 128 are the classic examples but the science park idea is still highly resonant with respect to local economic development where high tech production merges with its consumption in making such areas smart.

The development of urban services using contemporary ICT. In the form of networked data base, cloud computing and fixed and mobile networks, a force which is more central to our concerns here in coordinating diverse interests and sectors which will make the city smart in its design and planning.

The use of ICT to develop new urban intelligence functions. These are new conceptions of the way the city functions and utilise the complexity sciences in fashioning powerful new forms of simulation model and optimisation methods that generate city structures and forms that improve efficiency, equity and the quality of life.

The development of online and mobile forms of participation. In which the citizenry is massively engaged in working towards improving the city alongside planners and designers from government and business. Decentralised notions of governance and community action are central to these new forms of participations which use extensive ICT.

3 Our innovative approach

3.1 The essential tensions

The crucible for technological innovation is the cultural context in which it takes place. Technology is a social construction as much as it is a material or ethereal one, and its application is intrinsically social. There is an increasing consensus that cities represent the crucible for technological innovations and that larger cities with a highly educated workforce represent the best places where progress can be made with their invention and application. Globalisation complicates this picture but one of the reasons why so many companies and governments are embracing the idea of the smart city is that there is now a widespread view that to remain competitive and be ahead of the game, cities must mobilize ICT to become ever smarter in the pursuit of their competitive advantage. Alfred Marshall said it over 100 years ago but agglomeration economies, which come with cities growing ever bigger in terms of their populations and knowledge base lie at the core of the smart city. ICT holds the key to a better world and it will be most clearly demonstrated in large cities.

There are a number of innovations that our work on the smart city will establish. First and foremost, *FuturICT* is a programme that is founded on the application of complexity theory to human problems and as such, it is immediately apparent that the very subject of its focus in this project human systems are in and of themselves becoming ever more complex. This in turn is because of the invention of new modes of human functioning using ICT. This project is firmly at the forefront of understanding complex social systems using the very tools that are fashioning those social systems in the first place. As we noted earlier, McLuhans [32] notion that the tools that we shape then shape us is key. The science itself changes the very science that we are using. Cities, which adopt ICT in diverse forms, change the very nature of the adoption process by using that same ICT. The nexus is complex and we ignore this interwoven complexity at our peril. The problems that we deal with characterize all cities and are what many years ago [39] called wicked. When one tackled wicked problems, they became worse not better due to the unforeseen consequence and unanticipated effects which were ignored because the systems in question were treated in too immediate and simplistic terms. The great innovation of our program in smart cities as an exemplar of *FuturICT* is that we approach these issues in full knowledge of these dilemmas. We are particularly energised by concerns for privacy and confidentiality and the risks involved in the generation of new routinised individual based data that is emerging from all these initiatives.

3.2 The key themes

It might also seem at first sight that a programme related to smart cities would be strongly focused on hardware and networks but our focus will be much more on questions of organisation that imply software development and management of large scale computer resources, networks and data. As we have been at pains to point out throughout this paper, our focus is on integration of data, models, and users through ICT. This collective set of issues in ICT we might refer to as orgware, an old term dating back to the 1980s but one which symbolizes the constellation of issues that surround the use and development of new varieties of computation.

Developing and coupling databases which in turn are being forged using new kinds of media for collecting data through sensing, mining online transactions, and the automated recording of behaviour in the environment and communication, is one of the key foci in our project [46]. These kinds of coupling and the organisation that is developed will be part of new governance structures for the smart city new intelligence functions that utilize much wider participation in decision-making as well as real time construction and use of a variety of simulations and optimisations relevant to decision support. The research focus that we outline below pulls all these themes together in distinct projects, that inevitably will overlap with one another and with other projects in other domains of *FuturICT*.

One of our major themes will be the development of new forms of simulation model that embrace the new forms of complexity being developed in smart cities. An innovative aspect of our project is the development of a new class of simulation models for various activities in cities that will evolve as the city structures themselves evolve and become smarter. In other words, the models will simulate the city dynamics as self-organizing evolution processes, that mimic the Darwinian biological evolution in a balance between innovation and selection mechanisms. An example will suffice. Fifty years ago, cities were conceived in a manner that had barely changed for a thousand years: as a core dominated by the workplace and the movement of individuals to work and exchange goods (shop and receive services) in the car. This model was predicated on the basis that the city is a stable unchanging structure. Since then ICT and globalisation has dramatically shifted this model. Physically cities may not look very different from the material flows that inspired this past conception but in terms of their social networks and economic transactions, the old model simply cannot address current conceptions of a networked society. The world is now one that is as much dominated by flows of information that do not leave physical traces in the manner of the old. Our challenge is to build models that grapple with these changes and that have the potential to embrace very different conceptions of how the city might function. The rudiments of these types of models are already there in structures such as MATSim and Simulacra, in many of the agent-based models built for sectors of the city, and in new approaches to transportation modelling [35]. We intend to exploit these conceptions strongly within *FuturICT*.

Just as this then changes our conception of the way we might build models that respond to a changing and evolving system, our last theme relates to how the process of planning and decision might change to embrace the ability to sense the city in real time. Spatial scales and time scales are being collapsed by the emergence of real time data from the bottom up. Data sets are being created that show immediately the functioning of the real time city but also imply how long term changes in the city can be detected. In short, if all the data that we collected were in real time, at any instant, we could aggregate the data to deal with change in the city at any scale and over any time period. This prospect is a long way off and will never be reached (for once we reach it we will find more and different data that need to be collected) but what it does promise is an ability to have a real time view of change at different spatial scales and over different time scales. This will change both the models that we are able to build and the way in which these technologies can inform the planning and decision process with simulations and decision support being telescoped across space and time. This is crucial to the kind of citizen science that we will develop to provide a powerful participatory context to the future development of the smart city.

3.3 Our proposed projects

As we have noted, we cannot tackle all dimensions of ICT and the smart city and in fact, we will work with business interests such as IBM and Cisco to relate what they are doing to our wider quest. Therefore we will organize our research directions into seven distinct but overlapping areas that we show in aggregate form by the block diagram that is reproduced in Fig. 12 below. We define these areas as follows.

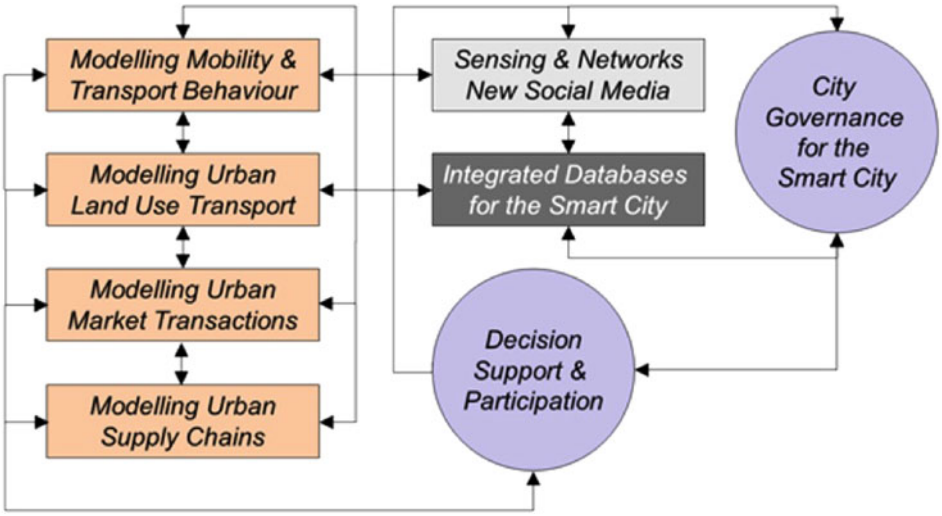


Fig. 12. The structure of *FuturICT*s smart cities programme.

3.3.1 Integrated databases for the smart city

We will select a series of databases that are being developed for different sectors of the city all of which rely on digital sources captured but not necessarily available in real time. We will consider how these data-bases can be enriched by adding data from more conventional sources such as recurrent cross-sectional censuses. We will explore error in such data, focus on standards of integration and provide an array of data mining and pattern recognition techniques, many based on machine learning to extract useful data for assessing the way the city is working. We will also enable these data bases to provide more aggregate possibilities useful for longer term decision support and we will focus on how new more unconventional and experimental sources of data through social media might also enrich these data sets. For example we envisage that we might link conventional online traffic data to utilities data such as water and electricity lines monitoring and location, and to social network data from STM (Synchronous Transport Module) network sources [17]. To this we will fuse population census data and consider how electronic data on land prices and residential transactions might be linked. There is nothing equivalent to this kind of integration available anywhere although there are attempts beginning in the GIS community.

In terms of transportation data, then we will collate available data streams from different European countries and this will require their translation into a common terminology and where necessary adjustments to make them comparable. The necessary contracts with their providers will have to be negotiated to make their collection smooth and reliable (APIs instead of web-bots). Examples are traffic counts, flight movements, container movements, shipping rates, TCP/IP traffic, tourism flows, mobile phone use, FAA 10 percent ticket sample, commercial air ticket data base, EUROSTAT statistics, National Population Census data, health warnings and so on which will constitute a particular focus of our database integration for Smart Cities. We will also derive key leading indicators using the best techniques to summarize the various data streams, while accounting for both their temporal and spatial nature. The publication of the key indicators and of our forecasts for them will be at least quarterly and we will extend this notion to other city data sets that we are intent on coupling together. The comprehensiveness of the indicators and the underlying data streams will be continuously improved.

3.3.2 Sensing, networking and the impact of new social media

We do not consider that we will develop our own sensors for we are more concerned about mining available data and adding value to its interpretation. Data from the usual social media sites have already been explored in great depth but largely in terms of its immediacy. We see the smart cities dimension of this task as involving how we might interpret the content of such media to very specific issues pertaining to the way cities work and the way citizens interact with their planning. In *FuturICT* there is a strong focus on networks and in this area, we will exploit and link to other network science initiatives and begin to explore how new online real time data sets can be mined so that various network and flows can be extracted from this data and used to provide a deeper understanding of how communities, markets, government and businesses relate to one another. We see this area as extracting social and economic data on which we have little information and understanding so far. Community detection in a liquid world is one obvious outcome from such research.

In terms of starting to extract social networks from travel data, two data streams are missing to understand the dynamics of transport movements: on the one hand, we are lacking the observations and the understanding of the social network geography of the Europeans and how their movements and their resources, and infectious diseases are channelled through them; on the other hand, there is no easy point-of-access to costs of long-distance travel, as all operators use price discrimination to maintain their margins. We will redress these issues in the programme. For example, first, we propose to establish a set of social network surveys in 10–12 European countries/regions, which cover the range of the current economic development and history of migration. In this initial survey the samples of 750 persons will be drawn in selected regions (NUTS 2-level) to be economical in the face-to-face survey administration. The expectation is that further research groups will join the initial effort over the years. The project will make some support for these teams available (data archiving; staff support; data analysis). Second, the long distance travel market (air, rail, ferries, gasoline prices) is characterised by strong price discrimination and the time of booking, season, service levels/class and origin/destination pair. We will set up an automated web-based observatory which will sample the prices systematically across the continent for European and Intercontinental journeys for a range of booking conditions (time to departure, class, length of stay) employing a number of servers simulating different customers by location. The dataset is enriched with suitable data on booking levels, school holidays, legal holidays, tourism flows, etc. The data set will be archived continuously and made available in regular intervals as a counterpart to the official 10 percent FAA ticket sample available for the US airline market.

3.3.3 Modelling network performance, mobility and transport behaviour

We will explore a broad range of theoretical developments to better understand traffic and transportation network performance, from the development of new theories to explain traffic breakdown, car following, and traffic kinetics, to the development of new route choice mechanisms, cooperative game behaviour under network uncertainty, and dynamic models for travel activity generation.

- large scale integrated travel behaviour and logistics simulation models with open-time horizons and explicit learning mechanisms at the agent-levels as tools for policy analysis
- simplified representations of the travel demand and logistics systems for integrated models across the infrastructure systems (energy, communications, water, long-haul-logistics)

- improved network and schedule design optimisation for resilient operation
- fast network design optimisation for basic built infrastructures (roads, railways, canals, pipelines)
- data fusion and data aggregation processes to allow continuous system performance
- new performance metrics for different users that combine efficiency reliability and equity.

We will also build the tools for crisis preparation in personal transport. The possible crises are manifold: e.g. a volcano eruption disrupting air travel across Europe; major population movements after a major chemical accident; shutdown of the railway services during a developing epidemic. The authorities and firms need models describing the current situation and its underlying interactions and dependencies reliably with theoretically sound models. These tools will be the basis for detailed models of disease spreading.

3.3.4 Modelling urban land use and transport

Several groups in *FuturICT* are working with land use transportation models of different kinds ranging from conventional social physics-urban economic style models to cellular automata models of urban development and agent-based models of spatial behaviour which extend to new directions in transportation modelling. We plan to use and extend the agent-based micro-simulation MATSim which provides a basis for extensive model implementation that links travel behaviour, land use, mobility issues and social networks. We want to extend it towards a multi-day timeframe to match the choices of long-distance travel. We have to speed up the already fast model implementation to enable a simulation of the 108 agents in a network of 107 to 108 links/services and 107 destinations in reasonable time. We will draw on the tools available for the GNU-public license framework at `www.matsim.org`.

To support the tool we will model the impacts of the social network geographies of Europeans: based on the new data collected, the project will model the interactions between the frequencies of contacts across all modes and the political, economic and transport performance across Europe and where appropriate the world. We will also extend these models to build tools for crisis preparation in logistics: The market for logistics services is more complex than the market for passenger transport due to lengths of the supply chains and the larger number of decision takers and actors involved. Based on on-going work to expand MATSim with the suitable data structures and models, we need to increase the computational speed to match the new scale of the implementation. We will implement the tools for Europe so that related projects in the *FuturICT* and elsewhere can explore scenarios of interest to them. We will consolidate the necessary data, choice models, generate the agent population, establish the networks, calibration and validation data for the first implementation and then the necessary biannual updates and five-years major updates. The contacts with the users and their experiences and results will be integrated on an on-going basis: the living model will continuously adapt and learn.

Other classes of more aggregate land use transportation models (such as the Simulacra suite of models (`www.simulacra.blogs.casa.ucl.ac.uk`) will be extended and linked to more disaggregate physical models. These simulate aggregate dynamics of development and we envisage that models of this kind can be linked to models of the MATSim variety. This style of model will be used to demonstrate how new planning and decision support systems can be fashioned for planning the smart city, and our focus will be on developing a wide portfolio of modelling tools that users and participants of various kinds can use in their dialogues.

3.3.5 Modelling urban transactional activities in labour, housing and transportation markets

We plan a new focus on modelling the market transactions that determine the way land, property and labour is developed, purchased, allocated and rewarded in the urban environment. We know little or nothing about how urban systems respond to macro-financial crisis (which will be explored elsewhere in *FuturICT*), and the booms and busts that plague national and international economies certainly play out at the city level. Indeed it might be argued that the origins of these crises in one sense lie at the heart of how cities function, or dysfunction. We will develop a series of housing market models built using agent-based technologies, synthesising various databases which relate to financial transactions in these markets, ways in which these markets clear and/or jam, how access to more global capital dictates the spatial behaviours of these markets, and how potential purchases and developers are affected by access to resources. And we will progress the design of smart cities with different kinds of transportation systems that both optimise efficiency and equity in mobility and access to opportunities.

In similar ways, we will explore how local labour markets behave with respect to the supply of businesses, the role of government, environmental quality, and the extent to which such labour relates to innovation. Migration is a key component linking housing to labour markets, and we will examine the impact of aging, the role of statutory instruments and regulations, and the role of capital provision from financial services in determining how such markets work. We envisage that the models that we will build will be closely linked to the transport and urban land use models that are being developed elsewhere but in the spirit of understanding that requires multiple models being used in parallel to explore complex systems such as cities, we envisage a fair degree of parallel and also counter modelling.

3.3.6 Decision support as urban intelligence: Real time modelling and participation in policy making

Templates and structures for decision support systems that involve the wide portfolio of models and tools that this project will focus on planning smart cities for the future, are in their infancy. Most attempts at building intelligence functions for city governments have in the past been top down and have been concerned with professional usage. We are suggesting that cities are so complex, that this kind of intelligence function is only one of many that needs to be coordinated in planning the city. Our work on integrating data bases and models will support the development of integrated intelligence but new ways of visualising data and urban problems, new ways of using tools which inform and predict the impacts of future scenarios, and new ways in which citizens might provide useful and focused advice all need to be refashioned into integrated systems that operate continuously and robustly. We will develop these environments for several different but related types of planning problem, at different scales and over different time scales, resolving issues of modelling that are multi-scale and multi-temporal following rudimentary ideas of integrated modelling of which there are a number of simple applications so far.

3.3.7 City governance structures for the smart city

What is required not only for cities but for government and governance at every level are new frameworks that take account of the extensive access to information that contemporary citizenship now makes possible. This is much wider than our project here and it clearly relates to issues of participation in the sixth project (3.3.6) above. This and the previous project are syntheses as Fig. 12 above shows pulling together much of the modelling capabilities that we intend which in turn are built on the integrated data systems that we foresee will be developed. Standards for data and model development, appropriate interfaces, security of who is able or not to access the material online, questions of confidentiality, IPR, privacy and so on will all feature under this area of the project. Just as the city is changing due to ICT and so are our models of its functioning, these kinds of institutions must also embody a degree of flexibility that is quite different from existing organisations that are tasked to deal with the future of our cities. In this, ICT will be central but so will issues of responsibility, openness, transparency, access to public data and the regulations that extra national government agencies may impose on what and how and where and why citizens are able to influence the governance of their cities.

3.4 Demonstrators

To date, we have not detailed exactly how we might demonstrate our research but it will be highly applied, developed and focused on real, live applications to cities that are manifestly planning to be smart and those that are becoming smart in a less self conscious manner. We intend to select a series of places and sectors that we consider typical of these types: for example, new planned smart cities, large cities that are clearly becoming smart such as our own London, Turin, Paris, Rome, Zurich and Tel-Aviv, for example, cities that have particular problems of economic decline whose future might be assured by explicit development of a smart city ethos, cities that have specific ethnic problems and high rates of migration, aging cities and so on. Our portfolio of tools for decision support will be developed on the basis of these demonstrators and we will ensure that these are linked to key initiatives on the smart city being developed by the worlds major ICT companies.

4 Expected paradigm shifts

Thomas Kuhn [29] who introduced the term paradigm in 1962, defined them as universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of researchers, in short, a world view that dominates science for a period of time during which that world view is extended. A paradigm shift occurs when this world view becomes eroded, when anomalies and inconsistencies mount up to such an extent that researchers can no longer work within the framework. The history of science has shown that the turbulence that sets in can lead to a paradigm shift that takes place over a very short period of time, years or decades rather than centuries. The development of modern computation could be seen as generating such a shift but in one sense, this is more because the very systems that we are focused on here have also been evolving just as our knowledge of how to understand them is evolving. This sea change we have alluded to previously is thus more than a paradigm shift for it may represent a once-for-all transition from a world based in energy and materials to one based on information. Nevertheless, it is quite clear from what we are suggesting that many of our traditional approaches to cities are no longer relevant, the planning systems that we currently work with are not fit-for-purpose, and thus the shifts that we will initiate here are paradigm changes of a kind that are unprecedented. In the more modest context of smart cities, we will present some of these here.

The first major shift which is the most obvious is the development of information infrastructure that underpins the city through distributed computing and networks available to everyone with devices that can access such infrastructure, Whether or not they can access this ICT depends on questions of governance and security but the fact that such infrastructure is now available, requires coordination so that services can be delivered most effectively. As an obvious spin-off from such service delivery, the data that is routinely collected is now being used to make cities smarter over different spatial and temporal scales. This is an unprecedented time for in the past, such data has not been available routinely and the fact that it can now be, yields opportunities for solving human problems that we have never had before. At the same time, cities will never be entirely automated and in this transition period, and maybe forever, we need to grapple with existing non-automated, non-digital technologies and enable these to merge and co-exist in an integrated fashion with the digital.

We are also realising that for the first time that we stand at a threshold in devising a new science of human behaviour and in our own domain, this will be a science of spatial behaviour. Routine data sensed in real time is yielding big data that will require new tools for their usage and analysis, and new methods of data mining that are applicable to individual observations are required. This will move the field of data analysis and statistical method forward in great leaps, if we can fashion new ways of dealing with millions of observations. Our traditional statistics aim to extract macroscopic average information using modest population numbers in the thousands, not the millions or the billions even, and these no longer work very well in this new world. We need new statistical methods to handle these data that are also able to recover generic information on the microscopic dynamics (usually not known for social systems) in a top-down approach. Visualisation is all important in this context and visualisation in the spatial domain leads the way. In developing new models of human behaviours, we need a focus on questions of location and mobility and in this sense the role of social media and web access is crucial [44].

These developments also impress upon us a new view of the spectrum from the local to the global, from bottom up to top down. Complexity theory stresses the evolution and development of systems from the bottom up but it is clear that systems reflect a mix of interventions across all scales notwithstanding these are all predicated by individual actions. In terms of localism, in cities this new science is finding its way down to the urban design, neighbourhood and building scales merging with models of how buildings function. It is here that behaviour hits physicality and in this domain of smart cities, we have opportunities to see how behaviour is conditioned by physical issues just as these same physical issues can be moulded by human behaviour. This is a paradigm shift in that it has not been possible to explore these issues hitherto.

5 The proposed research strategy

5.1 Relevant disciplines and fields

We have outlined seven key areas on which our research will be focused in Sect. 3.3 above and here we will set this in context with respect to the strategy we will adopt. Cities represent a focus for many different disciplines. Indeed in the human and social sciences, everything we study takes place in cities but our focus here is on cities as spatial systems, with our concern being directed at their spatial and physical organisation. Most of our proposals in understanding the smart city and how to make it smarter revolve around spatial issues and although we are conscious that non-spatial and a-spatial issues are relevant, these issues represent the boundary between what we will research and those of other groups within *FuturICT*.

The complexity sciences represent our focus in *FuturICT* but these sciences are linked to many different disciplines and professional fields that have the city as their concern. In particular, urban planning and transport planning are distinct areas that are central to this research while computer science is key to the development of large databases, networks, data mining and human computer interaction. Cities can only be studied in an interdisciplinary context and our perspective involves developing a social physics of cities that is consistent with treating their structure and evolution as complex systems [6,36]. In this sense, our quest involves ideas drawn from the mathematical physical and natural sciences but set within a social context that blends the qualitative with the quantitative.

5.2 Key references and patents

We intend to work with several companies who are building the smart city with respect to its hardware and software infrastructure. It is unlikely that we will develop hardware in any form but there are several areas where we might plan innovative software and database solutions to problems that involve coupling databases together and mining them in real time to extract patterns used to steer and control the way the city functions. These deliverables will be subject to IPR and we consider that some of these products might be patentable. However, much of the software we develop will be in the public domain, available under Creative Commons licenses which city governments and any agency or group acting in a non-profit context will be able to access in a transparent fashion. It is unlikely that we would be in a position to support such developments but arrangements with our partners, such as the developers of MATSim and such like software would be negotiated so that these kinds of software would be widely supported. Some of these systems are already open access.

5.3 Demonstrator outcomes

We provide a number of demonstrators as we indicated above in Sect. 3.4. These will be focused on a) specific problems types, b) specific model types and c) specific cities. In particular we envisage producing demonstrators which will present how the following problems can be articulated:

- Housing booms and busts in large cities, linked to financial crises
- Impacts of changes in energy on urban transportation systems and mobility
- The fracturing of transport networks due to short term problems related to urban conflict, weather and one-off events
- The efficiencies produced by synthesising different urban data sets
- The impact of climate change on cities in Europe, particularly sea level rise and rising temperatures on population location
- The participation of citizens in the development of plans for smart cities of the future focussing on mobility, housing, better design and aesthetics (the city beautiful) and access to opportunities
- The impact of immigration phenomena in a global world.

These are just a sample of possible issues that our science can be used to inform. We will match these against different styles of model and different city contexts developing some of these to the point where individual cities will be able to translate our applications to real contexts, particularly as we envisage that many of the these applications will be developed *in situ*.

5.4 Ethical issues

Because our projects deal with developing new data systems from the ground up, systems that involve sensing individual behaviours and merging these with secondary data sets constructed at an equivalent level, there are major privacy concerns. As we merge different data sets from different sources to produce integrated and coupled systems, there are clear issues of copyright and IPR which will have to be addressed. Much of our work on integrated databases will tackle these issues.

In terms of our focus on involving a large array of groups and citizens in planning the smart city of the future, we will need to take account of what information is accessible to whom and this becomes crucial when such information is available at a fine spatial scale where individuals can be identified. We will draw on an extensive knowledge base about how confidentiality can be assured in the construction of such data systems and their use and access. We will engage with the open data agenda being pursued by national and city government and NGOs and ensure that our research is entirely consistent with these developments.

6 Expected impacts

We will list these as a series of bullet points, which to an extent have been elaborated already in the previous presentation. These apply to all three sections on science, technology and society.

6.1 Impacts on science

We will tackle the essential conundrums of social evolution through the lens of the city. This will involve developing strategies and methodologies that deal with evolving systems that are becoming more complex due to technological innovation and increasing prosperity while at the same time, analysing these issues and anticipating them using the same technologies that are increasing that complexity. This also means the development of a statistical mechanics of cognitive systems.

We will progress the science and art of urban simulation which we believe is strongly rooted in robust theory of how the city functions in space and time as an economic entity and social artefact. This will involve embedding our models in new theories of the contemporary city which are grounded in the new economic geography, urban economics, agent-based conceptions of social and economic systems and new approaches to mobility and communications.

We will develop new methods of integrating spatial and related databases and continue to progress developments in data mining of very large data sets of the orders of terabytes. These will require new developments in neural nets, machine learning and evolutionary computation.

6.2 Impacts on technology and competitiveness

We will generate new ideas about enabling cities to realise their potential by getting smarter. Smart cities are incubators for ever smarter ideas and we will demonstrate this model on several exemplar cities with whom we will develop our science *in situ*.

Smart cities are competitive cities and we will identify ways in which cities that stand on a spectrum of the smartness scale can respond to new initiatives and increase their competitive advantage. We believe that smart cities and systems of smart cities need to embody this sense of competition in an interactive evolutionary context so that no city falls too far behind or progresses too far ahead.

We will develop new web-based interactive contexts that will enable a wider range of citizen activist and groups in the understanding and design of the city and community in which they have an interest and stake.

6.3 Impacts on society

Smart cities are equitable cities. We will develop infrastructures that are accessible to a wider range of interests and groups with differing levels of expertise and activism so that all are involved. Our focus on efficiency balanced against equity is central to this vision.

The web based interactive systems which we consider to be basic to the kind of citizen science that we assume should be normal in the smart city will enable fairness to be progressed and balanced against competition.

We believe that many of the methods that we will develop will be based on notions about how groups compete and cooperate and we consider that the sort of infrastructure, expertise and data that will characterise the smart city will enable equity to be easily established and such cities to improve the quality of urban life.

A more detailed research plan will follow but as yet our quest is simply to define the context, state the key problems and imply some sense of what solutions in terms of our research might focus on. We consider that this paper represents a basis for further discussion to argue the point that new technologies have both disruptive and synergetic effects, particularly on forms of social organisation that are required for future forms of governance and community action as well as business. A sense of what this research promises is available from the many contributions arrayed on the *FuturICT* web site: (`http://www.futurict.eu/`).

The publication of this work was partially supported by the European Union's Seventh Framework Programme (FP7/2007–2013) under grant agreement no.284709, a Coordination and Support Action in the Information and Communication Technologies activity area (*FuturICT* FET Flagship Pilot Project).

References

1. A. Aurigi. *Making the Digital City: The Early Shaping of Urban Internet Space* (Ashgate Publishing Company, Farnborough, UK, 2005)
2. K.W. Axhausen, Env. Planning B: Planning Design **35**, 981 (2008)
3. M. Batty, Technology highs. *The Guardian* **29** (1989)
4. M. Batty, Env. Planning B: Planning Design **17**, 247 (1990)
5. M. Batty, International Planning Studies **2**, 155 (1997)
6. M. Batty, *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals* (MIT Press, Cambridge, MA, 2005)
7. M. Batty, Env. Planning B: Planning Design **37**, 577 (2010)
8. A. Bazzani, B. Giorgini, L. Giovannini, R. Gallotti, S. Rambald, MIPRO2011, 1615 (2011)
9. A. Bazzani, B. Giorgini, S. Rambaldi, *Encyclopaedia of Complexity and Systems Science*, edited by R. Meyers **3**, 9411 (2009)
10. A. Bazzani, B. Giorgini, S. Rambaldi, R. Gallotti, L. Giovanninii, J. Stat. Mechanics: Theory Exper. **5**, P05001 (2010)
11. E. Beinhocker, *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics* (Harvard Business School Press, Cambridge, MA, 2008)

12. A. Caragliu, C. Del Bo, P. Nijkamp, *Smart Cities in Europe*. VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics, Research Memoranda 0048, Amsterdam, The Netherlands, 2009
13. W.H. Dutton, J.G. Blumler, K.L. Kraemer, *Wired Cities: Shaping the Future of Communications* (G. K. Hall, New York, 1987)
14. N. Eagle, A. Pentland, Pers. Ubiquitous Comp. **10**, 255 (2006)
15. N. Eagle, A. Pentland, Behavior. Ecol. Sociobiol. **63**, 1057 (2009)
16. R.K. Brail (Editor), *Planning Support Systems for Cities and Regions* (Lincoln Institute of Land Policy, Cambridge, MA, 2008)
17. E.O. Fernandes, et al., Smart Cities Initiative: How to Foster a Quick Transition towards Local Sustainable Energy Systems (`http://think.eui.eu/`) (2011)
18. L.M. Correia, et al., Smart Cities Applications and Requirements. White Paper, 2011-05-20, Net!Works European Technology Platform Expert Working Group (`http://www.networks-etp.eu/`) (2011)
19. R. Giffinger, et al., *Smart Cities Ranking of European Medium-Sized Cities*. Centre of Regional Science, Vienna University of Technology, Vienna, Austria (`http://www.smart-cities.eu/`) (2007)
20. P. Expert, T. Evans, V. Blondel, R. Lambiotte, Proc. Natl. Acad. Sci. USA **108**, 7663 (2011)
21. Edited by F. Giannotti, D. Pedreschi, *Mobility, Data Mining and Privacy* (Springer, Berlin, 2008)
22. F. Giannotti, M. Nanni, D. Pedreschi, F. Pinelli, C. Renso, S. Rinzivillo, R. Trasarti. Unveiling the complexity of human mobility by querying and mining massive trajectory data. *The VLDB Journal*, DOI: 10.1007/s00778–011–0244–8, 2011
23. F. Giannotti, M. Nanni, F. Pinelli, D. Pedreschi. Trajectory pattern mining. *ACM SIGKDD 2007, Proceedings, International Conference on Knowledge Discovery and Data Mining*, 330 (2007)
24. M.C. Gonzalez, C.A. Hidalgo, A-L. Barabasi, Nature **453**, 779 (2008)
25. M. Granovetter, Amer. J. Sociol. **78**, 1360 (1973)
26. C. Harrison, B. Eckman, R. Hamilton, P. Hartswick, J. Kalagnanam, J. Paraszczak, P. Williams, IBM J. Res. Develop. **54**, 1 (2010)
27. IFF. 2020 Forecast: The Future of Cities, Information, and Inclusion: A Planet of Civic Laboratories, Technology Horizons Program, Palo Alto, CA 94301 (`http://www.iftf.org/`) (2011)
28. J.E. Innes, D.E. Booher, *Planning with Complexity: An Introduction to Collaborative Rationality for Public Policy* (Routledge, London, 2010)
29. T.S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago, IL., 1962)
30. R. Lambiotte, V. Blondel, C. De Kerchove, E. Huens, C. Prieur, Z. Smoreda, P. Van Dooren, Physica A: Stat. Mech. Applic. **387**, 5317 (2008)
31. J. Larsen, J. Urry, K.W. Axhausen, *Mobilities, Networks, Geographies* (Ashgate Publishing Company, Farnborough, UK, 2002)
32. M. McLuhan, *Understanding Media: The Extensions of Man* (McGraw Hill, New York, 1964)
33. M. Nielsen, *Reinventing Discovery: The New Era of Networked Science* (Princeton University Press, Princeton, NJ, 2011)
34. A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, C.S. Mascolo. A tale of many cities: Universal patterns in human urban mobility [`arXiv:1108.5355`] (2011)
35. F. Pagliara, M. de Bok, D. Simmonds, A. Wilson (eds.), *Employment Location in Cities and Regions: Models and Applications* (Springer, Heidelberg, DE, 2013)
36. Y. Portugali, *Complexity, Cognition and the City* (Springer, Heidelberg, DE, 2011)
37. A. Pozdnoukhov, C. Kaiser, Proceedings of the Location-Based Social Networks Workshop, ACM SIGSPATIAL GIS'2011 **19** (2011)
38. J. Reades, F. Calabrese, C. Ratti, Environment and Planning B: Planning and Design **36**, 824 (2009)

39. H.J. Rittel, Panel on Policy Sciences, Amer. Association Adv. Sci. **4**, 155 (1969)
40. S. Sassen, *Talking Back to Your Intelligent City* (`http://whatmatters.mckinseydigital.com/cities/talking-back-to-your-intelligent-city`) (2011)
41. S. Scellato, A. Noulas, R. Lambiotte, C. Mascolo. Socio-spatial properties of online location-based social networks. *Proceedings of Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*, Barcelona, Spain (2011), p. 5
42. S. Schnfelder, K.W. Axhausen, *Urban Rhythms and Travel Behaviour: Spatial and Temporal Phenomena of Daily Travel* (Ashgate Publishing Company, Farnborough, UK, 2010)
43. IBM Global Business Services, A Vision of Smarter Cities: How Cities Can Lead the Way into a Prosperous and Sustainable Future, (IBM Institute for Business Value, Somers, NY) (`http://public.dhe.ibm.com/common/ssi/ecm/en/gbe03227usen/GBE03227USEN.PDF`) (2009)
44. C. Song, Z. Qu, N. Blumm, A-L. Barabsi, Science **327**, 1018 (2010)
45. F. Walsh, A. Pozdnoukhov, Pervasive Urban Applications workshop at PERVASIVE (2011)
46. D. Wang, D. Pedreschi, C. Song, F. Giannotti, A.-L. Barabasi, Human mobility, social ties, and link prediction. Proceedings, International Conference on Knowledge Discovery and Data Mining (`http://users.cis.fiu.edu/~lzhen001/activities/KDD2011Program/docs/p1100.pdf`) (2011)

Journal of Ethnic and Migration Studies

ISSN: 1369-183X (Print) 1469-9451 (Online) Journal homepage: http://www.tandfonline.com/loi/cjms20

Super-diverse migrants—similar trajectories? Ghanaian entrepreneurship in the Netherlands seen from a Mixed Embeddedness perspective

Robert C. Kloosterman, Katja Rusinovic & David Yeboah

JOURNAL OF ETHNIC AND MIGRATION STUDIES, 2016
VOL. 42, NO. 6, 913–932
http://dx.doi.org/10.1080/1369183X.2015.1126091

Routledge
Taylor & Francis Group

Super-diverse migrants—similar trajectories? Ghanaian entrepreneurship in the Netherlands seen from a Mixed Embeddedness perspective

Robert C. Kloosterman[a], Katja Rusinovic[b] and David Yeboah[c]

[a]Department Geography, Planning and International Development Studies, University of Amsterdam, Amsterdam, The Netherlands; [b]Department of Sociology, University of Amsterdam, Amsterdam, The Netherlands; [c]Department of Geography, University of Amsterdam, Den Haag, The Netherlands

ABSTRACT
The Ghanaian population in the Netherlands is relatively well-endowed in terms of human capital. In addition, a large number of them came when deindustrialisation had run its course and the Dutch economy, the service sector in particular, started growing again after 1985. On the basis of the Mixed Embeddedness model, we expected that the combination of, on average, higher levels of human capital and the transformation of the (urban) economy, would lead to rather different patterns of entrepreneurship when compared to their predecessors who came as guest workers. We explored this issue using interviews with 84 Ghanaian entrepreneurs in the Netherlands. Our data only partly corroborated our hypotheses. Notwithstanding, the higher levels of human capital and the shifts in the urban economies, a significant number of Ghanaian entrepreneurs still end up in the lower echelons of the opportunity structure.

KEYWORDS
Migrant entrepreneurship; Mixed Embeddedness; super-diversity; urban transformation; Ghanaians

Introduction

A small shop or cafe at the lower end of the market in an urban neighbourhood in transition run by either a former guest-worker or someone from a former colony. That is the archetype that cannot just be found in much of the academic literature on first-generation migrant entrepreneurship, but which has also been the image that has dominated much policy-making in the past decades (Kloosterman and Rath 2003; Panayiotopoulos 2006). Yet, new groups of migrants from less-developed countries have arrived and (urban)economies have undergone structural transformations and the question then becomes whether these newcomers have been able to exploit the opportunities offered by expanding urban economies (Kloosterman 2010; Ram et al. 2013; Jones et al. 2014). Below, we will explore this issue empirically by looking at the case of Ghanaian entrepreneurs in Dutch cities as they represent a group of newcomers and starting and running businesses in post-industrial urban environments.

These Ghanaian entrepreneurs can be seen as part of a wider trend in migration flows after 1980 which have significantly altered the composition of the migrant populations in

CONTACT Robert C. Kloosterman R.C.Kloosterman@uva.nl

many countries and, hence, the potential supply of migrant entrepreneurs. This so-called 'new' migration, consisting of a wide variety of political refugees, asylum seekers, and 'economic' migrants from a large number of both developed and less-developed countries, is much more diverse than its predecessor which was mainly about migrants from former colonies and 'guest-workers' from a limited number of countries (OECD 2001; for the Netherlands: see Obdeijn and Schrover 2008; Kloosterman 2014). It is not just that the number countries of origin has multiplied, but also, as Steven Vertovec (2007) has argued, that migrants have become much more diverse, even—"super-diverse"—in terms of key characteristics such as legal status, language, religion, location, transnational orientation, and, crucial for starting and running a business, human and social capital.

At about the same time, a large number of cities in the developed world experienced an urban renaissance when employment and population started growing again from the late 1980s onwards (Cheshire 2006; Clark 2009; Scott 2008, 2012). Migrants arriving then thus faced a rather different environment than their counterparts before that who had deal with shrinking urban economies. Deindustrialisation had more or less run its course and intensifying global competition had forced many firms in advanced economies to focus on types of production which require a high input of specialised kinds of knowledge. Notably, high-end producer and consumer services, education, healthcare, R&D activities together with creative and cultural industries have expanded rapidly (Scott 2012). Many of these economic activities have an outspoken urban orientation as they benefit significantly from agglomeration economies. Spatial concentration lowers the costs of (frequent) transactions, the search costs for specialised labour and those of using a dedicated infrastructure (physical and social) and, in addition, fosters innovation through face-to-face exchanges of knowledge. A very different type of urban economy, then, has emerged in the past few decades. However, as Allen Scott has shown in his penetrating analysis of the new urban economy, opportunities for (new)firms are certainly not limited to these high-end or *cognitive-cultural* activities (Scott 2008, 2012). The cognitive-cultural urban economy also generates demand for lower-skilled activities which are mainly to sustain the high-end part. Many of them are *servile activities*—such as dog-walking, cleaning, messenger services, janitors, shop-assisting, waiting (cf. Sassen [1991] 2001)—which are typically low-paid and often precarious. The emerging cognitive-cultural economy, accordingly, offers, in principle, opportunities for entrepreneurs with high levels of human capital as well for those with relatively low levels.

The Ghanaian entrepreneurs—who are, as we shall see below, relatively highly educated, mostly fluent in English, and who came to the Netherlands after 1980—may be seen, in principle, as a *typical* case of how 'new migrants' interact with the evolving post-industrial opportunity structure in advanced cities. Do they indeed depart from the archetypical first-generation migrant entrepreneur from the 1960s and 1970s? Have they started different kinds of businesses? Are Ghanaian entrepreneurs, to put it more succinctly, more oriented towards attractive expanding post-industrial growth sectors and do they, thus represent, in essence, a new kind of migrant business with a more promising outlook?

Methodology

To address these research questions a qualitative methodology was deployed. The study is based on interviews with 84 Ghanaian entrepreneurs in the Netherlands. Information about Ghanaian entrepreneurs and the location of their businesses was acquired with the help of Ghanaian immigrant associations in Amsterdam (RECOGIN and Sikaman), The Hague (Ghanatta), and Rotterdam (Ghanirom). Additional information was gathered through the informal network of one of the researchers who is from Ghanaian descent. He was also able to use his network in The Hague to get in touch with Ghanaian entrepreneurs. Most of the interviews took place in the (larger) cities of Amsterdam (N = 28), The Hague (N = 43), and Rotterdam (N = 4) and a small number of respondents (N = 9) were located in various smaller cities of the Randstad, such as Haarlem and Heemstede.

Most of the interviews (N = 82) were conducted face-to-face at the business premises, the other two interviews were done by e-mail. All interviews were conducted in English and, where necessary, the Ghanaian 'Twi' was used. The interviews were based on a semi-structured questionnaire with 165 questions. On average, the interviews took between one-half and two hours. The interviews with the entrepreneurs were in-depth, probing to uncover the motives for starting a business, hence an intensive research method was applied (see Rusinovic 2006). The comments and answers which could not be written directly on the questionnaire were put in a notebook. References were made to the dates, names, and cities of the respondents to avoid any misidentification.

The entrepreneurs were interviewed between 2005 and 2012. Some of the businesses were revisited as a follow-up in 2011 and 2013 to find out if there had been some important changes in the business activities, business size, business location or otherwise. Finally, in 2014 a complete update of the respondents was conducted. The fieldwork thus covers a period of nine years. This longitudinal aspect enabled us to examine the business success. In this article, a logistic regression analysis (binary logistic) is performed to assess the impact of key variables on the likelihood that the business still existed in 2014. However, given the way the sampling took place, representativeness is an issue and it would be hard to lay strong claims of external validity with respect to the larger population of Ghanaian entrepreneurs. This implies that the regression analysis was not used to generalise our findings. Instead, the logistic regression describes the relationships between some key variables and business success.

Below, we will explore these issues by using the Mixed Embeddedness model to refine the research questions (see Mixed Embeddedness and super-diverse, cognitive-cultural cities Section). We then present a brief description of the sample (see A closer look at the Ghanaian entrepreneurs Section) before embarking on an empirical exploration of the research questions (see Empirical results Section). In the last section, we will delve into the broader implications regarding the relationship between new migration flows and entrepreneurship (see Conclusions Section).

Mixed Embeddedness and super-diverse, cognitive-cultural cities

Entrepreneurs bring with them, obviously, both the resources and the drive to start and run a business. These businesses are anything but self-contained phenomena. Instead, they are part and parcel of a particular environment in the sense that they are aimed at exploiting (perceived) market opportunities within a concrete socio-economic, spatial, and regulatory context. The *Mixed Embeddeness* approach (Kloosterman, Rath, and van der Leun 1999; Kloosterman and Rath 2001; Rath 2001: Rusinovic 2006; Kloosterman

2010; Ram et al. 2013; Jones et al. 2014) aims at explaining patterns of entrepreneurship by systematically linking the supply side of entrepreneurs with their specific set of resources, on the one hand, with the opportunity structure and markets on the other. It partly builds on the interactionist model of entrepreneurship proposed by |Waldinger, Aldrich, and Ward (1990a, 1990b), but, borrowing from comparative political economy (Esping-Andersen 1990, 1999; Hollingsworth and Boyer 1997; Whitley 1999) takes explicitly (national and local) differences in the institutional and regulatory context and its impact on the structure of opportunities and markets for businesses into account. Moreover, in a more recent elaboration, it systematically incorporates variations in human capital on the side of (aspiring) entrepreneurs as well as divergent dynamics in the opportunity structure due to transformation processes in advanced urban economies (Kloosterman 2010). Below, we will use this latter version to analyse the patterns of Ghanaian entrepreneurship in the Netherlands.

In the Mixed Embeddedness model, while assuming that opportunities for businesses are accessible for aspiring entrepreneurs without much financial capital, there are two main dimensions. The first one concerns the dynamics of the opportunities for new businesses. To keep it simple, just two sets of possibilities are distinguished: on the one hand, opportunities which expand because of the underlying structural development of the urban economy, and on the other, opportunities which stagnate or even shrink. The latter may still offer chances for aspiring entrepreneurs if the outflow of the established entrepreneurs exceeds the pace of shrinkage (cf. Waldinger 1996). The second dimension addresses the differences in the level of educational qualifications of the aspiring entrepreneurs or (formal)human capital. Some types of businesses require relatively high levels of human capital in the form of formal educational qualifications beyond secondary of even primary schooling (e.g. consultancy or financial services), whereas others do not (e.g. a grocery or a temping agency). On the basis of this, a two-by-two matrix can be constructed with four distinct sets of opportunities (see Figure 1).

The left-hand side of the matrix comprises the sets of opportunities in structurally stagnating markets. The bottom-left quadrant represents the classic vacancy chain: migrant entrepreneurs replacing established businesses run by, typically, older entrepreneurs of either indigenous or (earlier)migrant origin as an economic counterpart of residential invasion-and-succession processes. These opportunities are characterised by low entry barriers. Easy entry often implies intense competition on price and, hence, low profits. To survive, entrepreneurs often deploy informal economic strategies, for instance, by employing family members off the books. Homogenous (within the migrant group itself) social networks foster the trust that is needed for such informal economic activities on a more permanent basis (cf. Portes and Sensenbrenner 1993; Panayiotopoulos 2006). Given the structural decline of the markets in general and the entrapment of migrant entrepreneurs in particular niches (both spatial and ethnic), chances of becoming successful are slim. The only way out (and up) is by breaking-out to more promising markets (Engelen 2001), but this requires a different set of resources on the part of the entrepreneur: more human and financial and/or heterogeneous social capital to cater to these other markets. Although these opportunities constitute the quintessential breeding grounds for immigrant entrepreneurs, the prospects of hard work and low profits makes that many of them are first and foremost pushed to become an entrepreneur as, for instance, job alternatives are lacking.

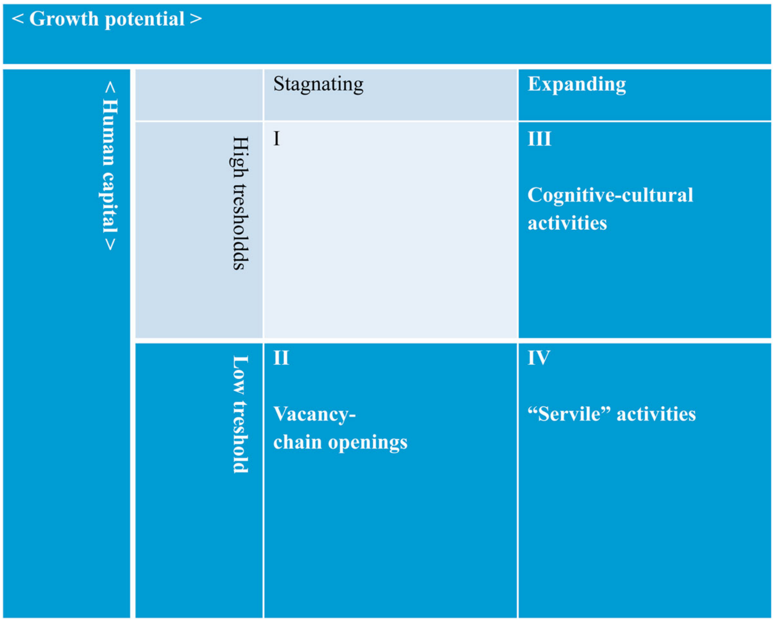


Figure 1. Schematic presentation of the Mixed Embeddedness model.
Source: Kloosterman (2010).

The top-left quadrant denotes the set of opportunities in stagnating markets which require relatively high levels of formal human capital. Migrants with higher vocational training or college degrees have, in principle (meaning if their educational credentials are acknowledged), a wider set of options both in terms of self-employment and employment to choose from than their counterparts who have only secondary schooling or less. It can be expected, therefore, that this will comprise a relatively small number of migrant entrepreneurs. Opting to set up shop in stagnating markets characterised by low levels of profit has to be explained in terms of either a lack of transparency and of information (e.g. due to not belonging to mixed social networks) on more promising opportunities, or a lack of access to these opportunities because of formal (not the right educational qualifications or, more likely, a lack of recognition of foreign credentials for the expanding activities) and informal obstacles (which may include discrimination), or socio-cultural or individual preferences which prioritise certain activities above others. Whereas the former refers to the reasons why highly educated migrants might start a business in sunset activities, it may also be the case that structural shifts could leave migrant entrepreneurs with relatively high educational qualifications out in the cold as markets, which seemed to expand first, start to contract.

Dynamic urban economies surely also comprise structurally expanding markets. In the Mixed Embeddedness model two types are distinguished: post-industrial/low-skilled and post-industrial/high-skilled (Kloosterman 2010). Here, we have renamed them, in line with Scott's (2008, 2012) terminology, cognitive-cultural and servile activities which are two poles of the emerging division of labour in advanced urban economies.

The servile activities, the bottom-right quadrant, are those low-skilled activities which sustain and support the high-end cognitive-cultural activities. According to Scott (2012,

43) '[t]he low-wage service-oriented economy segment of the new economy is focused on jobs like housekeeping, child care, health care, food preparation and serving, janitorial work, taxi driving, and home repair … '. The markets behind the opportunities, then, are generally created by outsourcing by firms and households. We expect that, although profit margins (and wages for workers) are typically low, the rising demand and expanding markets make these opportunities more rewarding and, hence, attractive than the stagnating low-skilled counterpart. Self-employment in these activities, the Mixed Embeddedness model assumes, is not so much out of necessity or push as well as based on a more positive motivation or pull. However, to be successful in these markets, migrant entrepreneurs have to have knowledge of the needs and practices of a more mainstream clientele instead of just their own co-ethnics. This implies heterogeneous social capital or, in other words, access to mixed social networks.

This latter also holds true for the top-right quadrant. There, however, besides heterogeneous social capital, entrepreneurs also need relatively high educational qualifications—beyond secondary schooling. The opportunities are located in the technology-intensive sectors, business and financial services, and in the creative or cultural economy (Scott 2012, 41–42). The software specialists from India in Italy (Cucculelli and Morettini 2012), but also Asians in the creative industries in London (Smallbone, Bertotti, and Ekanem 2005) or the Turkish consultants in Rotterdam (Rusinovic 2006) belong to this category. For migrant entrepreneurs with considerable resources regarding human as well as (mixed) social capital, these opportunities are attractive and offer good chances of upward mobility. This implies that entrepreneurs are *pulled* towards these opportunities and, given their options on the labour market, do not start a business out of necessity.

From the Mixed Embeddedness perspective, then, we expect new migrants to be mainly oriented towards the right-hand quadrants representing the expanding markets as they offer less competition and more profitable opportunities. Moreover, we expect that Ghanaians with higher levels of education will be able to exploit opportunities in the top-right quadrant. On the basis of this overarching hypothesis regarding the sorting pattern of Ghanaian entrepreneurs and their expected orientation towards expanding activities, we can also derive hypotheses from the Mixed Embeddedness model regarding the nature of their markets, the composition of their social networks, and their motivation. We expect them to cater to heterogeneous markets, be part of mixed social networks, being *pulled* towards entrepreneurship, and to be more successful in terms of survival rates in the expanding servile and especially cognitive-cultural activities. In the next section, we will explore these conjectures.

A closer look at the Ghanaian entrepreneurs

Ghanaian migration to the Netherlands is a rather recent phenomenon. Most Ghanaians came to the Netherlands in the last two decades of the 20th century. They came because of push factors, such as expulsion of many Ghanaians migrants from Nigeria, severe drought, mismanagement of the Ghanaian economy and political repression (Mazzucato 2006, 2008; Nieswand 2009). They were also pulled to the Netherlands which was seen at that time as an open and tolerant country with, moreover, a booming economy in the 1990s (see Kloosterman and Rath 2003). As a result, the formally registered Ghanaian population in the Netherlands increased from 2515 in 1987 to 21,922 in 2012, an increase of

870% in 25 years (CBS 2012a). This is, according to Mazzucato (2006), probably a serious underestimation as there are also a large number of undocumented migrants from Ghana. Ghanaians are highly concentrated in the larger cities of the western part of the Netherlands with more than half (52%) residing in Amsterdam which is not just the largest Dutch city, but also ethnically very diverse and, moreover, a buoyant urban economy (CBS 2012a; Kloosterman 2014).

Nearly all of our respondents (over 90%) migrated to the Netherlands in the last two decades of the 20th century (see Table 1). Most of the respondents ($N = 63$) migrated directly from Ghana to the Netherlands. The majority of our research population is between the ages of 40 and 59 (83%). About one third of the research population is female. The share of female entrepreneurs within the research population is comparable with that of all female entrepreneurs in the Netherlands (31%) and more specifically with the shares among Surinamese and Antillean entrepreneurs (source Van den Tillaart 2007; Mestres 2010, 31; EIM 2011, 69). However, compared to some other non-western groups, the share of female entrepreneurs among our Ghanaian sample is relatively high with respectively only 16% for Turkish and 13 for Moroccan female entrepreneurs (Van den Tillaart 2007, 82).

If we look at the endowment of human capital among our respondents, we see that about two thirds of the respondents have a high school diploma, finished a higher vocational education or have a university degree (see Table 1). In addition, almost half of the entrepreneurs ($N = 38$) had a relevant business experience prior to their migration. Among these respondents are female entrepreneurs who were gainfully self-employed 'market mammies' in Ghana, selling food, textile was prints, cosmetics and perfumes and many other goods.

A majority of the Ghanaian entrepreneurs of our sample are proficient in English, with 67% showing a good command of the language—a legacy of the colonial era (see Table 1). About one third of the respondents are fluent in Dutch, while more than half indicated to have a good 'spoken' level (i.e. they are able to understand and speak Dutch to a certain level but find it difficult to write). About 15% of the research population has poor Dutch language skills. However, these entrepreneurs are often able to communicate in English, which is spoken by many people in the Netherlands.

Table 1. Key characteristics of the Ghanaian entrepreneurs.

	N (Total = 84)	Percentage
Time of arrival		
Before 1980	1	1.2
1980–1989	44	52.4
1990–1999	34	40.5
2000–2009	5	6.0
Level of education		
Elementary school	25	29.8
High school	34	40.5
Lower vocational education	3	3.6
Higher vocational education	11	13.1
University	10	11.9
Other	1	1.2
Proficiency Dutch		
Poor	13	15.5
Good spoken level	44	52.4
Fluent	27	32.1

However, as the following quote illustrates, the ability to speak English can also lessen the urge to learn Dutch. The quote stems from an interview with Cynthia, who runs a beauty parlour in the centre of The Hague. She migrated to the Netherlands from Ghana in 1995. According to her the ability to speak English, hindered her to learn Dutch:

Q: How is your Dutch language proficiency?
R: Poor
Q: What do you mean by having poor Dutch language skills?
R: I have not made any effort to learn the Dutch language.
Q: How do you communicate with your clients who prefer to speak in Dutch to you?
R: I usually communicate with the non-Ghanaian clients in English and some of them even speak better English than me. Surprisingly, it seems in the Netherlands both the natives and immigrants alike, have a good spoken level of the English language. In a way everybody's ability to speak in English to me has not motivated me to learn the Dutch language.

To conclude, the Ghanaian entrepreneurs we interviewed are, in general, relatively well-educated (more than half of the Moroccans and Turks between 55 and 65 years in the Netherlands in 2011—i.e. the first-generation—only had basic education or even less, CBS 2012b, 87), had relevant business experiences prior to migrating to the Netherlands and most of them have a good command of Dutch and/or English. These are strategic human-capital resources which facilitate starting a business and, subsequently, can contribute to its survival (Sanders and Nee 1996; Unger et al. 2011; see Jacobs 2012). Given, then, the generally speaking, relatively high level of resources of Ghanaian entrepreneurs we expect the following:

(1) First-generation Ghanaian entrepreneurs who identified promising opportunities to start businesses were pulled rather than pushed into self-employment (see Pushed or pulled? Section).
(2) Given their relatively high level of educational qualifications, first-generation Ghanaian entrepreneurs were able to start businesses in the upper right-hand quadrant of the opportunity structure. In other words, we expect them to have started their businesses predominantly in the cognitive-cultural activities (see What kind of opportunity? Section).
(3) To be able to start a business in the cognitive-cultural activities catering to mainstream markets, a more heterogeneous social network is needed (see Social networks and starting a business Section).
(4) First-generation Ghanaian entrepreneurs who were embedded in more heterogeneous social networks are expected to have been able to get access to both formal and informal financial and non-financial resources enabling them to set up businesses in either the post-industrial low-skilled or post-industrial high-skilled quadrants of the Mixed Embeddedness model (see Financial assistance Section).
(5) First-generation Ghanaian entrepreneurs had the right mix of human, financial and social capital and personal characteristics which enabled them to identify promising business opportunities in the Netherlands and were motivated to start and run successful businesses (see Mixed networks? Section).

Below, we will investigate these hypotheses and try to assess whether Ghanaian entrepreneurs have followed a different track of self-employment than their counterparts from Turkey and Morocco who came in the 1960s and 1970s.

Empirical results

Pushed or pulled?

We expected that Ghanaian entrepreneurs given their human capital endowment and their time of arrival in the Netherlands were more likely to be pulled than pushed towards self-employment. As can be read from Table 2, this hypothesis is vindicated by our results. Only about a quarter indicated that they were pushed into self-employment. These entrepreneurs were, typically, unemployed before they became self-employed. The majority of our respondents , however, were pulled towards self-employment. These entrepreneurs said they decided to start their own business because they saw the opportunity to succeed as well as the possibility to become their own bosses. The predominance of pull factors would also imply, following the Mixed Embeddedness model, that most of these entrepreneurs are to be found in expanding markets where chances for making money are more evident than in stagnating sectors.

What kind of opportunity?

The Mixed Embeddedness model distinguishes four categories of opportunities. One of these, the one comprising high-skilled/structurally stagnating opportunities, is rather improbable for sheer economic reasons, so here we have included only the stagnating vacancy-chain opportunities and the expanding cognitive-cultural and servile opportunities. It appears that the Ghanaian migrant entrepreneurs we interviewed are mainly involved in servile activities such as hairdressing, barbering and beauty salons. Still, about one third of the respondents is active in low-skilled, vacancy-chain businesses,

Table 2. Pushed or pulled towards self-employment.

	N (Total = 84)	Percentage
Pushed	22	26.2
Pulled	62	73.8

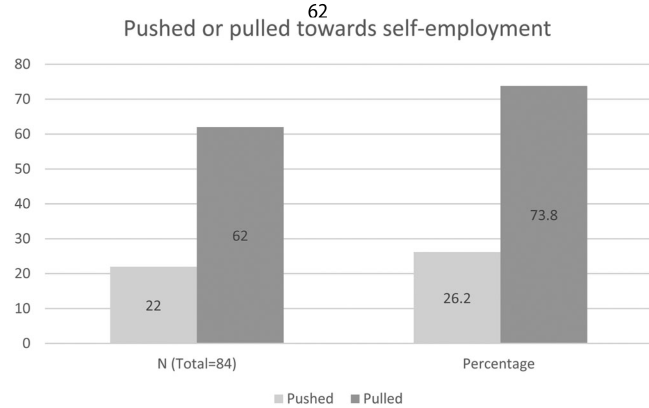


Table 3. Ghanaian entrepreneurs, by types of opportunity and sectors of activity.

Type of opportunity	Sectors of activity	Abs. number	%
Low-skilled stagnating		28	33
	Retail African (food) products	19	23
	Other retail	6	7
	Other	3	4
Servile activities		45	54
	Hair/beauty salon	7	8
	Travel services	8	10
	Catering	3	4
	Distribution of flyers	11	13
	Cleaning	4	5
	Day care	1	1
	Money transfer	2	2
	Removal household goods	2	2
	Other	7	8
C.-C. Activities		11	13
	B2B: among which IT, financial, shipping services	10	12
	Other	1	1
Total		84	100

mostly in retail. Less than 15% of the entrepreneurs are involved in cognitive- cultural activities requiring higher levels of human capital, such as shipping or financial businesses whereas about 25% of our sample has a higher vocational training or university degree. It would seem, then, that Ghanaian entrepreneurs face obstacles in getting access to cognitive-cultural opportunities. However, in accordance with our expectations, almost all entrepreneurs who are involved in the cognitive-cultural activities have a higher vocational education or a university degree ($N = 9$). In the Mixed Embeddedness model, it is assumed that to get access to expanding markets, one should possess mixed social capital, transcending the ethnic group of the entrepreneur, to be able to cater to mainstream markets. Below (see Social networks and starting a business Section), we will explore the composition of social networks of Ghanaian entrepreneurs in starting a business and getting funding (Table 3).

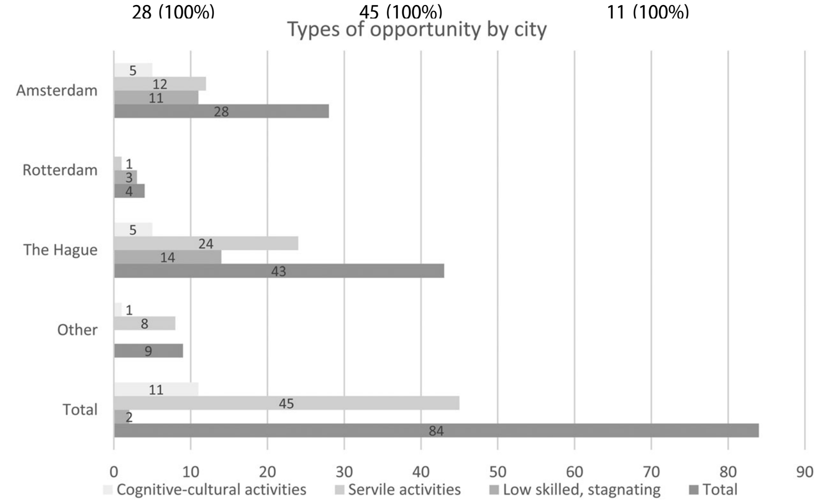
The distribution of the Ghanaian entrepreneurs across cities by types of opportunity (see Table 4) shows that those active in cognitive-cultural economic activities are mainly to be found in Amsterdam and The Hague. In Amsterdam, the fastest growing economy with the most outspoken cognitive-cultural profile of the three cities (Kloosterman 2013), the entrepreneurs are clearly tilted towards both servile and cognitive-cultural activities. The sample is clearly small, but in line with more general observations.

Social networks and starting a business

Starting a business entails more than combining a (nascent) entrepreneur's human capital and financial capital. Obtaining relevant non-financial assistance such as information on markets, workers, suppliers and the regulatory environment is crucial as well and social networks do play an important role in providing such information (see Kloosterman, Rath, and van der Leun 1999). Entrepreneurs, therefore, not only exploit informal and formal social networks for financial assistance but also use both sources to obtain non-financial resources. Immigrant entrepreneurs are typically embedded in informal networks which mainly consist of family, relatives, and (co-ethnic) friends (Light 1972;

Table 4. Types of opportunity by city.

	Types of opportunity			
	Low-skilled stagnating	Servile activities	Cognitive-cultural activities	Total
City				
Amsterdam	11 (39%)	12 (27%)	5 (46%)	28 (33%)
Rotterdam	3 (11%)	1 (2%)	–	4 (5%)
The Hague	14 (50%)	24 (53%)	5 (46%)	43 (51%)
Other	–	8 (18%)	1 (8%)	9 (11%)
Total	28 (100%)	45 (100%)	11 (100%)	84(100%)

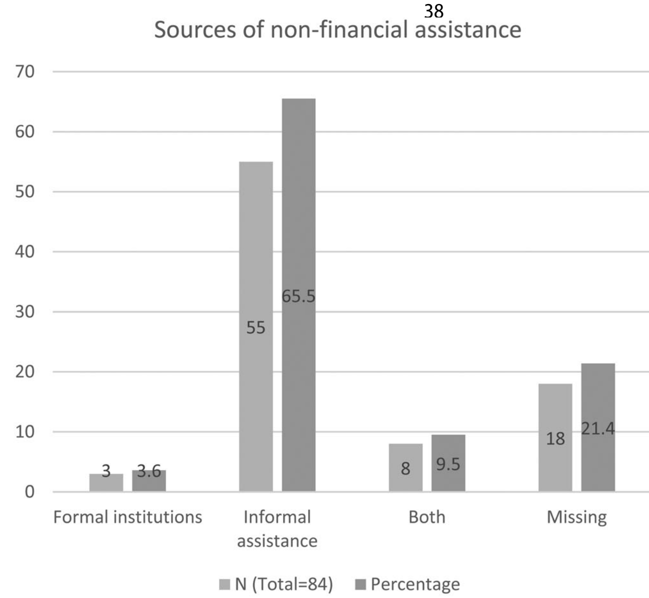


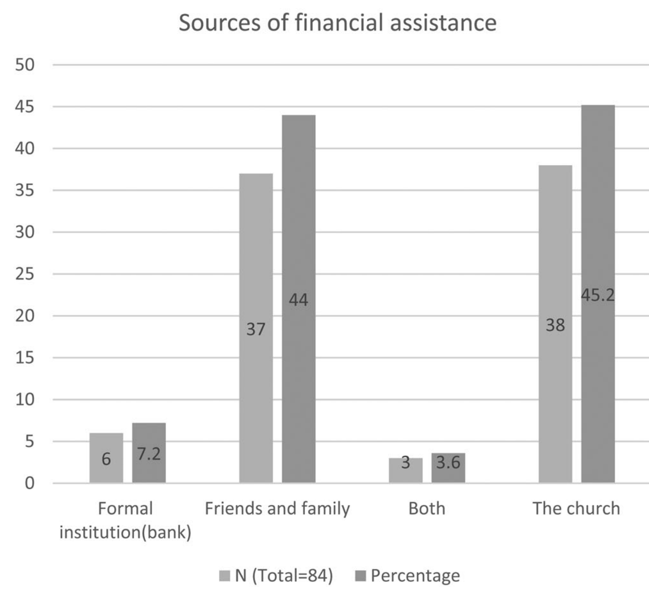
Portes 1998; Panayiotopoulos 2006; Rusinovic 2006; Jones et al. 2012). This also holds for the Ghanaian entrepreneurs (see Table 5). The majority of the respondents ($N = 64$) received assistance in the start-up of their business and in most cases they received help from family members or friends (informal assistance). One of them is Eric, who runs a barbering shop in The Hague. When he started his business, he saw that he could not work alone. He used his 'Asantemankuo' (Ashanti) ethnic association network and within a few weeks, he was able to employ an assistant from that ethnic group.

Churches and associated organisations are of great importance in the Ghanaian diaspora (Mazzucato 2006, 2008; Fumanti and Werbner 2010). They do not just provide spiritual leadership, but also assist their members in getting jobs, housing and in, some instances, to start a business and become self-employed. This can also be observed in our sample in which 38 respondents received assistance from the church in setting up/running their business. The assistance from the church mainly consists of helping to find customers and/or employees. A respondent who lauds the church for its support is James who established a money remittance agency in The Hague in 2000. After obtaining his Master's degree, he decided to become self-employed after all his job applications were turned down. However, James was unsure how the business could start successfully. With an advice from his pastor, James recruited two of the church members who assisted him to promote his company through visits to other churches and the gatherings of hometown associations in The Hague. 'The relentless support from my pastor and the employee-members of the church has been instrumental for smooth start of my business.'

Table 5. Sources of non-financial and financial assistance.

	N (Total=84)	Percentage
Non-financial assistance		
Formal institutions	3	3.6
Informal assistance	55	65.5
Both	8	9.5
Missing	18	21.4
Financial assistance		
Formal institution(bank)	6	7.2
Friend&Family→Informal assistance	37	44.0
Both	3	3.6
The church	38	45.2





Financial assistance

One of the main problems for entrepreneurs in general and particularly for immigrants is to raise capital from formal institutions, such as banks (Granovetter 1995; Rath 2000). Dominic is one of the few respondents who had his loan application approved by a bank. Dominic retails alcoholic and non-alcoholic beverages in Amsterdam 'Bijlmermeer'. The products he sells attract customers from different (ethnic)backgrounds in the Amsterdam Bijlmermeer neighbourhood. His loan request was eventually approved when he presented a business plan to a local bank. The loan was sufficient to start the business. His case, however, is rather exceptional as Table 5 illustrates. The majority receives informal financial assistance from family members, relatives, and friends indicating that the insertion in the more formal social networks of Dutch society is less well-developed. Among these respondents is Bernard, the respondent featured in the opening story of this chapter. Bernard is the co-owner of a travel agency with branches in Amsterdam and The Hague. During the interview he stated that a loan of €10,000 from a close friend as supplementary capital had enabled him to settle his financial obligations in the business partnership he signed with Kwame in Amsterdam. According to Bernard, 'a trusted friend is always willing to help in critical times of need. The loan prevented protected me from shame and ridicule'. Table 5 shows, almost half of the respondents depend on this kind of communal responsibility and support.

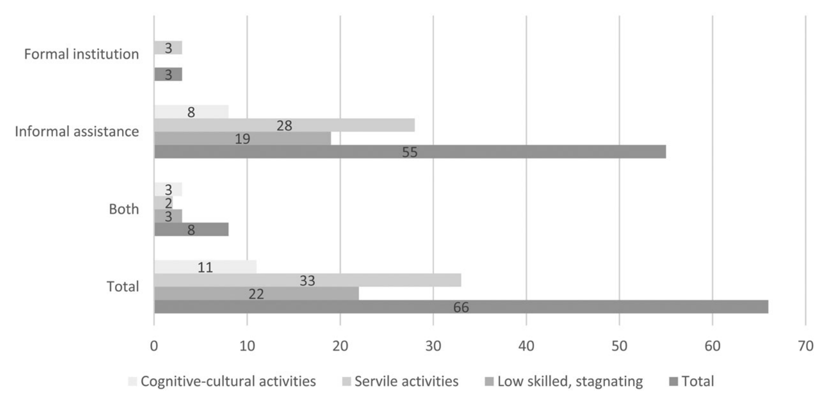
Mixed networks?

The Mixed Embeddedness assumes that to be able to start a business in the expanding activities catering to mainstream markets, a more heterogeneous social network is needed to obtain relevant information about these markets. Table 6 shows the forms of

Table 6. Sources of non-financial and financial assistance by types of opportunity.

	Types of opportunity			
	Low-skilled stagnating	Servile activities	Cognitive-cultural activities	Total
Who helped with the start?				
Formal institution	–	3 (9%)	–	3 (5%)
Informal assistance	19 (86%)	28 (85%)	8 (73%)	55 (83%)
Both	3 (14%)	2 (6%)	3 (27%)	8 (12%)
Total	22 (100%)	33 (100%)	11 (100%)	66(100%)

Sources of non-financial and financial assistance by types of opportunity



assistance by types of opportunity. Informal assistance is clearly the most important and very dominant in low-skilled stagnating or vacancy-chain activities and in servile activities. Concerning the latter, we can also observe forms of formal assistance indicating being embedded in more heterogeneous social networks. This also holds for those activities in cognitive-cultural activities where about one third received assistance from both informal and formal sources.

Although Ghanaian immigrant entrepreneurs have relatively high levels of human capital and many of them have a good command of Dutch and English, they still mainly rely on their co-ethnics –family, relatives and friends as well as churches and related organisations—for setting up and running their business. It seems, then, that their relevant networks are more ethnically homogeneous thereby limiting access to information pertaining to mainstream markets, suppliers and the regulatory environment. This may also impact on the ability to perceive and then exploit more promising activities even for those with high human capital. Table 7 shows that a larger number of the highly educated Ghanaian entrepreneurs (those with higher vocational training or a university degree) of our sample are to be found in vacancy-chain and servile activities than in cognitive-cultural activities. Apparently, there are significant obstacles which hamper access to these activities.

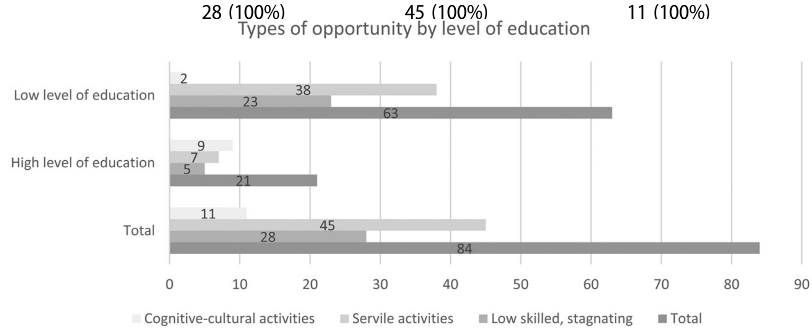
Factors behind business success

Defining what is ‘business success’ is rather problematic and there are various way of defining. We use a rather basic definition of success, namely survival (cf. Schutjens 2013). Businesses which were still in existence in 2014 are defined as successful. To determine whether the business still existed, we traced the entrepreneurs and their businesses through the internet or by telephone.

In 2014, only one third of the businesses ($N = 29$) had survived. To explore the underlying causal relationships a logistic regression (binary logistic) was performed to assess the impact of a number of variables on the likelihood that the business still existed in 2014. In

Table 7. Types of opportunity by level of education.

	Types of opportunity			
	Low-skilled stagnating	Servile activities	Cognitive-cultural activities	Total
Low level of education	23 (82%)	38 (84%)	2 (18%)	63 (75%)
High level of education	5 (18%)	7 (16%)	9 (82%)	21 (25%)
Total	28 (100%)	45 (100%)	11 (100%)	84 (100%)



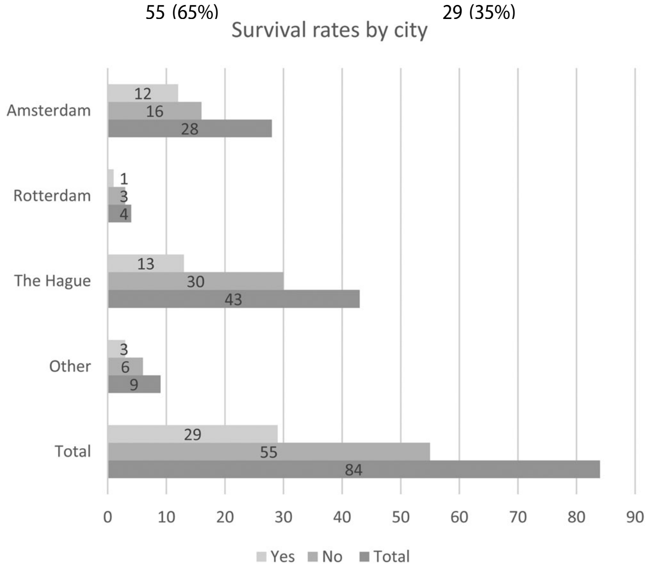
this logistic regression, we used key characteristics of the entrepreneurs as an independent variables: age, year of arrival in the Netherlands, level of education, proficiency in Dutch, motivation to start a business, social capital (informal or formal assistance in starting their business), and their business activities. The dependent variable was whether or not the respondent still existed in 2014.

The logistic regression shows that respondents who are less than 40 years old have a higher chance to survive compared to older respondents (odds ratio 1.875). Also, the duration of stay in the Netherlands influences the chances to survive: entrepreneurs who have been longer in the Netherlands display a higher chance to survive (odds ratio .519). In addition, a positive motivation to start a business ('pull') also increases the chances to survive, compared to entrepreneurs who are pushed into entrepreneurship. If we include proficiency in Dutch, the analysis shows that entrepreneurs who have a good command of the Dutch language have a higher chance to survive, as well as higher educated entrepreneurs (university or higher vocational education) compared to lower educated entrepreneurs (significant, 5% reliability, odds ratio .277). These results are an indication of the importance of human capital for survival.

With regard to the importance of social capital, the logistic regression shows that entrepreneurs who receive both informal as well as formal assistance in the start-up phase have the highest chance to survive in comparison (significant, 1% reliability) to entrepreneurs who receive formal or informal assistance. The findings indicate that in addition to human capital, social capital also influences the chances to survive.

Table 8. Survival rates by city.

	Business still exists?		
City	No	Yes	Total
Amsterdam	16 (57%)	12 (43%)	28 (100%)
Rotterdam	3 (75%)	1 (25%)	4 (100%)
The Hague	30 (70%)	13 (30%)	43 (100%)
Other	6 (67%)	3 (33%)	9 (100%)
Total	55 (65%)	29 (35%)	84 (100%)



We have also included the business activities in which the entrepreneurs are involved. It appears that entrepreneurs who are involved in cognitive-cultural activities have a higher chance to survive compared to entrepreneurs in servile activities (significant, 10% reliability) which corroborates the predictions based on the Mixed Embeddedness model. The smallest chance to survive is to found among entrepreneurs involved in low-skilled activities.

The Mixed Embeddedness model emphasises the role of the wider economic context. Table 8 shows the survival rates broken by city. As could be expected, chances for survival are relatively high in Amsterdam with its fast-growing economy and smaller in both The Hague and Rotterdam.

Conclusions

According to Sepulveda, Syrett, and Lyon (2011, 470) 'much research on ethnic enterprise has concentrated on particularly well-established groups', which in the Dutch case would refer to Chinese, Surinamese, Antilleans, Turks, and Moroccans (Beckers and Blumberg 2013). Newcomers, so far, have been given only short shrift in the Netherlands and also elsewhere (Ram and Jones 2008). These new groups may display rather different characteristics compared to the more established ones thereby adding to the notion of 'super-diversity' as proposed by Vertovec (2007). From our perspective, these different characteristics may also translate into different profiles of self-employment of these newcomers (Ram, Theodorakopoulos, and Jones 2008; Jones et al. 2014).

It is, however, not just the composition of the migrant populations which have changed in many cities, but also the urban economy and, hence, the opportunity structure for small businesses (Kloosterman 2010; Sepulveda, Syrett, and Lyon 2011). The leading sectors of this evolving urban economy, the, following Allen Scott (2008; 2012), so-called *cognitive-cultural* activities are dependent on the input of highly skilled labour offer opportunities for small businesses. On the other end of the continuum, we find the *servile* activities which typically require much less human capital (in terms of formal educational qualifications) and which also offer chances for small businesses.

Focusing on the self-employment profiles of Ghanaian migrant entrepreneurs in Dutch cities is one way, then, to explore then how increasing diversity of the immigrant population interacts with the emerging cognitive-cultural economy. The Ghanaian migrant entrepreneurs constitute a pertinent case of newcomers as most of them arrived after 1980 and, moreover, if we look at their resources, we can observe that they are, on average, relatively well-educated, most of them fluent in English and quite a few in Dutch as well, and, in addition, many of them have entrepreneurial experience prior to migration which sets them apart from their earlier counterparts (Rusinovic 2006; Beckers and Blumberg 2013). Most of them settled in the larger cities of the western part of the Netherlands—Amsterdam, Rotterdam and The Hague—which have through thorough processes of de-industrialisation and subsequent expansion of high-end service activities.

We have used the Mixed Embeddedness approach for a first exploration of Ghanaian entrepreneurship in the Netherlands. Given their time of arrival, we expected that a substantial part of our sample would have set up shop in expanding activities. We, moreover, also expected that, given their endowment of formal (and informal) human capital, Ghanaians with a higher vocational training or university degree would be able to enter the more rewarding expanding cognitive-cultural activities. These expectations were only partly vindicated. True enough, we did find that most of our respondents were pulled towards entrepreneurship and also that nearly two third could be found in expanding markets. We also found that more than half of them were to be found in servile activities and only about one in nine of the entrepreneurs that we interviewed were active in cognitive-cultural activities and indeed that they were highly educated. About one third was to be found in vacancy-chain activities. We also observed that only half of the highly educated respondents had been able to enter these markets. It seems, then, that there are certain obstacles hampering the entrance of even highly educated migrants in cognitive-cultural activities.

These findings are in line with what Jones et al. (2014, 2) observed for the UK: '… despite the novelty of their geographical and social backgrounds, the newcomers seem to be engaged in much the same restricted range of low-value businesses as earlier groups' and even 'reproducing the marginality of their forerunners'. To account for this, they pointed to the difficulties of getting money capital to start and run a business, to the difficulties in operating in an alien linguistic and regulatory environment and to outright discrimination and racism.

It may be the case that due to a lack of recognition of educational qualifications certain openings are blocked for aspiring Ghanaian entrepreneurs. It may also be true that discrimination impacts on the sorting process but we have no concrete evidence for that. However, if we compare our findings with those for second-generation migrant entrepreneurs in the Netherlands (Rusinovic 2006; Beckers and Blumberg 2013), we find that the highly educated among these second-generation migrants are more capable of getting access to high-skilled activities. Being educated in the Netherlands, the issue of not acknowledging their credentials does not arise and, in addition, being brought up in the Netherlands, they know the ins and outs of the Dutch regulatory system.

Looking at our data, we noted that Ghanaians are mainly embedded in ethnically homogenous social networks in which Ghanaian churches and their associations play a big role thereby blocking the formation of heterogeneous social capital. According to Burt (2001, 32) 'Social capital is the contextual complement to human capital.' Consequently, if highly educated Ghanaians are not able to get access to more promising opportunities, this may be attributed to how their social capital is structured. Being part of rather closed-off dense social networks with no or just a few (weak)ties to other social networks, will seriously hamper the flow of rich information (Burt 2001; Elfring and Hulsink 2003). Ghanaian entrepreneurs, hence, while benefiting from their dense networks in terms of access to financial capital, may display herd behaviour in starting businesses in less-promising sectors as they lack crucial information on mainstream markets and business opportunities and, notably, also on the regulatory environment (especially the tax system). This might even be the case for highly educated migrants. Arrighetti, Bolzani, and Lasagni (2014) observed that so-called multicultural hybrid firms which deployed "ethnically diverse human resources at the management and personnel levels" stand a greater chance of survival which may indicate that they indeed are able to gather richer information by bridging distinct networks. Our logistic regression analysis indicated that the length of stay is a significant factor in determining the chances for survival of the Ghanaian businesses. This also seems to point in the direction that knowledge of the wider business environment is helpful in running a business.

The sample we used is, as we mentioned, is quite small allowing only a limited scope of generalisation. Our study is first and foremost exploratory, aiming to tease out if the causal relationships anticipated by the Mixed Embeddedness model can actually be traced empirically by looking at a new group of migrants setting up shop in post-industrial urban economies. We were indeed able to vindicate key relationships spelled out in the model. However, the funnelling of a relatively large share of highly educated Ghanaian entrepreneurs towards more marginal activities shows that more subtle social sorting processes are producing an ethnically segmented population self-employed.

Acknowledgements

We would like to thank Sjoerd de Vos for his advice on the statistical part, Tom Elfring for his comments and Miranda Jonker for her help in collecting the data and her text suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Arrighetti, A., D. Bolzani, and A. Lasagni. 2014. "Beyond the Enclave? Break-Outs Into Mainstream Markets and Multicultural Hybridism in Ethnic Firms." *Entrepreneurship & Regional Development* 26 (9–10): 753–777.

Beckers, P., and B.F. Blumberg. 2013. "Immigrant Entrepreneurship on the Move: A Longitudinal Analysis of First-and Second-Generation Immigrant Entrepreneurship in the Netherlands." *Entrepreneurship & Regional Development* 25 (7–8): 654–691.

Burt, R. 2001. "Structural Holes Versus Network Closure as Social Capital." In *Social capital: Theory and Research*, edited by N. Lin, K. S. Cook and R. S. Burt, 31–56. New York: Aldine de Gruyter.

CBS (Central Bureau of Statistics). 2012a. *Allochtonen in Nederland 2012.* Voorburg: Centraal Bureau voor de Statistiek.

CBS (Central Bureau of Statistics). 2012b. *Jaarrapport voor de integratie.* Voorburg: Centraal Bureau voor de Statistiek.

Cheshire, P. C. 2006. "Resurgent Cities, Urban Myths and Policy Hubris: What We Need to Know." *Urban Studies* 43: 1231–1246.

Clark, P. 2009. *European Cities and Towns 400–2000.* Oxford: Oxford University Press.

Cucculelli, M., and G. Morettini. 2012. "Transnational Ties in Technology-Based Sectors: The Case of Indian Software Entrepreneurial Firms in Italy." *International Journal of Entrepreneurship and Small Business* 17 (3): 319–337.

EIM. 2011. *Trendstudie MKB en Ondernemerschap. Ontwikkelingen, vooruitblik en beleidssignalen.* Zoetermeer: Panteia/EIM.

Elfring, T., and W. Hulsink. 2003. "Networks in Entrepreneurship: The Case of High-Technology Firms." *Small Business Economics* 21 (4): 409–422.

Engelen, E. 2001. "Breaking-in and Breaking-Out; A Weberian Approach to Entrepreneurial Opportunities." *Journal of Ethnic and Migration Studies* 27 (2): 203–223.

Esping-Andersen, G. 1990. *Three Worlds of Welfare Capitalism.* Cambridge: Polity Press.

Esping-Andersen, G. 1999. *Social Foundations of Postindustrial Economies.* Oxford: Oxford University Press.

Fumanti, M., and P. Werbner. 2010. "The Moral Economy of the African Diaspora: Citizenship, Networking and Permeable Ethnicity." *African Diaspora* 3 (1): 2–11.

Granovetter, M. 1995. *Getting a Job: A Study of Contacts and Careers.* Chicago: University of Chicago Press.

Hollingsworth, J. R. and R. Boyer. 1997. "Coordination of Economic Actors and Social Systems of Production." In *Contemporary Capitalism: The Embeddedness of Institutions,* edited by J. R. Hollingsworth, and R. Boyer, 1–47. Cambridge: Cambridge University Press.

Jacobs, F. 2012. "The Dynamics of Dutch Immigrant Entrepreneurship. A search for the critical factors for immigrant entrepreneurs in the Dutch urban environment." Master thesis in Human Geography, Universiteit van Amsterdam, Amsterdam.

Jones, T., M. Ram, P. Edwards, A. Kiselinchev, and L. Muchenje. 2012. "New Migrant Enterprise: Novelty or Historical Continuity?" *Urban Studies* 49 (14): 3159–3176.

Jones, T., M. Ram, P. Edwards, A. Kiselinchev, and L. Muchenje. 2014. "Mixed Embeddedness and new Migrant Enterprise in the UK." *Entrepreneurship & Regional Development* (ahead-of-print): 1–21.

Kloosterman, R. C. 2010. "Matching Opportunities with Resources: A Framework for Analysing (Migrant) Entrepreneurship from a Mixed Embeddedness Perspective." *Entrepreneurship and Regional Development* 22 (1): 25–45.

Kloosterman, R. C. 2013. "The Amsterdam Economy and Its Impact on the Labor Market Position of Migrants, 1980–2010." In *Immigration and the New Urban Landscape, New York and Amsterdam,* edited by N. Foner, J. Rath, J. W. Duyvendak, and Rogier van Reekum, 107–121. New York: NYU Press.

Kloosterman, R. C. 2014. "Faces of Migration: Migrants and the Transformation of Amsterdam." In *Migration and London's Growth,* edited by B. Kochan, 127–143. London: LSE London.

Kloosterman, R. C. and J. Rath. 2001. "Immigrant Entrepreneurs in Advanced Economies: Mixed Embeddedness Further Explored." *Journal of Ethnic and Migration Studies* 27 (2): 189–201.

Kloosterman, R. C. and J. Rath. 2003. "Introduction." In *Immigrant Entrepreneurs: Venturing Abroad in the Age of Globalization,* edited by R. C. Kloosterman and J. Rath, 1–16. Oxford: Berg/University of New York Press.

Kloosterman, R. C., J. Rath, and J. P. van der Leun. 1999. "Mixed Embeddedness, Migrant Entrepreneurship and Informal Economic Activities." *International Journal of Urban and Regional Research* 23 (2): 252–266.

Light, I. 1972. *Ethnic Enterprise in America.* Berkeley: University of California Press.

Mazzucato, V. 2006. "Migrant Transnationalism: Two-Way Flows, Changing Institutions and Community Development Between Ghana and the Netherlands." *Economic Sociology_the European Electronic Newsletter 7* (3): 8–16.

Mazzucato, V. 2008. "The Double Engagement: Transnationalism and Integration. Ghanaian Migrants' Lives Between Ghana and the Netherlands." *Journal of Ethnic and Migration Studies* 34 (2): 199–216.

Mestres, J. 2010. "Migrant Entreprneurship in OECD Countries: Magnitude, Contribution, to Employment and Specific Migration Policies." In *Open for Business; Migrant Entrepreneurship in OECD Countries,* edited by OECD, 23–61. Paris: OECD.

Nieswand, B. 2009. "Development and Diaspora: Ghana and Its Migrants." *Sociologus,* 59 (1): 17–31.

Obdeijn, H., and M. Schrover. 2008. *Komen en Gaan; Immigratie en Emigratie in Nederland Vanaf 1550.* Amsterdam: Uitgeverij Bert Bakker.

OECD. 2001. *Trends in International Migration; Continuous Reporting System on Migration Annual Report 2001 Edition.* Paris: OECD.

Orozco, M. and R. Rouse. 2012. "Migrant Hometown Associations and Opportunities for Development." In *The Community Development Reader* (2nd ed.), edited by J. DeFilippis, and S. Saegert, 280–285. New York, Oxon: Routledge.

Panayiotopoulos, P. 2006. *Immigrant Enterprise in Europe and the USA.* London: Routledge.

Portes, A. 1998. "Social Capital: Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 24: 1–24.

Portes, A., and J. Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *The American Journal of Sociology* 98 (6): 1320–1350.

Ram, M., and T. Jones. 2008. "Ethnic-Minority Businesses in the UK: A Review of Research and Policy Developments." *Environment and Planning C: Government & Policy* 26 (2): 352.

Ram, M., T. Jones, P. Edwards, A. Kiselinchev, L. Muchenje, and K. Woldesenbet. 2013. "Engaging with Super-Diversity: New Migrant Businesses and the Research–Policy Nexus." *International Small Business Journal* 31 (4): 337–356.

Ram, M., N. Theodorakopoulos, and T. Jones. 2008. "Forms of Capital, Mixed Embeddedness and Somali Enterprise." *Work, Employment & Society* 22 (3): 427–446.

Rath, J. 2000. *Immigrant businesses: The Economic, Political and Social Environment*. St. Martin's Press in association with Centre for Research in Ethnic Relations University of Warwick.

Rath, J. 2001. "Sewing Up Seven Cities." In *Unravelling the Rag Trade; Immigrant Entrepreneurship in Seven World Cities*, edited by J. Rath, 169–92. Oxford: Berg and University of New York Press.

Rusinovic, K. 2006. *Dynamic Entrepreneurship; First and Second-Generation Immigrant Entrepreneurs in Dutch Cities.* Amsterdam: Amsterdam University Press.

Sanders, J., and V. Nee. 1996. "Immigrant Self Employment: The Family as Social Capital and the Value of Human Capital." *American Sociological Review, 61*: 231–249.

Sassen, S. ([1991] 2001). *The Global City: New York, London, Tokyo.* Princeton: Princeton University Press.

Schutjens, V. 2013. *Etnisch Ondernemerschap: Blijf Kijken*. Oratie Universiteit van Amsterdam http://www.oratiereeks.nl/upload/pdf/PDF-2557weboratie_Schutjens.pdf.

Scott, A. J. 2008. *Social Economy of the Metropolis: Cognitive-Cultural Capitalism and the Global Resurgence of Cities.* Oxford: Oxford University Press.

Scott, A. J. 2012. *A World in Emergence; Cities and Regions in the 21st Century.* Cheltenham: Edward Elgar.

Sepulveda, L., S. Syrett, and F. Lyon. 2011. "Population Superdiversity and new Migrant Enterprise: The Case of London." *Entrepreneurship & Regional Development* 23 (7–8): 469–497.

Smallbone, D., M. Bertotti, and I. Ekanem. 2005. "Diversification in Ethnic Minority Business: The Case of Asians in London's Creative Industries." *Journal of Small Business and Enterprise Development* 12 (1): 41–56.

Unger, J.M., A. Rauch, M. Frese, and N. Rosenbusch. 2011. "Human Capital and Entrepreneurial Success: A Meta-Analytical Review." *Journal of Business Venturing* 26: 341–358.

Van den Tillaart, H. 2007. "Etnisch Ondernemerschap in Nederland: Ontwikkelingen en Perspectieven." *Migrantenstudies* 23 (2): 76–98.

Vertovec, S. 2007. "Super-Diversity and Its Implications." *Ethnic and Racial Studies* 30 (6): 1024–1054.

Waldinger, R. 1996. *Still the Promised City? African-Americans and New Immigrants in Postindustrial New York.* Cambridge: Harvard University Press.

Waldinger, R., H. Aldrich, and R. Ward. 1990a. "Opportunities, Group Characteristics, and Strategies." In *Ethnic Entrepreneurs*, edited by R. Waldinger, H. Aldrich, and R. Ward, 13–48. London: Sage.

Waldinger, R., D. McEvoy, and H. Aldrich. 1990b. "Spatial Dimensions of Opportunity Structures." In *Ethnic Entrepreneurs*, edited by R. Waldinger, H. Aldrich, and R. Ward, 106–130. London: Sage.

Whitley, R. 1999. *Divergent Capitalisms: The Social Structuring and Change of Business Systems: The Social Structuring and Change of Business Systems.* Oxford: Oxford University.

Divergent resilience: the employment growth paths of Amsterdam and Rotterdam, 2000-2014

J.J. Trip[1], R.C. Kloosterman[2] & A. Romein[1]
[1] *Faculty of Architecture and the Built Environment, Delft University of Technology, the Netherlands.*
[2] *Centre for Urban Studies, University of Amsterdam, the Netherlands.*

Abstract

Amsterdam and Rotterdam have followed rather different trajectories after 1970 when deindustrialisation set in. Amsterdam benefited strongly from the growth in financial services, creative and cultural industries (CCIs) and tourism after 1990. Moreover, it has been quite successful in attracting highly skilled workers. Rotterdam, meanwhile, has fared less well. This has also become apparent after 2008, when employment growth in Amsterdam bounced back while that in Rotterdam stagnated. The Amsterdam economy, in other words, appears to be more resilient than that of Rotterdam.

Resilience is a complex concept and many potential explanations are on offer. In this paper, we focus on the role of the sectoral composition in explaining the divergent growth paths of Amsterdam and Rotterdam using a shift-share analysis of employment data over the period 2000-2014 as a strategic window. This is a first explorative step to a more comprehensive understanding of these cities' growth paths.

Keywords: resilience; sectoral composition; employment; shift-share analysis; growth path

1 Introduction

According to Mark D'Eramo [1: 85] "[t]he contrast between the two port cities of Amsterdam and Rotterdam, less than half an hour apart, could hardly be more dramatic". He focuses on the strikingly different urban landscapes of both cities – Amsterdam with its dense and functionally mixed urban environment and

Rotterdam characterised by high-rise towers and broad avenues – and points at how they were shaped by port activities. Pre-industrial marine technology in the case of Amsterdam and industrial/container technology in the case of Rotterdam. It is not just the built environment that sets these two cities, with more or less comparable numbers of inhabitants (in 2014, Amsterdam had 811,000 and Rotterdam 618,000 inhabitants [2]), within commuting distance, and part of the same national institutional framework, apart. Amsterdam and Rotterdam also have followed rather different trajectories of development, notably after 1970 when deindustrialisation set in. Amsterdam, historically the financial and cultural capital of the Netherlands, benefited strongly from the growth in financial services, creative and cultural industries (CCIs) and tourism after 1990 [3, 4]. Rotterdam, has, on the whole, fared less well and still seems to struggle with the aftermath of industrial decline and the large-scale automation of ship-handling (based on containers) which has resulted in significant job losses. Amsterdam has been much more successful in attracting and retaining highly educated workers than Rotterdam [5: 41] thereby creating a virtuous circle with urban amenities (such as shops, restaurants, cafés, and galleries) making the city even more attractive.

Below, we will explore how the two cities have fared with respect to their employment growth paths. We will compare the overall and the sectoral development of the number of jobs both before and after the outbreak of the credit crisis in 2008. We will examine the responses to the external shock of the credit crisis thereby assessing the resilience of the two urban economies. Resilience in our view refers “to the ability of a region to accommodate shocks” but also to “the ability of regions to reconfigure their socio-economic and institutional structures to develop new growth paths” [6: 734]. The concept of resilience is useful when analysing how the shock of the credit crisis has influenced the developmental trajectories of both cities [7].

Diversity figures prominently among the factors which have been proposed to explain differences in resilience [6–9]. We address the question if and to which extent the sectoral composition can explain differences in economic resilience between Amsterdam and Rotterdam. To do so, we use a shift-share analysis which enables us to distinguish the impact on the urban employment growth paths of (1) factors related to the national developments; (2) those related to the sectoral composition or industrial mix; and (3) those that are first and foremost local.

We first discuss the methodology and the data (Section 2). After that, we present the pre-shock and post-shock employment growth paths of both cities as well as a breakdown of the sectoral composition (Section 3). We then go into the results of the shift-share analysis (Section 4). We conclude by pointing at potential explanations for the observed patterns (Section 5).

2 Methodology and data

We apply a twofold approach to analyse the development of employment and sectoral composition in Amsterdam and Rotterdam. First, we analyse sectoral composition in terms of the employment shares of different sectors as well as location quotients. This will show for each of the two cities which are the main economic sectors, but also which sectors are overrepresented or underrepresented. Second, we use a shift-share analysis [10]. This allows us to disentangle the growth or decline of employment into three components:

1) national growth (NG) or national share: the growth that can be attributed to the national factors such as interest rates and national demographics and policies;
2) industrial mix (IM): the growth that may be attributed to a city's specific economic composition and specialisation; and
3) regional share (RS) or competitive effect: the growth that may be attributed to region-specific factors, such as accessibility, agglomeration economies and quality of life.

For each city, these factors can be distinguished on a city level as well as for specific sectors. The sum of the three components equals the shift, i.e. the total employment growth in a city or sector. Together, these methods provide more detailed insights in the underlying trends shaping the employment growth paths of the two cities. However, it must be stressed that a shift-share analysis in itself does not offer explanations.

For the analyses we used data on the employment ("All jobs of employees in the Netherlands, in practice all jobs that fall under the Dutch wage tax legislation. Jobs of Dutch people working abroad are excluded, jobs of foreigners working in the Netherlands are included. There are no restrictions based on age or weekly working hours. Jobs of self-employed workers are excluded" [2]) for each sector in Amsterdam and Rotterdam, as well as for the Netherlands as a whole. Data is classified according to the Dutch Standaard Bedrijfsindeling (SBI), which follows the international SIC and NACE classifications. Data covered the period from 2000 to 2014 [2].

Two issues should be noted regarding these data. First, due to discontinuities in the way data have been collected, it is virtually impossible to construct consistent time series broken down by sector for the entire period:

1) a change in the method of data collection. Before 2006 data were collected by means of the yearly questionnaire on employment and payment (EWL), but starting from 2006 they were obtained from the social security registration (UWV), which was based on a slightly different definition. This leads to overall somewhat higher employment figures; and
2) a change in classification. In 2008, the SBI 2008 classification replaced the SBI'93 classification used until then. This reflected a similar change in the international NACE classification.

Given only slight differences, it is safe to assume that the 2006 change in data collection method did not significantly affect the relative economic composition in terms of sectors' shares. This implies that calculations based on individual years (e.g. location quotient) were only affected by the 2008 change in SBI classification. Fortunately from the perspective of our analysis, this change coincides with the economic crisis, making it still possible to analyse ‘before’ and ‘after’ the crisis, which would also correspond to the maximum interval advised for shift-share analysis [11]. Due to the changes in 2006, however, three periods had to be distinguished regarding the shift and share analysis: 2000-2005, 2006-2008 and 2008-2014.

Second, data on several specific sectors was classified as ‘confidential’. The total employment in these sectors is known, just not the size of the individual sectors. As this concerned only sectors which are quite small in both cities (e.g. agriculture, mining and energy production), we decided to group these sectors; this entails sectors A-C and E for SBI’93 and A, B, D and E for SBI 2008.

3 Analysis

3.1 Employment

Between 2000 and 2014, employment in Amsterdam grew from around 470,000 jobs to nearly 570,000 or with about 21 per cent. The growth path of Rotterdam is less steep with 330,000 in 2000 to nearly 370,000 in 2014 or some 12 per cent (Figure 1). These percentages are around 8 to 10 per cent point less if we account for the aforementioned change in definitions, but the difference remains the same. Resilience is about the ability of a city or a region to bounce back and return to its former growth path after a shock. Figure 2 shows the employment growth paths in both cities and the Netherlands (base year 2008). The pre-shock growth paths of Amsterdam and Rotterdam (and the nation as whole) are quite similar with a slightly steeper increase of employment in Rotterdam compared to Amsterdam. After the financial crisis broke out in 2008, however, a striking divergence occurred: Amsterdam displayed a strong growth of employment whereas Rotterdam showed a stagnation. The growth paths of the two cities, then, embody empirically the two contrasting cases theoretically distinguished by Martin and Sunley [7: 22] of a city which “did not return to its previous growth path, emerges from the shock on a lower growth trajectory” and a city where the shock had the effect of “propelling the region on a recovery trajectory that is much more favourable than the region’s pre-shock growth trend: its economic base emerges from the shock with a greater growth potential.” Below, we will attempt to disentangle the effects of the sectoral composition from more city-specific factors on these diverging growth paths.

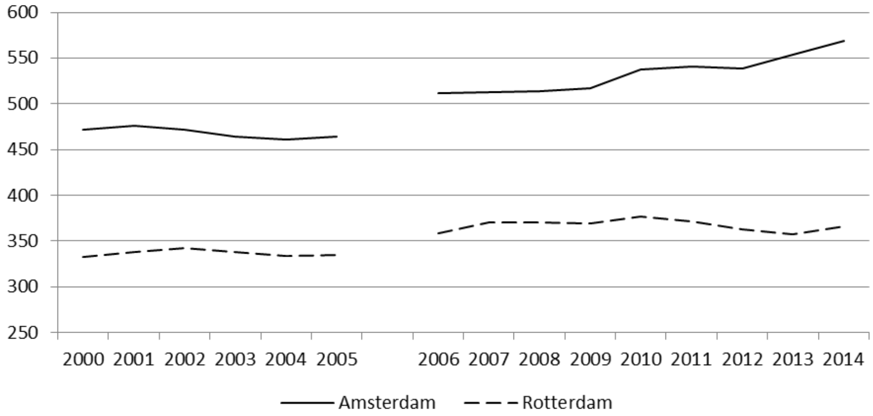


Figure 1: Employment trends, Amsterdam and Rotterdam, 2000-2014 (break due to definition change in 2006).

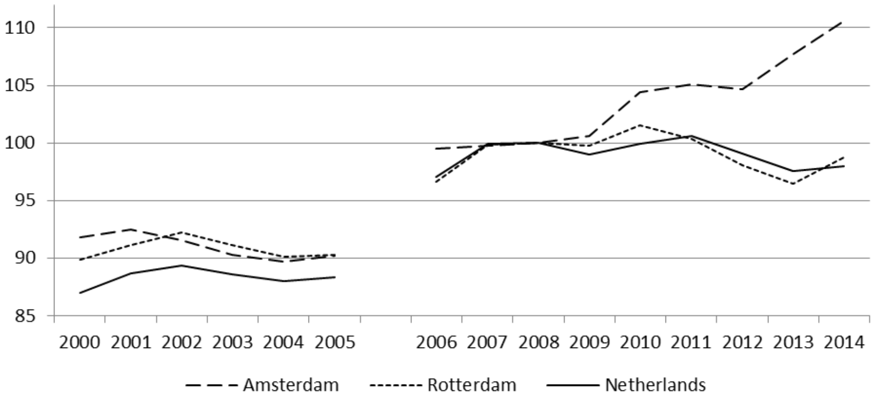


Figure 2: Development of employment (index: 2008=100. Break due to definition change in 2006).

3.2 Sectoral composition

Figure 3 shows the sectoral composition of Amsterdam and Rotterdam in 2008. Both cities are dominated by service activities, but Rotterdam shows a relative specialization in manufacturing and transport, due to its seaport which is largely included within its municipality boundaries (in contrast to Schiphol Airport, which is located outside Amsterdam). It also has a slightly larger employment share for healthcare. In contrast, Amsterdam has higher shares of tourism, IT and financial and business services.

The picture becomes somewhat more nuanced, but not very different, if we consider location quotients (Table 1 and Table 2). These show the overrepresentation or underrepresentation of sectors compared to the national average. Financial services are particularly strong, relatively, in Amsterdam, but they are overrepresented in Rotterdam as well. The same is true to a lesser extent for business services and culture, sports and recreation, reflecting the fact that while Amsterdam has a stronger service, cultural and knowledge economy than Rotterdam, the latter still has service dominated large-city economy.

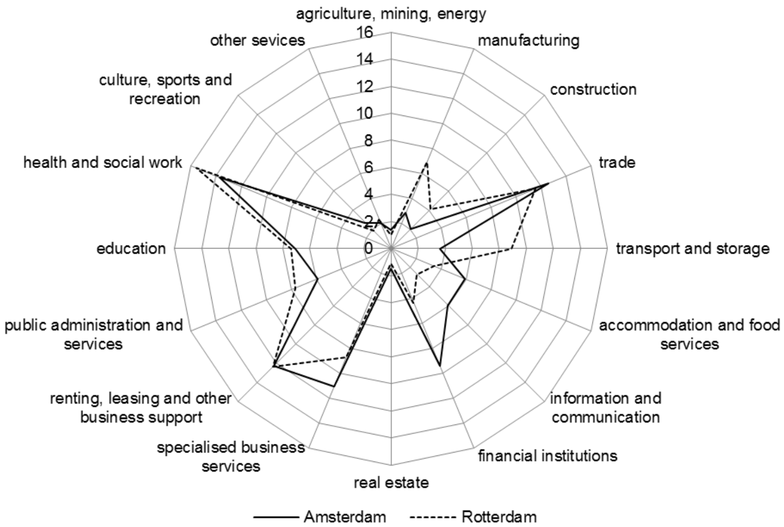


Figure 3: Sectoral composition (percentage of total employment) in 2008.

Table 1: Sectoral composition location quotients for Amsterdam and Rotterdam, 2000-2008.

		Amsterdam		Rotterdam	
SBI'93	Sector description	2000	2008	2000	2008
A-C; E	Agriculture, mining, electricity, gas and water supply	0.32	0.21	0.53	0.42
D	Manufacturing	0.39	0.44	0.65	0.69
F	Construction	0.45	0.39	0.75	0.81
G	Trade	0.80	0.76	0.76	0.70
H	Hotels and restaurants	1.55	1.48	0.91	0.86
I	Transport and communication	1.13	0.89	1.76	1.70
J	Financial intermediation	2.70	2.35	1.44	1.18
K	Business activities	1.41	1.34	1.25	1.17
L	Public administration	1.04	0.92	0.99	1.21
M	Education	1.01	1.07	1.12	1.12
N	Health and welfare	0.90	0.89	1.03	1.00
O	Culture and other services	1.45	1.55	0.99	1.12
Total		1.00	1.00	1.00	1.00

Table 2: Sectoral composition location quotients for Amsterdam and Rotterdam, 2008-2014.

		Amsterdam		Rotterdam	
SBI 2008	Sector description	2008	2014	2008	2014
A-B, E-D	Agriculture, mining, energy and water supply, waste management	0.66	0.53	0.50	0.79
C	Manufacturing	0.28	0.23	0.68	0.63
F	Construction	0.41	0.40	0.82	0.97
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	0.76	0.74	0.70	0.73
H	Transportation and storage	0.74	0.62	1.81	1.93
I	Accommodation and food service activities	1.48	1.53	0.86	0.90
J	Information and communication	1.94	2.21	0.88	0.76
K	Financial institutions	2.73	2.67	1.26	1.12
L	Renting, buying and selling of real estate	1.56	1.55	1.11	1.15
M	Consultancy, research and other specialised business services	1.64	1.71	1.29	1.17
N	Renting and leasing of tangible goods and other business support services	1.12	1.35	1.12	1.18
O	Public administration, public services and compulsory social security	0.93	0.85	1.22	1.21
P	Education	1.06	0.95	1.11	1.08
Q	Human health and social work activities	0.89	0.82	1.01	1.01
R	Culture, sports and recreation	1.64	1.50	1.11	1.22
S	Other service activities	1.08	1.10	1.24	0.69
Total		1.00	1.00	1.00	1.00

3.3 Shift-share analysis

We have calculated the contributions of the different components to the employment growth paths using shift-share analysis for each city as a whole and their individual sectors for a pre-shock (2000-2005) and a post-shock period (2008-2014).

The National Growth (NG) component was positive for both cities in the pre-shock period. In Amsterdam this component contributed 7,176 and in Rotterdam 5,060 to the overall employment trend between 2000 and 2005 (see Table 3). The contribution of the sectoral composition or Industrial Mix (IM) was also positive, respectively 7,753 for Amsterdam and 3,552 for Rotterdam. The contribution of the specific urban factors, the Regional Share (RS), however, was for both cities negative, with -23,070 for Amsterdam and -6,962 for Rotterdam. The overall result was a slight decline of employment in Amsterdam (-8,140), whereas the national and industrial mix factors more than compensated for the negative local impact in Rotterdam with a net growth of 1,650. Broken down by sectors, Amsterdam shows almost across the board negative Regional Shares. In Rotterdam, manufacturing, construction and public administration show positive Regional Shares (see Table 3). This confirms the earlier observation that

Rotterdam was doing slightly better in terms of its pre-shock employment growth path than Amsterdam.

Table 4 shows a break-down of the growth paths of the two cities into three components for the post-shock period 2008-2014. In marked contrast to the pre-shock period, we find negative National Growth components regarding the overall employment trend for both cities, respectively -10,224 for Amsterdam and -7,364 for Rotterdam thus reflecting the deep and prolonged recession on a national level after 2008. The impact of the Industrial Mix (IM) is positive in Amsterdam with 2,227 and slightly negative in Rotterdam with -175. The big surprise can be found in the contribution of local factors. The calculated effect of the Regional Shares (RS) in Amsterdam stands at a staggering 62,526, amply compensating for the negative impact of the national factors and resulting in a net growth of 54,530. The picture for Rotterdam is rather different with a much smaller Regional Share of 3,059 which falls short of the loss due to national factors and resulting in a net loss of employment of 4,480.

We can observe a negative IM component for sectors that have been especially hard hit by the crisis, such as construction, finance and other business services (which include e.g. advertisement and consultancy, but also sectors directly related to construction and finance, such as architecture and legal services). The IM component is positive for healthcare and, in Amsterdam, accommodation and food services, i.e. tourism. Remarkably, however, is the strong regional share RS in Amsterdam, which adds to positive IM components and mitigates the effect of negative ones in the above sectors. This is in sharp contrast to Rotterdam, where RS components are all relatively small (even taking into account the smaller size of the local economy) and partly negative. For several sectors they are in fact more negative than IM components.

Table 3: Shift-share analysis 2000-2005.

	Amsterdam				Rotterdam			
SBI’93	NG	IM	RS	Shift	NG	IM	RS	Shift
A-C; E	50	-382	113	-220	57	-439	-1,098	-1,480
D	387	-3,627	-2,210	-5,450	458	-4,292	1,164	-2,670
F	189	-1,222	-837	-1,870	225	-1,456	1,171	-60
G	993	-3,327	-3,215	-5,550	669	-2,242	-3,377	-4,950
H	411	-294	-587	-470	171	-122	-389	-340
I	532	-2,600	-4,643	-6,710	586	-2,860	154	-2,120
J	748	-2,048	21	-1,280	281	-771	-1,160	-1,650
K	1,640	5,215	-4,455	2,400	1,029	3,273	-3,642	660
L	527	561	-2,908	-1,820	353	375	1,612	2,340
M	449	2,656	-1,405	1,700	351	2,076	-417	2,010
N	854	11,824	-1,698	10,980	688	9,527	-55	10,160
O	398	998	-1,245	150	192	482	-924	-250
Total	7,176	7,753	-23,070	-8,140	5,060	3,552	-6,962	1,650

Table 4: Shift-share analysis 2008-2014.

	Amsterdam				Rotterdam			
SBI 2008	NG	IM	RS	Shift	NG	IM	RS	Shift
A-B, E-D	-140	281	-721	-580	-76	151	2,384	2,460
C	-292	-1,054	-773	-2,120	-513	-1,851	-1,446	-3,810
F	-210	-2,275	755	-1,730	-303	-3,279	2,292	-1,290
G	-1,289	925	6,463	6,100	-852	612	2,100	1,860
H	-372	-702	-986	-2,060	-654	-1,237	2,281	390
I	-603	4,425	5,768	9,590	-254	1,866	648	2,260
J	-606	-26	8,572	7,940	-199	-8	-1,312	-1,520
K	-964	-4,282	4,425	-820	-321	-1,426	-1,503	-3,250
L	-154	-846	819	-180	-79	-435	134	-380
M	-1,131	-3,261	9,182	4,790	-640	-1,846	-2,375	-4,860
N	-1,258	1,927	23,120	23,790	-910	1,394	2,626	3,110
O	-598	1,206	1,262	1,870	-565	1,140	-5	570
P	-723	-421	424	-720	-546	-318	-496	-1,360
Q	-1,404	7,142	2,432	8,170	-1,146	5,831	335	5,020
R	-274	284	440	450	-134	139	705	710
S	-207	-1,096	1,343	40	-172	-908	-3,311	-4,390
Total	-10,224	2,227	62,526	54,530	-7,364	-175	3,059	-4,480

A closer look at the different sectors reveals marked differences between Amsterdam and Rotterdam with respect to the IM and RG components, i.e. between the growth based on sectoral economic composition and that based on regional competitiveness. In the period 2000-2005 the overall regional share is negative in both cities, especially in Amsterdam. The IM component is positive, but whereas Amsterdam has an advantage here it is not sufficient to compensate for the negative RS component. After the shock of the outbreak of the credit crisis, we observe a negative IM component for sectors that have been especially hard hit by the crisis, such as construction, finance and other business services (which include e.g. advertisement and consultancy, but also sectors directly related to construction and finance, such as architecture and legal services). The IM component is quite positive for healthcare and, in Amsterdam, accommodation and food services, i.e. tourism. Amsterdam clearly outperforms Rotterdam in almost every single sector when it comes to the contribution of local factors (see Figure 4).

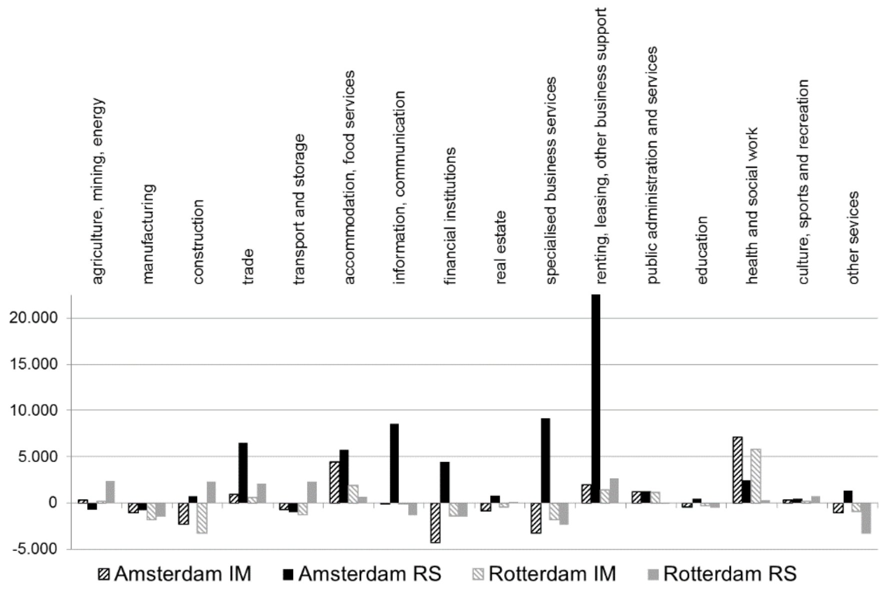


Figure 4: Industrial mix (IM) and regional share (RS) components, 2008-2014.

4 Divergent growth paths

Our findings indicate that, in terms of employment, the Amsterdam economy is recovering much faster than that of Rotterdam after the 2008 crisis, and has bounced back far beyond its pre-crisis level. Our results also suggest that this divergence cannot be primarily explained by the differences in sectoral composition . Whereas Amsterdam indeed has a slight advantage over Rotterdam in terms of its industry mix or sectoral composition, this effect is smaller than before 2008, and is dwarfed by the much larger contribution of specific local factors to employment growth. It appears, as our shift-share analysis shows, that particularly city-specific factors play a decisive role in the strong recovery of the city after 2008. This is in line with recent findings of Martin *et al* [13] for the UK.

Shift-share analysis has its limitations. It considers employment change over one specific time interval, and focuses on the city as a self-sufficient system that does not depend on the development of other regions [12]. In addition, the analysis is dependent on the delineation of its principal building blocks, the sectors, and cannot provide insights in intra-sectoral trends within subsectors (or within the firms). Nonetheless, the employment trends we have identified seem very clear with Amsterdam and Rotterdam more or less displaying similar growth paths before the credit crisis and strongly different trajectories after that with Amsterdam showing a much stronger performance both when compared to its own pre-shock path and to Rotterdam's post-shock growth path.

However, another limitation of the method applied is that while we can quantify the various components of employment growth, our analysis does not provide an explanation of what they actually entail. To explain why Amsterdam has been able to enter a post-shock new growth path, displaying thereby a more fundamental type of resilience, we have to look beyond the shift-share analysis. According to Boschma [6: 735], a new growth path is based on the "ability to adapt and reconfigure their industrial, technological and institutional structures in an economic system that is restless and evolving" At this stage, we can only speculate on what might constitute these emerging structures. We offer a few suggestions which are analytically distinct but which in reality may be interrelated and/or overlap. Urban economies do not change overnight but tend to move along path-dependent trajectories, so we have to look first at the existing strengths of Amsterdam and how they might be recombined to create new structures.

First, it may be that significant shifts have occurred within the sectors. It might be, for instance, that the broader delineation of the sectoral composition hides rapidly expanding subsectors. Secondly, it may be that on a firm level, new strategies with respect to markets, products and ways of production have been developed. Third, Amsterdam may offer a very suitable environment for foreign firms given its quality of place [5, 14] and its combination of amenities and a large pool of highly educated workers with a cosmopolitan outlook.

Our findings are only the starting point for further research that should reveal the factors behind the observed patterns. A more extensive analysis of the factors behind the divergent trajectories of the two largest Dutch cities and, more specifically, of the underlying drivers of Amsterdam's new growth path is high on our agenda.

Acknowledgements

We would like to thank Miranda Jonker from the University of Amsterdam for her assistance in the data collection, and Erik Louw from Delft University of Technology for his advice on the analysis.

References

[1] D'Eramo, M., Dock life. *New Left Review*, **96**, pp. 85-99, 2015.
[2] CBS Statline, Statistics Netherlands, http://statline.cbs.nl/Statweb.
[3] Kloosterman, R.C., Double Dutch: trends of polarisation in Amsterdam and Rotterdam after 1980. *Regional Studies*, **30(5),** pp. 367-376, 1996.
[4] Kloosterman, R. C., Faces of migration: migrants and the transformation of Amsterdam. *Migration and London's Growth*, ed. B. Kochan, London School of Economics: London, pp. 127-143, 2014.
[5] Vermeulen, W., Teulings, C., Marlet, G. & de Groot, H., *Groei & Krimp. Waar moeten we bouwen – en waar vooral niet?*, VOC Uitgevers: Nijmegen, 2016.

[6] Boschma, R., Towards an evolutionary perspective on regional resilience. *Regional Studies,* **49(5)**, pp. 733-751, 2015
[7] Martin, R. & Sunley, P., On the notion of regional economic resilience: conceptualization and explanation. *Journal of Economic Geography,* **15(1),** pp. 1-42, 2015.
[8] Scheffer, M., *Critical Transitions in Nature and Society*. Princeton University Press: Princeton, New Jersey, 2009.
[9] Diodato, D. & Weterings, A.B., The resilience of regional labour markets to economic shocks: Exploring the role of interactions among firms and workers. *Journal of Economic Geography*, **15(4),** pp. 723-742, 2015.
[10] Loveridge, S., A practical approach to shift-share analysis*. Journal of the Community Development Society,* **26(1),** pp. 110-124, 1995.
[11] Quintero, J.P., Regional economic development: an economic base study and shift-share analysis of Hays County, Texas. *Applied Research Projects, Texas State University-San Marcos*, Paper 259, 2007.
[12] Mayor, M. & López, A.J., Spatial shift-share analysis versus spatial filtering: an application to Spanish employment data (Chapter 7). *Spatial econometrics; methods and applications*, eds. G. Arbia & B.H. Baltagi, Fysica-Verlag: Heidelberg, pp. 123-142, 2009.
[13] Martin, R., Sunley, P., Gardiner, B. & Tyler, P., How regions react to recessions: resilience and the role of economic structure. *Regional Studies,* **50(4),** pp. 561-585, 2016.
[14] Trip, J.J., Assessing quality of place: a comparative analysis of Amsterdam and Rotterdam. *Journal of Urban Affairs,* **29(5),** pp. 501-517, 2007.

File ID	uvapub:159455
Filename	446658
Version	final

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	book chapter
Title	Faces of migration: migrants and the transformation of Amsterdam
Author(s)	R.C. Kloosterman
Faculty	FMG: Amsterdam Institute for Social Science Research (AISSR)
Year	2014

FULL BIBLIOGRAPHIC DETAILS:
http://hdl.handle.net/11245/1.446658

Migration and London's growth

Final report of LSE London's HEIF 5 project on Migration and the Transformation of London led by Christine Whitehead, Ian Gordon and Tony Travers

Edited by

Ben Kochan

LSE London 2014

Published by LSE London
The London School of Economics and Political Science
Houghton Street
London WC2A 2AE
UK

LSE London acknowledges funding for the development and production of this book from HEIF 5.

British Library Cataloguing in Publication Data.
A catalogue record for this book is available from the British Library.

ISBN 978-1-909890-11-4

Printed in Great Britain by RAP Spiderweb Limited, Lancashire

Layout and design by Ben Kochan
Photographs by Ben Kochan unless otherwise credited

This publication, as well as papers and presentations associated with the project are available on LSE London's web-site: http://lselondonmigration.org/

Contents

Section 3. Migration, communities and local services

Image: Robert Kloosterman

"Not just Amsterdam, but many other cities in the world are also displaying ever more diverse faces of migration spanning low-paid cleaners to high-flying CEOs of multi-nationals"

Faces of migration: migrants and the transformation of Amsterdam

Robert C. Kloosterman

Multicultural markets: the street and the financial species

The daily Dappermarkt street market is clearly a meeting place for people from almost everywhere in the world with its sights and sounds and smells of all kinds of food being sold. The market, in the eastern part of Amsterdam, built in the 19th/early 20th century, is an urban renewal area with a highly diverse resident population in terms of their place of birth, length of stay, family status, education, socio-cultural orientation and life prospects. There are members of what used to be the 'traditional' (white) working class, migrants from less-developed countries who came as guest workers and their descendants, migrants from the former colonies and their descendants, political refugees and asylum seekers and young urbanites, mostly students or recent graduates from the Netherlands and increasingly from abroad.

Only a few kilometres to the south, there is a quite different type of street life. In the Zuidas business district with its gleaming office towers, there are people who work in advanced business services sitting outside posh cafés. Their more or less uniform dress style and conversation often in English, the lingua franca of their trade, belies their diverse national backgrounds from the Netherlands, the US, the UK, Germany, France, Japan, China and many other countries.

Both city snapshots are images which reflect the presence of migrants in Amsterdam. Not just Amsterdam, but many other cities in the world are also displaying ever more diverse faces of migration spanning low-paid cleaners to high-flying CEOs of multinationals (Vertovec, 2007). In addtion, ethnic diversity associated with migration tends to overlap with socio-economic fault lines resulting in segmented urban societies. To put the Amsterdam experience in perspective, I briefly sketch the relationship between the emerging production system and social stratification. After that, I set out very briefly recent migration trends in Amsterdam. I then explore how this diverse migrant population has fared in socio-economic terms using quantitative data from the local

Amsterdam office (O+S) Dutch and the national statistical office (CBS). I conclude by trying to draw out more general lessons.

A new wave of urbanisation

Cities are always changing but I want to distinguish between the different phases of urbanisation under industrial capitalism and to focus on the more enduring generic characteristics, specifically those concerned with the production system: the dominant technology, the leading industrial sectors, the key products, the prevailing labour practices and managerial strategies, the locational patterns, and the main cleavages in the social division of labour. The current phase, labelled cognitive-cultural capitalism by Allen Scott (2008 & 2012), is first and foremost characterised by a production system in which 'highly qualified human capital' constitutes the crucial input and more customised products are an increasingly important output (Scott, 2012: p.x). Within the leading sectors, then, the key input consists either of highly formalised knowledge - as in advanced producer services and high-tech industries - or of a less-formalised form of aesthetic, cultural or craft knowledge as in creative industries such as fashion design or music making. Both forms of human capital may enable firms to compete more on quality than on price. Thus they are able to find or create a particular niche and focus on a limited range of products, and escape, at least temporarily, cut-throat competition from other firms.

These leading activities tend to thrive when, on the one hand, firms are able to reap agglomeration economies by being concentrated in local clusters, and, on the other, when they are able to tap easily into global networks to maintain contacts with suppliers, customers and others. Those cities or city-regions with the right conditions have benefited and seen employment and population growth. They have a certain critical mass and diversity, an orientation towards advanced producer services, high-tech industries and/or creative and cultural industries, and a tradition of openness and global linkages and a supportive infrastructure. In general "… capitals demonstrated a striking ability to recover, aided by their dominance of national and international business and transport networks. The upturn in world banking and finance since the 1980s consolidated their primacy." (Clark, 2009: p.242) Global cities such as New York (not a political but certainly a financial capital) and London evidently belong to this category, but on a more modest level, smaller cities like Amsterdam have displayed similar dynamics achieving an urban renaissance with strong growth of cognitive-cultural activities, a resurgence of their inner-cities and widespread gentrification (Engelen, 2007; Fernandez, 2012; Kloosterman, 2013). Amsterdam's economy is currently dominated by producer services which provided 39 per cent of the city's total employment in 2013, (O+S, Jaarboek 2013) and creative and cultural industries which provided 10.5 per cent of total employment in 2012 (O+S, Jaarboek 2013; http://www.rotterdam.nl/sectorenenclusters).

While the competitiveness of the evolving urban production system hinges on the leading sectors, these activities do not exist in a vacuum, but "are almost directly coupled with an adjunct penumbra comprising a flexible low-wage employment segment … focused on jobs like housekeeping, child care, health care, food preparation and serving, janitorial work, taxi driving, and home repair" (Scott, 2012: p. 43). Their labour market position is not just characterised by low pay but also tends to be more insecure, and, moreover, are dead-end jobs in terms of opportunities for upward social mobility. Given the often precarious situation these workers find themselves in, they belong to a new vulnerable group, the 'precariat' (Scott, 2012: p.101). Migrants from non-western countries are typically overrepresented in this class of servile workers.

Accordingly, with this new phase of urbanisation, a central fault line is emerging between those who work in cognitive-cultural activities and those in supporting activities (Sassen, 1991/2001; Scott, 2008 & 2012). Both poles of the labour market are, in principle, able to attract migrants, especially in global cities with already firmly established linkages with various parts of the world. However, given the diverging requirements for the jobs in both segments, multifaceted sorting processes in the urban labour market create a complex mosaic along lines of human capital, social and cultural capital, and, often ethnicity. This mosaic is also spatially articulated as the position in the labour market is strongly related to the status in the housing market resulting in neighbourhoods with very different ethnic compositions. Migration, hence, is not just about those who live close to the Dappermarkt, but also about those who live in the expensive parts of the Canal District in the historical centre of the city.

This is, of course, a highly schematic rendering of what is happening in advanced urban economies. It does, however, provide a few key reference points for positioning a city and its migration experience. Before locating Amsterdam in this broader scheme of contemporary processes of urbanisation, we first consider the migration trends and some key indicators of the socio-economic position of the migrants.

Migration trends

In 1959, Amsterdam's population peaked at about 872,000 inhabitants and the nadir was in 1985 at just under 676,000. Without the arrival of migrants from abroad this decrease could have been much larger, as the birth rate had declined and, moreover, many people, especially middle-class families with children, left the city in droves for the green promise of suburbia (Kloosterman, 1994). Migrants from Turkey and Morocco came as guest workers to fill low-skilled jobs and later as part of family reunion schemes. There were also migrants from Surinam who left when the former Dutch colony became independent in 1975. These migrants partly replaced those leaving for Purmerend, Hoorn, Almere, and other smaller cities in the vicinity of Amsterdam (Mak, 1995: p. 328). The socio-economic effects of this relatively rapid change in Amsterdam's population became painfully evident in the 1980s when a deep recession hit the Netherlands and Amsterdam in particular. Deindustrialisation rapidly pushed up unemployment, especially among migrant workers. When employment started growing again with an expansion of the service sector in the second half of the 1980s, the local labour force in Amsterdam turned out to be less well qualified for the new jobs. It was not just that the educational qualifications of many members of the labour force were deemed inadequate, employers preferred workers from outside Amsterdam who they saw as better educated and more eager (Kloosterman, 1994). As a result, a significant part of the Amsterdam labour force found itself at the end of the hiring queue and, consequently, the overall rate of unemployment in Amsterdam went up to about 25 per cent – one of the highest in the country - with even higher rates among migrants. Amsterdam in the 1980s, with its riots, squatters, and highly visible population of junkies close to the Central Station seemed on its way to becoming an intractable urban mess (Bosscher, 2007a,b and c).

In retrospect, however, we can now observe that, although unemployment remained very high throughout the 1980s, the signs of an urban renaissance were already there. A period of economic and demographic decline, of sharp political, social and cultural turmoil ended more or less in 1985.

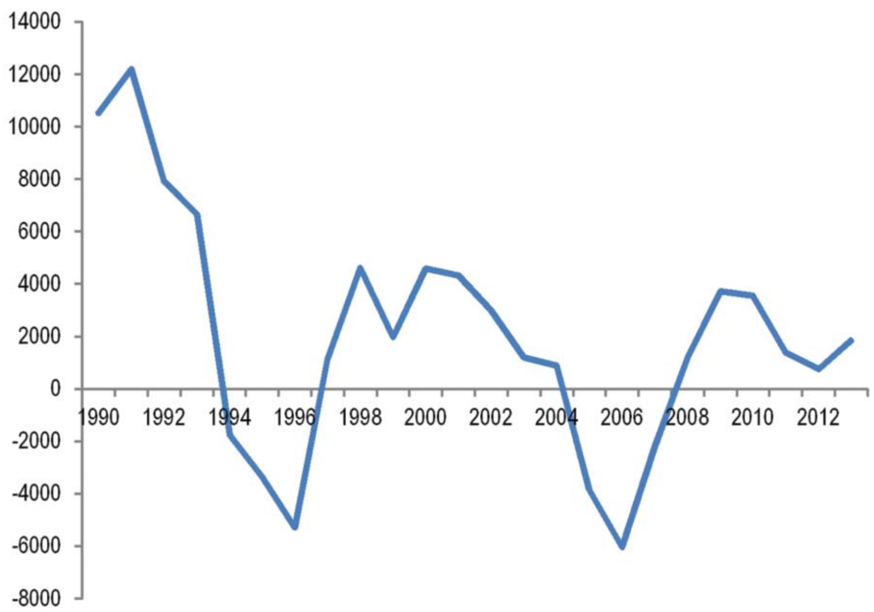
Geert Mak (1995: p.326), in his history of Amsterdam, has dubbed this period from 1965 to1985 the 'Twenty Years Urban War'. It was not just that after two decades of deindustrialisation, the urban economy picked up again in the mid-1980s with the expansion of the services sector and growth of employment, but city life itself had become more popular among large numbers of higher educated young people who, instead of leaving the city soon after graduation, stayed in the more central parts of the city even when they had children (Boterman, 2012), and widespread processes of gentrification started to alter the outlook of the city. The Jordaan, a rather dilapidated and deserted neighbourhood in the 1970s, became a prominent playground for the new urbanites. Teachers, academics, and increasingly people working in the producer services had become the successors to the indigenous working-class Amsterdam residents who had left the city (Mak, 1995; p. 326). The fear expressed in the 1970s

and early 1980 that Amsterdam would be stuck with a lower educated labour force, did not materialise. Instead, in 2007, Amsterdam ranked second, only behind much smaller Utrecht, in terms of the proportion of higher educated workers in the city's labour force with slightly more than 50 per cent (Marlet, 2009: p. 121). Indeed, a new phase of urbanisation had set in, driven by an unfolding production system strongly oriented towards cognitive-cultural activities. Amsterdam benefited greatly from the combination of a production system dependent on proximity because of agglomeration economies and connectivity because of global linkages with the shift in the residential preferences of high-skilled workers towards urban living.

These changes are reflected in the demographic trends of the Dutch capital. In 1985, after almost three decades of continuous decline, the Amsterdam population started growing again. After 1985, the birth rate in Amsterdam exceeded the death rate and has remained positive ever since so contributing to the population growth. Even more important at least until the late 1990s was the decline in net domestic outmigration, while the net international migration surplus remained until 1993 when outward migration temporarily exceeded inward (see Table 1; O+S, 2014). After 2008, all components contributed to population growth and at the beginning of 2014 (Latten and Deerenberg, 2013), the Amsterdam population went over the 800,000 mark again for the first time since 1971 (O+S, 2014a) (see figure 1).

International migrants from roughly 1965 until 1985 were mainly relatively low-skilled non-western guest workers and their family members or they came from Surinam. From 1990 onwards they were far more diverse (Obdeijn and Schrover, 2008; Latten and Deerenberg, 2013). There were political refugees, asylum seekers from the former Yugoslavia, Iraq, Iran, Ghana, Afghanistan, Somalia and other countries. They tended to be generally better educated than their predecessors. Moreover, with the liberalisation and extension of the European Union, citizens of other EU member states have acquired the right, in principle, to settle in the Netherlands. The mgrants from notably Great Britain, Germany, Spain, Belgium, France, Italy come mainly for work, study or to join a family member and are typically relatively highly educated (Obdeijn and Schrover, 2008 p.339). Amsterdam, with its expanding economy, its cosmopolitan atmosphere, its two universities and other higher education institutions, and its already sizeable population of western migrants, unsurprisingly, has the largest number of non-Dutch citizens of all Dutch cities (O+S, 2014). Whilst Dutch citizens make up about 695,000 of the city's 799,500 population, figure two shows that the remainder come from a wide range of countries. Non-Dutch citizens in Amsterdam

Figure 1: Net international migration to Amsterdam, 1990-2013

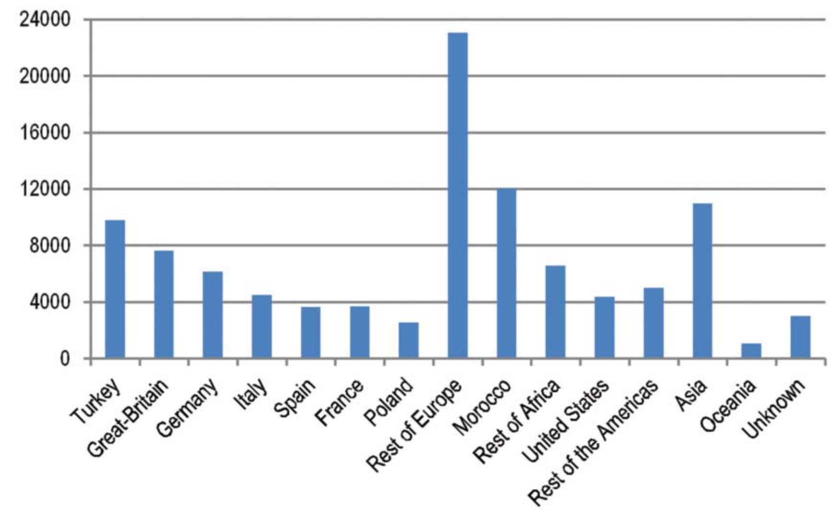


Source: O+S (2014a), http://www.os.amsterdam.nl/feiten-en-cijfers/#

(from all EU member states, the rest of Europe and the Americas), total about 60,500 persons, which is about 7.5 per cent of the overall population (see Figure 2).

Figure 2, however, grossly underestimates the diversity of the city's population as many people of Turkish, Moroccan or Surinam descent and other countries (and especially their offspring) have Dutch nationality. Table 1 shows a more nuanced picture of the share of migrants and their offspring in the population of Amsterdam. This table reveals that Amsterdam has become a majority/minority city in that people of Dutch origin have become a minority. Migrants from non-western countries have the largest share, with about a third of the total Amsterdam population. Migrants from western countries constitute about 16 per cent of the total Amsterdam population. The second generation (those born in the Netherlands of parents born elsewhere) count for little less than half of the population of non-Dutch origin Amsterdam.

Figure 2 Amsterdam's non-dutch population by nationality, January 1, 2013



Source: O+S (2014), Jaarboek 2014

Amsterdam in the 21st century has again become a city of migrants just like it was in its Golden Age, the 17th century, when people from Germany, Belgium, Spain, Portugal, among them many Huguenots and Jews, flocked to a booming and tolerant, or to put it more cynically, indifferent city. Many migrants encountered difficulties in finding their way and historical studies have shown that substantial numbers of migrants were part of a pre-industrial underclass living separate lives from the rest of the Amsterdam population (Obdeijn and Schrovers, 2008 p.59-63). Other migrants, though, were crucial in boosting the competitiveness of the Amsterdam economy by bringing knowledge of products and production processes (e.g. Huguenots and woollen cloth, Prak, 2002: p.158-159) and trade networks (e.g. Sephardi Jews, Israel, 1987 p.73). What can we say about their contemporary counterparts?

Migrants and the new wave of urbanisation

A comprehensive overview of the role of migrants in the Amsterdam economy is beyond the scope of this contribution. Instead, the focus is on the labour market position of migrants in Amsterdam. I will look at the unemployment trends among specific groups of migrants and especially among the youth, the differences between men

Table 1 The Amsterdam population by country of origin and generation

	Actual figures				%		
Country of origin	1st generation	2nd generation	Netherlands	Total	1st generation	2nd generation	% of total Amsterdam population
Surinam	37393	30097		67490	55.4	44.6	8.3
Dutch Antilles	6877	5211		12088	56.9	43.1	1.5
Turkey	21921	20290		42211	51.9	48.1	5.2
Morocco	34115	39196		73311	46.5	53.5	9.0
Other non-western countries	57047	29920		86967	65.6	34.4	10.7
Total non-western countries	157353	124714		282067	55.8	44.2	34.8
Western countries	73531	55494		129025	57.0	43.0	15.9
Netherlands			400093	400093			49.3
Total	230884	180208	400093	811185	28.5	22.2	100.0

Source: O+S (2014), Jaarboek 2014: 218. Figures at January 1 2014

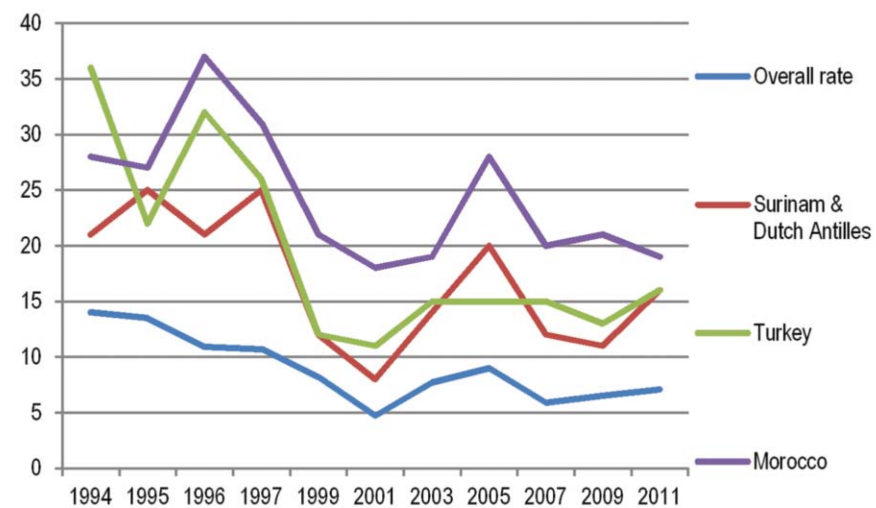
and women in groups of migrants and the nature of their work contracts and consider whether a 'precariat' is emerging along ethnic lines.

The recovery of the Amsterdam economy after 1985 first benefited highly educated young workers who to a large extent came from outside the city and largely bypassed the local pool of lesser educated unemployed (Kloosterman, 1994). As a result, the overall rate of unemployment in Amsterdam remained relatively high for quite some time (Kloosterman, 2013). Only after 1997 did the rate drop below 10 per cent and reached friction level in 2001 (see Figure 3). After that, the rate of unemployment remained relatively modest and even the credit crisis in 2008 did not affect the rate very much – at least until 2011. The picture for migrants and their descendants from Surinam, the Dutch Antilles, Turkey and Morocco, however, is rather different. Their

unemployment rates are consistently higher and, in some cases, much higher than the overall rate. These rates also appear to be more erratic, but this might be due, at least partly, to problems with data collection. Whatever the quality of the data, it seems beyond doubt that unemployment rates among 1st and 2nd generation migrants are much higher than those of indigenous Dutch (and western migrants). The data also suggests that these migrants from non-western countries are among the first to lose their job and the last to be hired.

If we focus on the labour market situation of young people up to the age of 26, these differences become even more salient. Recently, the municipality of Amsterdam published a Factsheet which gives a detailed picture of the labour market situation of young people in Amsterdam in 2014 (Monitor jeugdwerkloosheid Amsterdam 2014, 2014). This report shows that unemployment rates among young people of respectively Moroccan descent (48 per cent), Surinam and the Dutch Antilles (38 per cent) and Turkey (33 per cent) are considerably higher than the overall rate for young people in Amsterdam (24 per cent). These differences can only be partly explained by differences in educational qualifications as even highly educated young people from non-western countries are twice as likely to be unemployed as their counterparts of

Figure 3 Unemployment rates in Amsterdam by country of origin, 1994-2011 (%)



Source: O+S, Jaarboek 2013

Dutch descent (19 per cent compared to 8 percent; Monitor jeugdwerkloosheid Amsterdam 2014, 2014 p.2).

The same pattern is evident across all the working population. Table 2 shows the unemployment rates and the net labour market participation rates among different ethnic groups. Again, there are much higher rates of unemployment among migrants from non-western countries and, accordingly, much lower rates of net labour participation than among those from western countries and the Netherlands itself. Even more striking are the differences between women of the different groups. Unemployment among women from Surinam and Dutch Antilles (24 per cent), Turkey (27 per cent) and Morocco (a staggering 36 per cent) are much higher than among women from western countries and those of Dutch descent (both 5 per cent). These figures are very much in line with those mentioned by Vertovec (2007:1040) for

Table 2 Labour market participation and unemployment in Amsterdam by country of origin and gender, 2013.

	Men		Women		Men and women	
	Labour participation net rate (%)	Unemployment rate (%)	Labour participation net rate (%)	Unemployment rate (%)	Labour participation net rate (%)	Unemployment rate (%)
Surinam & Dutch Antilles	54	20	51	24	52	22
Turkey	57	18	24	27	41	21
Morocco	51	26	29	36	40	30
Other non western countries	54	22	42	23	48	22
Western countries	76	7	72	5	74	6
Netherlands	77	7	69	5	73	6
Total	69	12	59	11	64	11

Source: O+S (2014), Jaarboek 2014: 214

London: "Employment rates are especially low for women born in South Asia (37 per cent) and the Middle East and North Africa (39 per cent)". The high unemployment rate for women of Turkish and Moroccan descent in Amsterdam partly explains the very low rates of net participation, but they also reflect the fact that a much smaller number enter the formal labour market in the first place. The figures for workers of Dutch descent and those from other western countries, both men and women, are comparable but contrast sharply with that of migrants from non-western countries.

Labour market position is not just about rates of unemployment but also the quality of the jobs. The pay, security, how dangerous and/or dirty they are and whether they offer possibilities for upward social mobility are also key issues. There is no data on most of these characteristics, but there is a breakdown according to job security for different groups of workers for 2013 (see Table 3). This gives only a very limited insight into the nature of the jobs, but still marked differences are revealed between the various groups. It indeed seems the case that migrants from non-western countries are more likely to have flexible contracts, which is one of the defining characteristics of the 'precariat'. The proportion of migrants on flexible contracts from Morocco and Surinam and the Dutch Antilles at 40 per cent and 30 per cent respectively stands out. Interestingly, we do find a gap here between workers of Dutch descent and those

Table 3 Labour market position of the active Amsterdam labour force broken down by country of origin as a percentage of the total group, 2013

	Tenured (%)	Flexible contract (%)	Self-employed (%)	Combination of self-employed and job (%)
Surinam and Dutch Antilles	62	30	6	3
Turkey	50	24	26	1
Morocco	56	40	2	3
Other non-western countries	45	31	21	4
Western countries	48	31	18	3
Netherlands	55	21	18	6
Total	53	26	17	5

Source: O+S (2013), Jaarboek 2013, 4213

from other western countries. The proportion of workers from the latter group on flexible contracts is 10 percentage points higher than for those of Dutch descent.

Among workers from non-western countries, there are marked differences regarding self-employment with a very high rate of self-employment among those from Turkey with one in four being self-employed. This contrasts sharply with the self-employment rates among Surinamese and Dutch Antilleans and a significant proportion of Moroccans are on flexible contracts. From the data, it is impossible to tell to what extent this self-employment is out of necessity - pushed towards self-employment because of obstacles in the labour market - or out of choice (pulled) for reasons of higher remuneration or self-esteem (Kloosterman, 2010a). It is also unclear to what extent self-employment fits the description of the 'precariat' and thus could be described as a highly uncertain and low-paid activity. However, research seems to indicate that the recent rise of self-employment is intimately related to the increasing flexibilisation of the economy and that part of the self-employed can be characterised as a kind of buffer, or even a contemporary version of Karl Marx's Reserve Armee of workers who can easily be called upon when demand is high.

The demographers Jan Latten and Ingeborg Deerenberg (2013) have shown that Amsterdam, as an international centre of producer services and creative and cultural industries, attracts most of the workers from western countries within the Netherlands. These workers are, on average, relatively well paid (e.g. expats working with multinationals) thereby significantly contributing to the local economy.

Conclusions

Amsterdam has shown a similar development pattern to many other (capital) cities. The population decline which started in the 1960s as manufacturing shed jobs and middle-class people left the city was partly countered by the arrival of lesser educated migrants from non-western countries seeking work and a better life in the west. With the transformation of the production system after 1985 and a shift towards cognitive-cultural activities requiring highly skilled workers, the city changed significantly. Central urban locations became attractive again for both firms and residents thereby altering the face of many urban neighbourhoods. The emerging production system also fostered contacts and ever more flows of highly skilled workers between different advanced urban economies. Migration before 1990 was confined to people from less-developed countries, but after that increasing flows of mainly highly skilled workers (and students) came from other advanced economies between global nodes.

Amsterdam, a second or third-tier global city, but nonetheless a global player in particular producer services, creative and cultural industries, and tourism (recently ranked, together with the rest of the Dutch Randstad as the 16th most influential city in the world (Kotkin, J. 2014), basically fits this pattern. The city can be labelled as a highly developed cognitive-cultural urban economy with a very diverse population.

A closer look at the relationship between migration and the processes of urban transformation suggests that there are no simple, clear-cut answers to the questions regarding this relationship. The more generic structural economic changes intersect with existing more specific local socio-cultural and institutional contexts. The labour market statistics revealed a city with very evident ethnic fault lines especially between, on the one hand, migrants from non-western countries and, on the other, those from western countries and the Netherlands itself. Surinamese, Turks and especially Moroccans, although the majority were born or have lived most of their lives in Amsterdam, still find themselves at the end of the hiring queue. This is also true if they have higher educational qualifications. The deeply rooted division between insiders and perceived outsiders, a dominant trait of Dutch corporatist society for much of the post-war period, apparently still holds sway (Kloosterman, 1994). Moreover, it is very hard not to suspect discriminatory practices on the part of employers towards these groups. Even at the height of the economic boom, unemployment rates among these categories were still significantly higher than among native Dutch and migrants from western countries who do not seem to suffer from negative stereotyping. The danger then exists that some may become less attached to the labour market and develop subcultures, which are even less conducive to getting a job. The process might then feed off itself and create an ever greater distance between these detached subcultures and more mainstream ones. A self-reinforcing process is initiated which is in many ways similar to what William Julius Wilson (2011) has described regarding unemployed Afro-Americans in American cities, but in this case (a specific interpretation of) religion (Islam) rather more as a defining characteristics than race as such.

Within the highly diverse migrant population, then, a continuum can be found comprising the top-end of highly paid bankers, consultants, lawyers, managers mainly from other western countries and a bottom end with long-term unemployed migrants originating from non-western countries - the Zuidas and the Dappermarkt. What is in between, is less clear; there are some indications of the emergence of a precariat along ethnic lines but the jury is still out, as flexible, low-paid jobs are not confined to ethnic minorities and time may tell who is getting stuck in which jobs.

It is clear, though, that migration and migrants are here to stay. Amsterdam as a minority/majority city is testimony to that. Insertion processes of migrants are obviously dependent on the trajectory of the urban economy which is partly a function of a city's place and role in the global urban hierarchy. They are also determined by the national regulatory framework regarding migration as well as labour and welfare. In addition, the specific history of a city as embodied in the minds of people, its local traditions and its built environment also impacts on processes of insertion (Kloosterman, 2010b). These factors determine the variations on a common theme of migration and transformation and the more idiosyncratic characteristics of the Amsterdam case can contribute to them.

The analysis above has shown that the pocket-size global city of Amsterdam is increasingly benefiting from migration especially from other western countries. Some groups are however tending to be left behind. Migrants and their descendants from non-western countries who came in the 1960s, 70s and 80s have alarmingly high unemployment rates which point to the existence of an unemployed, ethnic outsider population, which may be even worse than the emergence of a precariat. The costs and benefits of migration are, then, quite unequally distributed. There are clear winners, those working in high-paid jobs, but there are also losers: migrants barely able to get a decent job, but also established citizens who feel they face unfair competition from newcomers or who feel alienated by the rapid demographic changes in their neighbourhoods. It would be hard for a compact and open city like Amsterdam to afford sharp socio-spatial divisions.

The former mayor of Amsterdam, Job Cohen, talked about "keeping the city together". This in a nutshell captures the huge task faced not just by the current mayor of Amsterdam, but probably by also by mayors of other cities. It obviously requires creating the conditions for a thriving urban economy, but it also needs to go beyond seeing it in purely economic terms. Cities should be seen as more than just localised machines of commodification and accumulation. Cities also have, according to Aristotle, a *telos* or goal beyond the economic realm and should provide a home and a haven to its citizens (Sandel, 2010). This requires at the very least some sense of a shared destiny as well as considerable collective action to balance the different claims and to pursue policies of integration. Given the increasing fragmentation of urban (and national) societies along various lines (not just ethnic) and the retreat of the state in the domain of social policies (Streeck, 2014), urban policymakers in Amsterdam and elsewhere face an uphill struggle.

References

Bosscher, D. (2007a) 'De oude en de nieuwe stad', In: P. de Rooy (ed.), *Geschiedenis van Amsterdam; Tweestrijd om de stad, 1900-2000*. Amsterdam: SUN: 337-397

Bosscher, D. (2007b) 'Een stad van en voor wie?'. In: P. de Rooy (Ed.), *Geschiedenis van Amsterdam; Tweestrijd om de stad, 1900-2000*. Amsterdam: SUN: 399-449

Bosscher, D. (2007c) 'Geen woning, toch een kroning; Herwonnen zelfbewustzijn'. In: P. de Rooy (Ed.), *Geschiedenis van Amsterdam; Tweestrijd om de stad, 1900-2000*. Amsterdam: SUN: 451-493

Boterman, W.R. (2012) Residential Practices of Middle Classes in the Field of Parenthood. Dissertation Universiteit van Amsterdam: Amsterdam

Clark, P. (2009) *European Cities and Towns 400-2000*. Oxford: Oxford University Press

Engelen, E. (2007) '"Amsterdamned"? The uncertain future of a financial centre', *Environment and Planning A* (39: 6): 1306-1324

Fernandez, R. (2011) Explaining the Decline of the Amsterdam Financial Centre; Globalizing Finance and the Rise of a Hierarchical Inter-City Network. Dissertation Universiteit van Amsterdam: Amsterdam

Israel, J.I. (1989) *Dutch Primacy in World Trade 1585-1740*. Clarendon Press: Oxford

Judt, T. (2010), *Ill Fares the Land; A Treatise On Our Present Discontents*. Londen: Allen Lane

Kloosterman, R.C. (1994) 'Amsterdamned: The rise of unemployment in Amsterdam in the 1980s', *Urban Studies* (31:8): 1325-1344

Kloosterman, R.C. (2010a) 'Matching opportunities with resources: A framework for analysing (migrant) entrepreneurship from a mixed embeddedness perspective', *Entrepreneurship and Regional Development* (22:1): 25-45

Kloosterman, R.C. (2010b) 'This is not America: embedding the cognitive-cultural urban economy', *Geografiska Annaler: Series B, Human Geography* 92 (2): 131-143

Kloosterman, R.C. (2013) "The Amsterdam Economy and Its Impact on the Labor Market Position of Migrants, 1980–2010." In: N. Foner, J. Rath, J.W. Duyvendak, and Rogier van Reekum (eds.) *Immigration and the New Urban Landscape, New York and Amsterdam*. New York: NYU Press: 107-121

Kotkin, J. (2014) The world's most influential cities in New Geography 15th August http://www.newgeography.com/content/004475-the-worlds-most-influential-cities accessed 28/08/20014

Latten, J. and Deerenberg I. (2013) 'Nieuwkomers grote steden per saldo economisch sterker', Bevolkingstrends 2013. Centraal Bureau voor de Statistiek, http://www.cbs.nl/NR/rdonlyres/14C473A8-FDDD-4A9D-A485-58746C39AA1D/0/20131004b15art.pdf

Mak, G. (1995) *Een kleine geschiedenis van Amsterdam*. Atlas: Amsterdam/Antwerpen

Marlet, G. (2009) De aantrekkelijke stad; Moderne locatietheorieën en de aantrekkingskracht van Nederlandse steden. Dissertation Universiteit van Utrecht: Utrecht

Monitor jeugdwerkloosheid Amsterdam 2014 (2014) Factsheet Nummer 6, juni 2014: Gemeente Amsterdam

Obdeijn, H. and M. Schrover (2008) Komen en gaan; Immigratie en emigratie in Nederland vanaf 1550. Uitgeverij Bert Bakker: Amsterdam

O+S (2013) *Jaarboek 2013*. http://www.os.amsterdam.nl/publicaties/

O+S (2014a) http://www.os.amsterdam.nl/feiten-en-cijfers/#

O+S (2014b) *Jaarboek 2014*. http://www.os.amsterdam.nl/publicaties/

Prak, M. (2002) Gouden Eeuw; Het raadsel van de Republiek. Uitgeverij Sun: Nijmegen

Sandel, M. (2010) *Justice. What's the Right Thing To Do?*. Penguin Books: London

Sassen, S. (1991/2001) *The Global City. New York, London, Tokyo*. Princeton University Press, New Jersey

Scott, A. J. (2008) *Social Economy of the Metropolis: Cognitive-Cultural Capitalism and the Global Resurgence of Cities*. Oxford University Press, Oxford

Scott, A.J. (2012) *A World in Emergence; Cities and Regions in the 21st Century*. Cheltenham (UK)/Northampton (US): Edward Elgar

Streeck, W. (2014) 'How will capitalism end', *New Left Review* (87): 35-64

Vertovec, S. (2007) 'Super-diversity and its implications', *Ethnic and Racial Studies*, 30(6), 1024-1054

Wilson, W. J. (2011) *When work disappears: The world of the new urban poor*. Random House LLC

OPEN ACCESS Freely available online

Where Am I? Location Archetype Keyword Extraction from Urban Mobility Patterns

Vassilis Kostakos*, Tomi Juntunen, Jorge Goncalves, Simo Hosio, Timo Ojala

Department of Computer Science & Engineering, University of Oulu, Oulu, Finland

Abstract

Can online behaviour be used as a proxy for studying urban mobility? The increasing availability of digital mobility traces has provided new insights into collective human behaviour. Mobility datasets have been shown to be an accurate proxy for daily behaviour and social patterns, and behavioural data from Twitter has been used to predict real world phenomena such as cinema ticket sale volumes, stock prices, and disease outbreaks. In this paper we correlate city-scale urban traffic patterns with online search trends to uncover keywords describing the pedestrian traffic location. By analysing a 3-year mobility dataset we show that our approach, called Location Archetype Keyword Extraction (LAKE), is capable of uncovering semantically relevant keywords for describing a location. Our findings demonstrate an overarching relationship between online and offline collective behaviour, and allow for advancing analysis of community-level behaviour by using online search keywords as a practical behaviour proxy.

Citation: Kostakos V, Juntunen T, Goncalves J, Hosio S, Ojala T (2013) Where Am I? Location Archetype Keyword Extraction from Urban Mobility Patterns. PLoS ONE 8(5): e63980. doi:10.1371/journal.pone.0063980

Editor: César A. Hidalgo, MIT, United States of America

Received January 16, 2013; **Accepted** April 7, 2013; **Published** May 21, 2013

Funding: This work was partially funded by the Academy of Finland, TEKES - the Finnish Funding Agency for Technology and Innovation, the City of Oulu, and the commercial partners of the UbiOulu program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The list of commercial partners of the UbiOulu program include Nokia, Bluegiga Technologies, Liidea Ltd. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: vassilis@ee.oulu.fi

Introduction

Research has shown that the study of urban mobility can offer insight into human behaviour [1–3]. One mechanism that makes this possible is the concept of a *location attractor* [4] in urban environments, i.e. that certain locations attract individuals exhibiting a particular behaviour deviating from routine movement. For instance, a sports stadium becomes such an attractor when it hosts sports events, causing individuals interested in sports to alter their daily routine and visit the stadium. Along the same lines researchers have also attempted to understand the relationship between mobility and behaviour by linking mobility traces corpuses to geospatial configuration [5].

However, capturing urban mobility traces is challenging and often requires infrastructure that can be expensive and complex to maintain. For this reason, researchers have increasingly studied online behaviour datasets to uncover patterns in collective human behaviour. Online datasets are convenient to capture and analyse, and previous work has already noted their potential. For example, trends in online search activity have been shown useful in providing models of real world phenomena such as influenza outbreaks, stock market activity, and consumer behavior [6–10]. One example is the search volume of the keyword "flu" which correlates with the outbreak of influenza as recorded in national health statistics [6,8,11]. These models typically rely on the careful choice of search keywords that are expected to correspond with a particular real-world phenomenon.

In this paper we set out to investigate the relationship between urban mobility and online keyword search volume. Because *both* urban mobility [1–5] and keyword search volume [6–10] correlate with collective behaviour, we hypothesise that they may be linked directly. Evidence of such a relationship would suggest that urban mobility can be studied by using online data as a proxy, provided that the appropriate keywords can be identified.

In this paper we make a number of contributions. Specifically:

- We show that urban mobility patterns can be modelled using a set of search keywords. This means that the mobility at particular locations can be modelled using the popularity of certain keywords over time.
- We show that these keywords appear to be semantically relevant to the respective locations. This suggests a strong relationship between urban mobility and online search behaviour.
- We fully describe our process (called Location Archetype Keyword Extraction, or LAKE), which relies on publicly available tools. This means that other researchers can immediately validate our approach by using LAKE with their own data.
- We demonstrate that the reason for LAKE's effectiveness is the existence of location archetypes. This means that mobility across semantically relevant locations correlates highly, and therefore LAKE identifies keywords that are relevant to semantically similar locations.
- We also provide evidence that the correlation of mobility at any two locations is also geographically bound, specifically inversely related to their distance.

The implications of our findings are diverse. First, we provide evidence of a link between urban mobility and online search behaviour. Furthermore, our work presents a cheap way to estimate pedestrian mobility at urban locations using search keywords as a proxy. Finally, our findings show that our method can also characterise urban locations given mobility data, which is a valuable input for interactive systems operating in urban environments.

Materials and Methods

Data Provision

Our approach uses online search data provided by Google. According to analytics firm NetMarketShare, Google's share of global online searches was 84.14% in November 2012, making it the biggest online search engine in terms of volume of executed searches. Thus, we argue it provides the most comprehensive search data for us.

Google Trends is a publicly available online tool by Google that provides insights into temporal and spatial volumes of search queries from the search engine. Its main purpose is to model over time the popularity of keywords used in search queries. Building on top of this tool, *Google Correlate* can identify web search queries whose temporal frequency best correlates with any user-specified temporal pattern [12]. Google Correlate requires a time series data in date-value pairs as an input. The format of the input data is a two-dimensional vector of form *[dd/mm/yyyy, value]*. The tool then attempts to find search queries whose temporal patterns match the input data. The tool operates on a national level, such that the popularity within a particular nation is considered.

We independently collected urban mobility traces using a network of WiFi access points across our city. An access point was installed in each of the locations in our study, and we calculated the number of unique devices detected by the access point per day of the study. We normalise this data and use it as an approximation of the pedestrian flows in various locations. One pedestrian flow vector can therefore be used to describe pedestrian activity at one location for the duration of our study. Using one such vector as an input, the Google Correlate tool provides a list of the top-10 search keywords whose popularity best matches the pedestrian flow vector over time. The top search keywords are ranked in terms of the correlation's r coefficient.

Data Collection

To investigate the relationship between urban mobility and online search activity we collected a longitudinal dataset of city-scale mobility, and used the Google Correlate tool to match time-series mobility data with search keywords. We collected mobility traces using a municipal WiFi network between January 2009 and May 2012. The network consists of 1300 access points covering a large portion of the area of the city of Oulu, Finland (shown in Figure 1). The network is free, open, and available in a number of public locations including schools, libraries, health clinics, and public squares.

We also calculated the daily number of visitors at a variety of locations by considering the number of unique devices that used the WiFi network at a particular location. For example, Figure 2 shows the number of unique devices detected in our University over a period of 4 weeks in 2009. This analysis is possible because each device uses a unique identifier, and each of the 1300 access points is uniquely identifiable by our software. Each device that accesses the network is assigned a unique identifier (ID). If a device connects to the network for the first time, the device ID will be determined based on the MAC address of the device using a one-way hash function. When a device connects to the network, our software records the device ID, the date and time, and ID of the WiFi access point that can be used to infer the location. The ID of a device or access point does not change over time. Thus, we were able to determine how many unique devices have accessed the network at specified location on each day for a given time period. A limitation of this approach is that if a user has multiple WiFi devices, our analysis will infer that multiple users have accessed the network.

Some locations have multiple overlapping access points installed to cope with high network demand. In these cases, our analysis treats those access points as an aggregate "virtual" access point. This ensures that a device is only recorded once in that area or building, even though it may actually associate with multiple access points. We manually inspected the locations in our analysis and we defined virtual points in this manner. For example, all access points across the 3 floors of the library were aggregated into a single virtual access point, and similarly the lobbies in our university were all aggregated into a single virtual access point. The same is true for the access points high schools, where we coded the data so that each high school is a single virtual point. Overall, we defined 12 virtual access points in our analysis representing 84 real access points.

Visual Interpretation of the Data

We are able to generate time-series data describing the number of visitors across the city as well as at a particular location, for example at our University. The total number of devices detected across the city during the whole study is shown in Figure 3, and clearly demonstrates an upward trend underlying the varying seasonal patterns. We hypothesize that this upward trend is due to the increasing penetration of WiFi enabled lightweight devices such as smartphones and tablets during the study period [13].

The time-series data for a particular location, our University, is shown in Figure 4 along with a normalisation process aimed at retaining the seasonal patterns but discarding the overarching upward trend in the number of WiFi devices detected. This upward trend affects the results obtained in our analysis, because in addition to the seasonal patterns it also takes into account the upward trend of the data. As a result, search queries that have become increasingly popular since 2009 are ranked higher when we conduct our analysis using just the raw pedestrian flow data. Because our analysis is more focused on the seasonal patterns of the data, we perform the normalization process described next.

Data Normalization

We illustrate our normalisation procedure using the data collected from the main lobby areas of our University. Figure 4 visualizes the effects of our normalization method. The pedestrian flow data shown in Figure 4a was collected in our University, and shows that the beginning of the semester in autumn is relatively busy when freshmen start their studies, while the amount of students decreases gradually during the year. A sharp drop in the amount of pedestrians can also be seen during the Christmas holidays.

To filter out the upward trend component and maintain the seasonal patterns, we denote the university time-series data as a vector K, where a value K_i is the number of devices detected on i^{th} day of our study in that location. The vector T is a vector where value T_i is the total amount of devices detected on i^{th} day in the whole city.

Therefore, the percentage of devices across the whole community that were detected in the university is

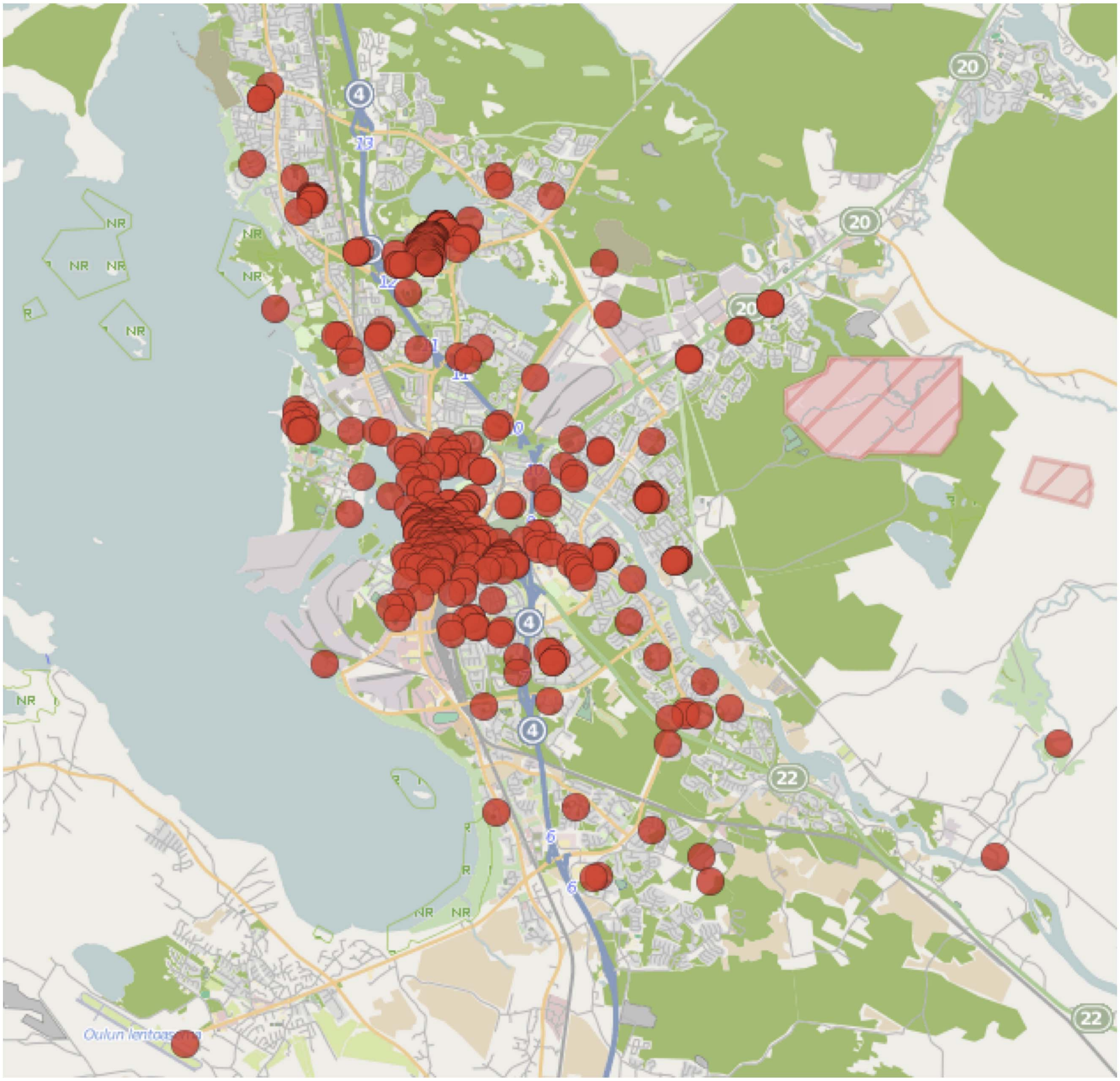


Figure 1. A map showing all the WiFi access point locations. The map covers an area approximately 20km×20km.
doi:10.1371/journal.pone.0063980.g001

$$W = \frac{K_{uni}}{T}$$

where W is a vector containing the relational frequency of the detected devices for each day of the study. The result of this normalization is shown in Figure 4b but still entails an upward trend. Given the strong annual pattern across all our data, we apply minimum-maximum normalization on each year separately to highlight the seasonal variations. The resulting normalized vector can be denoted as

$$\{W_{2009},\ldots,W_{2012}\}$$

where W_{year} is a minimum-maximum normalized annual subset from the vector W and is denoted as

$$W_k = \frac{W_i\text{-}\min(W_k)}{\max(W_k)\text{-}\min(W_k)},$$

where $k = \{2009,2010,2011,2012\}$

This normalization step further reduces the upward trend component, and the results are shown in Figure 4c.

To validate our normalization method we rely on linear regression and inspect the slope x_i, r^2 and *p-value* of the regression lines in Figures 4a, 4b and 4c. This approach was used to validate whether the upward trend component, estimated with a linear

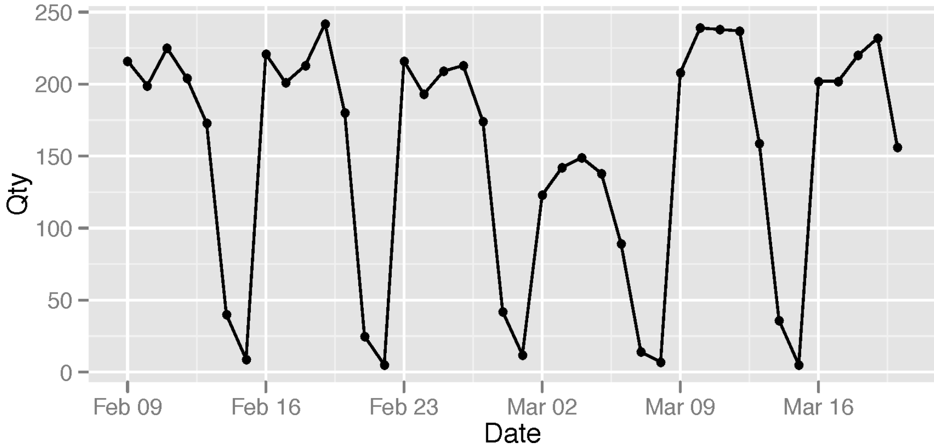


Figure 2. Total number of unique devices detected at one location. The data demonstrates how over a period of 4 weeks in 2009 the number of devices fluctuates on a daily pattern. The first data point in this plot is a Monday. The data also shows that weekends have distinctly less activity.
doi:10.1371/journal.pone.0063980.g002

regression line, can be progressively eliminated. Our aim is to derive a "stable" oscillating function with a linear regression line whose slope is near zero and has low r^2 value. The results of this validation process are presented in the results section. Note that our intention is to verify that a linear regression has a *poor* fit, thereby suggesting that the linear trend in the data has been largely eliminated.

Keyword Extraction

Using our normalisation approach we derive one time series per location of interest in the city, and use each one to identify search queries with similar temporal patterns on Google Correlate. Effectively, we uploaded pairs of values of the format <date, value>, using each day as a data point. Google correlate aggregates this data into weekly values. Our analysis results in the top 10 online search keywords used in Finland whose search volume time series matches the pattern of pedestrian flows (using Pearson's correlation) at each location respectively.

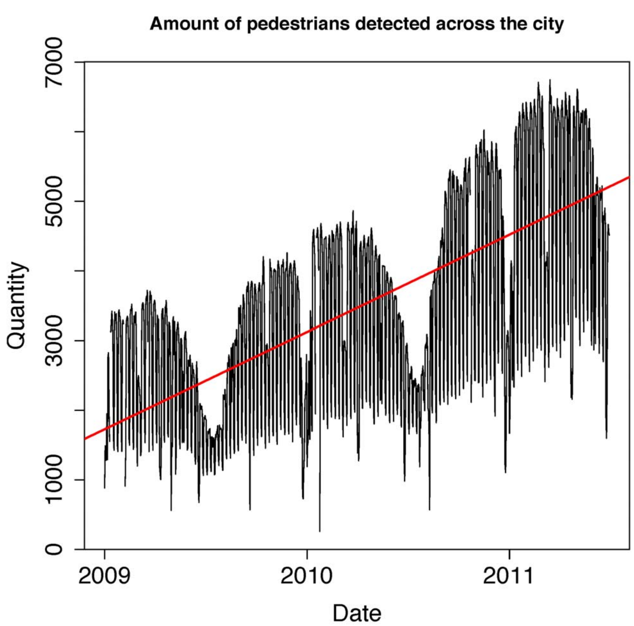


Figure 3. Total number of unique devices detected in the city during the study period. The data demonstrates how over a period of three years the volume of devices doubles, much unlike the population of our city that has grown at more modest rates.
doi:10.1371/journal.pone.0063980.g003

Our results show that, when appropriate keywords are used, the pedestrian flows at a location can be approximated with a degree of accuracy of 0.7–0.9. However, our results did have an unexpected element. Surprisingly, the results were keywords that appeared to be semantically relevant to the locations. For example, given the time series from the University data (Figure 4c) some of the keywords we obtained are "research", "scholar" and "lecture" (Table 1). These terms were derived by feeding the normalised data for each location to the Google Correlate tool and deriving the top-10 search terms whose popularity matches the pedestrian flow volume. We refer to this whole process as Location Archetype Keyword Extraction (LAKE).

To verify that the results of LAKE were indeed semantically relevant to the respective locations, we issued a questionnaire to 29 residents of our city (22 male, 7 female) with an average age of 30.5 ($\sigma = 9.5$). Each respondent was given the following:

- a set of location names (High school, University, Library, Camping, Ice Hockey Hall)
- a set of 16 keywords
- instructions to rate on a 5-point likert scale the semantic similarity of each keyword to the respective location.

We constructed the set of words for each location by aggregating and mixing.

- the top 10 keywords for that particular location derived using our analysis
- 3 words selected at random
- 3 keywords from a different location.

Thus, each of the 15 respondents was given 5 locations and 16 words per location. The order of words for each location was randomized. Respondents rated each word on a 5-point Likert scale to indicate the semantic relevance of that word to the respective location. No information was given to respondents about how the words were derived, the respondents were recruited

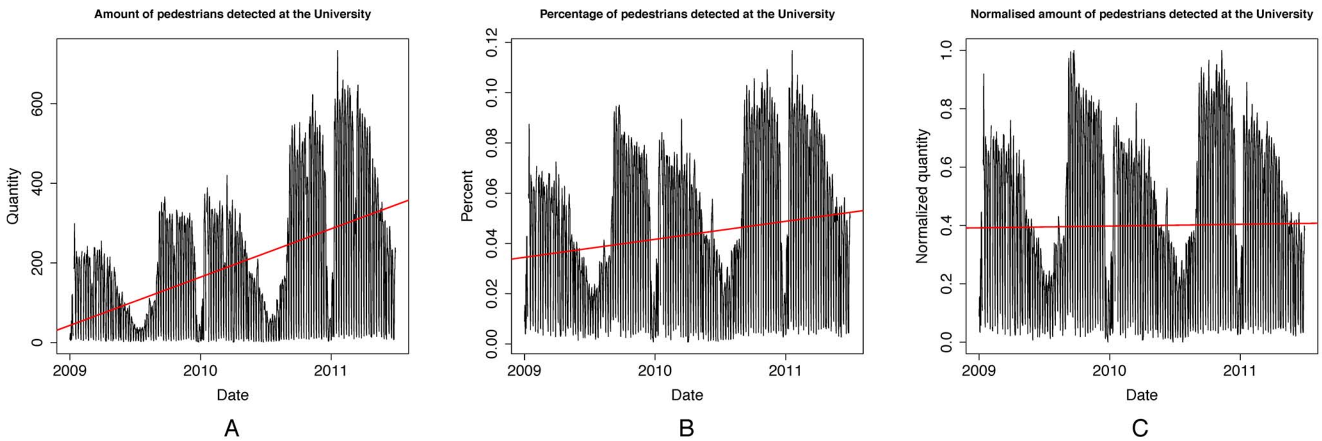


Figure 4. Normalisation of pedestrian flows time series. Figure 4a shows the number of WiFi devices detected during the study period in the lobby areas of our university. The data demonstrates strong seasonal patterns, for example showing that summer time is relatively quiet at the University. At the same time, however, this time series shows an upward trend, reflecting the increasing number of WiFi devices using the network across the city over time. To account for this increasing trend while maintaining the seasonal patterns, we normalise the data using a two-step approach. First we calculate the percentage of all devices seen on any day that visit a particular location on that day. Results of this normalization step are shown in Figure 4b. Subsequently, we apply an annual minimum-maximum normalisation filter to derive the time series in Figure 4c. The effectiveness of our normalization method is verified using linear regression (in red), which shows that compared to the original data ($x = 0.3324$, $r^2 = 0.2192$, $p < 2.2e\text{-}16$) the normalization in Figure 4b ($x = 1.964e\text{-}05$, $r^2 = 0.04205$, $p = 3.318e\text{-}12$) and subsequently Figure 4c ($x = 1.724e\text{-}06$, $r^2 = -0.000879$, $p = 0.9438$) retain mostly seasonal patterns in variation.
doi:10.1371/journal.pone.0063980.g004

via email, and all were natives and residents of our city. The respondents were instructed to consider the broader context of the locations, for instance the search term "marketing" refers to "marketing" as university studies (economics) or "product marketing". The respondents were also encouraged to use online search if they were unfamiliar with any of the queries. Locations were identified only as their categorical names. No more details about the locations or their whereabouts were given. Using this method, we hypothesise that the top 10 keywords included within the scrambled set of 16 words will receive higher ratings than the set of random keywords, suggesting that LAKE produces semantically relevant results. Furthermore, we hypothesize that the top 10 keywords will receive higher rating that the 3 keywords from a different location, suggesting that LAKE distinguishes between locations. As we describe in the results section, both our hypotheses were confirmed by the independent respondents.

Interpretation of LAKE Analysis

The mechanism that enables LAKE to produce semantically relevant keywords – without direct human intervention – is not immediately evident. Besides the lack of direct human input, we were puzzled by the fact that pedestrian flows are collected at a *particular location* in our city, while the Google Correlate tool operates on a *national level*. Given our findings, we hypothesise that there must exist "location archetypes", instances of which are the particular locations we observed with our WiFi network. For example, we hypothesise that our university must be just one of many instances of the "University archetype", and therefore nation-wide online search queries must correlate with the pedestrian flows at all instances of this location archetype. For this reason we refer to the process we have described so far as Location Archetype Keyword Extraction (LAKE), because we hypothesize that it allows us to extract keywords referring to the archetype of locations we study.

To investigate the existence of location archetypes, we analyse longitudinal pedestrian flows collected from various locations that we intuitively believe must belong to the same archetype: five high schools in our city. The correlation of daily visitors at a number of unrelated locations is shown in Figure 5a and contrasted with the data for five distinct high schools in Figure 5b. The results demonstrate that pedestrian flows at all the high schools correlate strongly with each other ($r > 0.8$), and thus are likely instances of a "High School archetype". Subsequently, the pedestrian flow data from all high schools produce semantically relevant keywords using our LAKE analysis.

This analysis relies on daily flows of pedestrians at various locations. In our analysis we also considered using hourly-intervals to compare locations, but the results were unreliable. When considering hourly intervals, the Circadian rhythm becomes very prominent. In other words, the "nominal" daily patterns take over, with high activity during work hours and low activity during nighttime and weekends. For this reason the analysis results in high correlation patterns between most locations, and effectively reduces the ability of our technique to differentiate between locations.

Distance Effect on Correlation

An alternative explanation for the results in Figure 5 on the correlation of pedestrian flows may be distance: the high correlation in urban mobility between two locations may be due to their physical proximity. This is a spatial phenomenon that has previously been hypothesised [4]. We investigate whether the strong correlations we identified may be due to spatial proximity in addition to the existence of location archetypes. Because the LAKE analysis operates on pedestrian flows, we expect that locations close to each other are likely to strongly correlate with each other regardless of their semantic relationship. Effectively, we assume that a group of people moving across the city is likely to register at multiple locations that may not be semantically relevant to each other but be spatially proximal to each other.

To investigate this relationship, we constructed two matrices: a distance matrix M_D shown in Table 2 and a correlation matrix M_C, which can be seen in Table 3. The distance matrix has all the locations of the study denoted with L_i as rows and columns, and

Table 1. Top 10 search engine queries for various locations.

University	r-value	Ice hockey hall	r-value
Helka (Helsinki university library)	0.914	sjl (Finnish ice hockey union)	0.726
google scholar	0.895	finnhockey	0.724
scholar	0.894	keilahalli (bowling hall)	0.722
tutkimus (research)	0.893	kiekko kaleva (hockey in local newspaper)	0.716
learning	0.887	sm-liiga nelonen (tv-channel with hockey)	0.715
optima oulu (student environment)	0.878	nelonen sm (same as above)	0.713
funktio (function)	0.876	jääkiekkoliitto (ice hockey union)	0.711
luento (lecture)	0.874	nelonen sm liiga (tv-channel, ice hockey)	0.707
development	0.872	lihapata (meat stew)	0.699
nelli (university e-library portal)	0.871	finhockey	0.699
Library	**r-value**	**High school**	**r-value**
hietsun kirppis (flea market)	0.889	wilma kempele (student environment)	0.795
pistiäinen (stinging bee)	0.870	wilma kiiminki (student environment)	0.791
viinitila (wine farm)	0.864	helmi (student management system)	0.781
reitti (path)	0.863	wilma oulu (student environment)	0.771
lättähattu (50′s dressing style/old train)	0.862	edu (news and info on education)	0.763
hietaniemen kirpputori (flea market)	0.860	wilma kuusamo (student environment)	0.758
hietalahden kirpputori (flea market)	0.859	pedanet (support for online learning)	0.747
museorautatie (museum railroad)	0.857	wilma raahe (student environment)	0.744
korppoo (island in Turku archipelago)	0.854	wilma kemi (student environment)	0.736
tammisaari (town)	0.853	varoitusmerkit (warning signs)	0.726
Camping	**r-value**		
festivaali (festival)	0.804		
naantali majoitus (housing)	0.793		
kalajoki camping	0.788		
hanko majoitus (housing)	0.761		
vesipuisto serena (water amusement park)	0.760		
rauhalahti camping	0.759		
rengastie kartta (map)	0.758		
Högsåra (island in Hittinen archipelago)	0.756		
kuhan uistelu (zander fishing)	0.740		

The r-values reported are calculated byGoogle Correlate, and here we report the top-10 results for each location. In brackets are English translations where necessary.
doi:10.1371/journal.pone.0063980.t001

their respective distances D_{ij} from each other. The correlation matrix is structured in a similar manner where correlation is denoted with C_{ij}. The diagonals of both matrices are left empty.

We then use these two matrices to construct the scatterplot in Figure 6, where each point in the scatterplot can be denoted with coordinates (D_{ij}, C_{ij}) from the two matrices. In this figure each data point is for a pair of locations. For each pair we use the Pearson's r correlation in pedestrian flows (Table 4) and the physical distance between those two locations in meters (Table 3). We then plot these two values along the *y* and *x* axes of the figure respectively. In addition we use colour coding (green-blue-red) to indicate an increasing expected semantic similarity that we expect for those pairs. Our results confirm that distance is inversely related to the correlation between locations, but only for locations that are *not* semantically relevant. In the latter case, distance has little impact on the correlation.

Results

In Figure 3 we verify that the total number of unique WiFi devices detected across the city has indeed an upward trend highlighted in red ($x_a = 3.383$, $r^2_a = 0.3125$, $p_a < 2.2e\text{-}16$), and we also verify that this trend is present in the University raw data in Figure 4a ($x_b = 0.3324$, $r^2_b = 0.2192$, $p_b < 2.2e\text{-}16$). Our first normalisation step (Figure 4b) partially filters this trend by relying on the percentage of devices detected in the University area for each day of the study ($x_c = 1.964e\text{-}05$, $r^2_c = 0.04205$, $p_c = 3.318e\text{-}12$). Our final normalisation step (Figure 4c) utilizes the annual minimum-maximum normalization and adequately filters the upward trend by resulting in a non-significant linear fit ($x_d = 1.724e\text{-}06$, $r^2_d = -0.000879$, $p_d = 0.9438$). We summarize the regression coefficients for each dataset in Table 4.

This normalization procedure shows that by calculating the total percentage of the devices detected in the university in relation

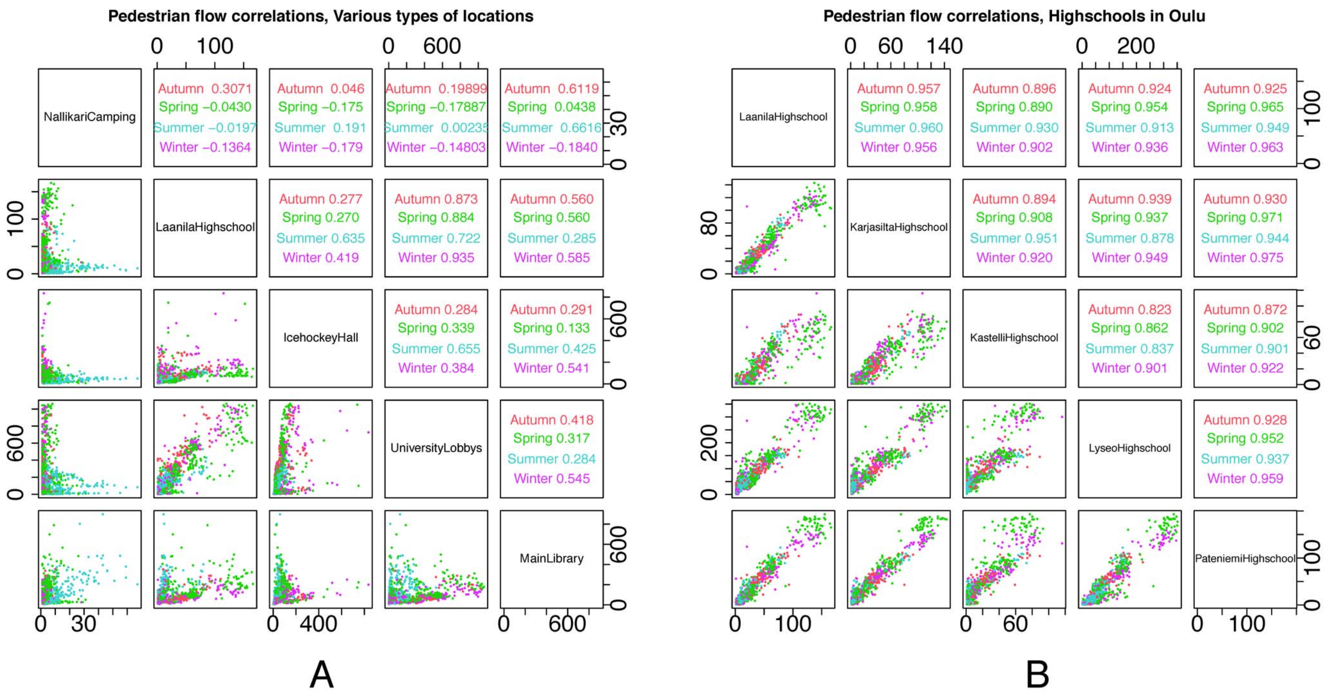


Figure 5. Scatterplot matrices showing pedestrian flow. Here we can see pedestrian flows for (a) various types of locations and (b) high schools. In this figure each scatterplot is for a pair of locations. Each scatterplot data point is for a particular day of our study, and indicates the correlation of pedestrian flows for the two locations on that particular date. Data points are color-coded by season to account for seasonal variations. Reading the scatterplots: to locate the scatterplot for a pair of locations, the row-column intersection cell need to be inspected below the diagonal. Similarly, the correlation values are at the row-column intersection cell above the diagonal.
doi:10.1371/journal.pone.0063980.g005

to the whole city network, and by then performing annual minimum-maximum normalization we can filter the trend component in the data ($x_b > x_c > x_d$ and $r^2_b > r^2_c > r^2_d$). By filtering the trend component we highlight the underlying seasonal patterns of pedestrian flows in the original pedestrian flow data and thus are able to identify search engine queries and keywords that fluctuate with seasons, rather than queries that are trending over time.

The LAKE analysis was also applied to five locations of semantically different archetypes. The results of the human assessment of the search queries can be seen in Table 5. The keywords obtained through our analysis were rated consistently more relevant one a 5-point Likert scale ($\alpha = 0.982$, $R_{avg} = 3.407$) than random keywords ($\alpha = 0.948$, $R_{avg} = 1.620$) or keywords from different locations ($\alpha = 0.977$, $R_{avg} = 2.124$). Our analysis found semantically relevant keywords for a number of locations such as high schools, a camping site, the university, and an ice hockey arena.

Figure 5 demonstrates the effect of the physical distance between pairs of locations on their correlation of pedestrian volumes. In green we show location pairs that are not semantically relevant, which as a set demonstrate an inverse effect between distance and correlation of pedestrian volumes. In blue we show location pairs that are semantically related to each other (pairs consisting of the university and one high school). Finally, in red we show location pairs that belong to the same location archetype (pairs of high schools).

We find that for location pairs that are not semantically relevant there exists an inverse effect between distance and correlation of pedestrian volumes. This is not the case for pairs of locations belonging to the same location archetype, or being semantically relevant. The results show that for the green set distance has an inverse effect on pedestrian flows correlation volumes ($x_{random} = -6.285e\text{-}05$, $r^2_{random} = 0.11$, $p_{random} = 0.007$). The effect of distance is minimal in the blue and red groups ($x_{university} = 2.647e\text{-}07$, $r^2_{university} = -0.327$, $p_{university} = 0.910$; $x_{highschools} = 4.075e\text{-}06$, $r^2_{highschools} = 0.051$, $p_{highschools} = 0.172$).

Finally, we constructed a scatterplot for more than 30000 random pairwise combinations of locations from our dataset. Figure 7 shows the results, which confirm an inverse underlying relationship between physical distance and pedestrian flow correlation between pairs of locations ($x_{random} = -1.188e\text{-}02$, $r^2_{random} = 0.01012$, $p_{random} < 2.2e\text{-}16$). The scatterplot contains some vertically clustered set of points at distance = 10, 12 and 15 kilometres, highlighting spatial clusters of WiFi access points and the polycentric nature of the region.

Discussion and Conclusion

Our results demonstrate two important findings. First, we demonstrate that LAKE can be used to identify keywords that

Table 2. Distance matrix for all locations.

Location	**L_1**	**L_2**	**…**	**L_j**
L_1	–	D_{12}	…	D_{1j}
L_2	D_{21}	–	…	D_{2j}
…	…	…	–	…
L_i	D_{i1}	D_{i2}	…	–

Each cell D_{ij} denotes the distance in meters between locations L_i and L_j.
doi:10.1371/journal.pone.0063980.t002

Table 3. Pedestrian flow correlation Matrix for all locations.

Location	L_1	L_2	...	L_j
L_1	–	C_{12}	…	C_{1j}
L_2	C_{21}	–	…	C_{2j}
...	…	…	–	…
L_m	C_{i1}	C_{i2}	…	–

Each cell C_{ij} denotes the Pearson's correlation in daily pedestrian flows between locations L_i and L_j.
doi:10.1371/journal.pone.0063980.t003

Table 4. Summary of the linear regression analysis.

Figure	r^2	x_i (slope)	p-value
3	0.31250	3.3830	<2.2e-16
4a	0.21920	0.3324	<2.2e-16
4b	0.04205	1.964e-05	3.318e-12
4c	−0.00879	1.724e-06	0.9438

Values of the normalization procedure from each dataset in Figures 3 and 4.
doi:10.1371/journal.pone.0063980.t004

strongly correlate with urban mobility at particular locations. This means that subsequently monitoring the popularity of these keywords can indeed offer insights into the pedestrian flows of that particular location. An important benefit of this approach is that it can be cheap and convenient to collect data on keyword popularity on large scale, thus making this analysis more accessible.

A second, unexpected, finding of our work is that the keywords produced by LAKE are semantically relevant to the respective locations. Verification with human raters suggests that indeed the keywords LAKE produced are semantically relevant to the locations. In some cases this semantic relevance is contextually derived, so we would expect only a resident of our city to be able to identify this relationship.

Furthermore, our analysis of the distance between locations and their mutual correlation shows that locations of the same archetype correlate strongly with each other regardless of the distance between them, while random locations show declining correlation as their distance increases. These findings suggest that pedestrian correlations are not only affected by seasonal variations, but show also causal relationships in a local scale.

We note that do not argue that our technique can replace existing ways of studying urban mobility. In fact, our technique relies on the existence of detailed mobility data – if this data exists in the first place, it can be used directly to conduct robust mobility analyses. What our technique complements existing urban mobility tools in two ways. First it allows us to identify keywords that can subsequently be monitored more easily than collecting urban mobility data. Second, it provides a way to derive "qualitative" data about the location in question, by generating keywords that are semantically relevant to this location. In some ways, these keywords can be understood as the broader context of the environment in which the data was collected.

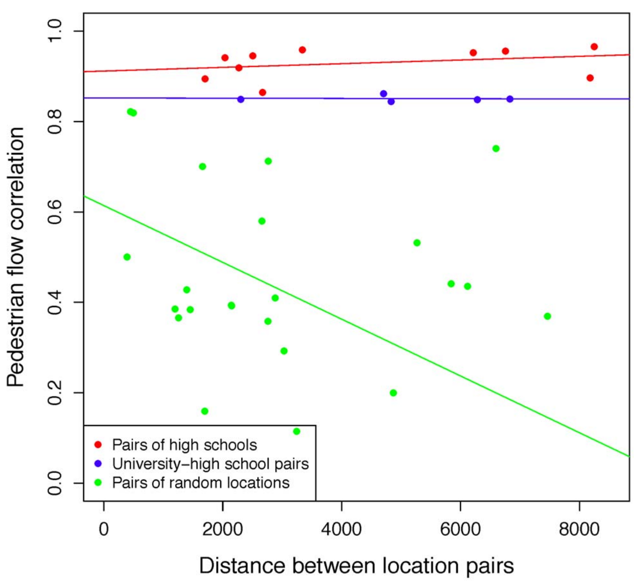


Figure 6. Distance affects pedestrian flow correlations. Correlation in pedestrian flows is affected by distance (in meters) between two locations. Orange dots are pairs of high schools, and blue dots are pairs consisting of the university and high schools. Green dots are pairs of semantically irrelevant locations. Regression lines are included with the colour of the respective category. We identify two trends in this data. With the green colour we show location pairs that are not semantically relevant, which demonstrate an inverse effect between distance and correlation of pedestrian volumes ($x_{random} = -6.285e\text{-}05$, $r^2_{random} = 0.11$, $p_{random} = 0.007$). In orange we show location pairs that are semantically relevant (pairs of high schools) and in blue we show location pairs that are highly related to each other (pairs consisting of the university and one high school). We find that for both sets of pairs distance has no significant effect on the pair's correlation of pedestrian flows ($x_{university} = -2.647e\text{-}07$, $r^2_{university} = -0.327$, $p_{university} = 0.910$; $x_{highschools} = 4.075e\text{-}06$, $r^2_{highschools} = 0.051$, $p_{highschools} = 0.172$).
doi:10.1371/journal.pone.0063980.g006

How can Location-bound Data Correlate with National-level Data?

It is not clear why location-bound pedestrian flows correlate with search engine query data, which is collected at a national level. The collected location-bound pedestrian flows represent a non-negligible part of the population of our country and the city. Even if other potential queries are correlated and compared with our data, only the relevant search queries show significant correlations ($r > 0.8$).

Therefore, we find these results meaningful and believe that this phenomenon indeed demonstrates a latent relationship between online and offline community behaviour in terms of spatial and temporal activity. Specifically, we believe that the locations act as population attractors, in which the visitors' movement frequencies and patterns create a unique "fingerprint" for each location, which becomes the main factor in the correlation process. These unique, archetype-specific pedestrian flow characteristics can then be identified using LAKE.

The findings of this paper would clearly be stronger if a larger number of location archetypes were investigated. Unfortunately there are no more location archetypes included in our data that can be reliably analyzed. This is because our data is captured from a public WiFi network that is only deployed in public spaces and governmental buildings. Many of these buildings house a mixture of facilities or services, and therefore they data we collect at those locations are not "clean". The archetypes we have presented in this paper are all locations that have a rather well-defined purpose, and therefore our results can be reliably verified.

Table 5. Results of the questionnaire on LAKE analysis.

Location	Wordset (words)	Cronbach's α	Average relevance (1–5)
Highschool	LAKE analysis (10)	0.966	3.662
Highschool	Random (3)	0.971	2.080
Highschool	LAKE from other locations (3)	0.995	2.816
Highschool	All words (16)	0.979	3.207
University	LAKE analysis (10)	0.890	4.479
University	Random (3)	0.862	1.391
University	LAKE from other locations (3)	0.975	2.678
University	All words (16)	0.986	3.563
Library	LAKE analysis (10)	0.768	1.479
Library	Random (3)	0.528	1.437
Library	LAKE from other locations (3)	0.402	2.609
Library	All words (16)	0.912	1.683
Camping	LAKE analysis (10)	0.966	3.610
Camping	Random (3)	0.969	1.701
Camping	LAKE from other locations (3)	−0.230	1.241
Camping	All words (16)	0.984	2.808
Ice hockey hall	LAKE analysis (10)	0.989	3.803
Ice hockey hall	Random (3)	0.955	1.494
Ice hockey hall	LAKE from other locations (3)	0.561	1.275
Ice hockey hall	All words (16)	0.993	2.897
All	**LAKE analysis (10*5)**	**0.982**	**3.407**
All	**Random (3*5)**	**0.948**	**1.620**
All	**LAKE from other locations (3*5)**	**0.977**	**2.124**
Total	**All words (16*5)**	**0.985**	**2.831**

Here we show for each location three sets of keywords and their respective results. Each row is an individual test case. The interrater agreement (Cronbach's alpha) across all results was $\alpha = 0.976$ suggesting a strong agreement between the raters and the relevance or non-relevance of keywords to all the locations. For most of the cases, respondents agree that the words obtained using LAKE are more relevant to a location than words from random location wordsets or totally random words.

doi:10.1371/journal.pone.0063980.t005

Application of the Keywords' Semantic Relevance

We have described a methodology to "convert" pedestrian flow data [14] into phrases that are semantically relevant to particular areas and contexts and also that can be used as a proxy for estimating the volume of that data. This approach can have a substantial impact on how we design and implement contextual and situated computer systems. These keywords can be used to identify relevant media, services, data, and advertisements in urban computing systems.

One straightforward mechanism to implement this is to actually use web search tools to identify relevant media. For instance, using Google's "image search" with the queries our technique identifies, we believe that we are able to retrieve location images that are semantically relevant to the locations in our study. Similarly, our approach could be used for advertising purposes. For example Google's advertising tools can be used to retrieve advertisements relevant to the keywords identified in our analysis. We believe that these example cases would provide substantial impact on urban pervasive computing and advertising, and therefore they need further studying and assessments by human participants.

Furthermore, our findings also point to a future direction in applying our technique. Using LAKE we can initially assign keywords to each location. Subsequently, we can conduct pairwise comparisons between locations to identify their degree of correlation, as shown in Figures 6 & 7. This degree of correlation can be used as a "weight" parameter to extend each location's set of keywords by incorporating the keywords from the other location, coupled with a weight. Hence, for each location we can derive an extended set of associated keywords, some of which would be weighted. What this approach offers us is a way to extend LAKE by taking advantage of our findings in Figure 7: nearby locations are more likely to correlate with each other. Therefore, the extended set of keywords for a particular location could include keywords that are relevant to nearby locations as well. This is an approach that we intend to evaluate in the future.

Limitations

Our data collection system utilizes hardware MAC addresses to identify unique devices in the network. If a *single* pedestrian is carrying or using *multiple WiFi devices* to access the network, he will be detected as if *multiple* pedestrians in the vicinity of the access point. Similarly, our method does not take into account if the detected device is stationary. For instance, residents living in the vicinity of an AP will be detected and registered as pedestrians every day, even if the AP was accessed from a residence and is stationary for the whole duration of the study. In future work these limitations could be assessed by detecting the direction from which the WiFi connection requests came to the AP and the signal strength of the connection.

Furthermore, our analysis found that adopting a different normalisation algorithm leads to some changes in the extracted keywords, particular beyond the top-20. This is not surprising since a different normalisation process results in changes to the

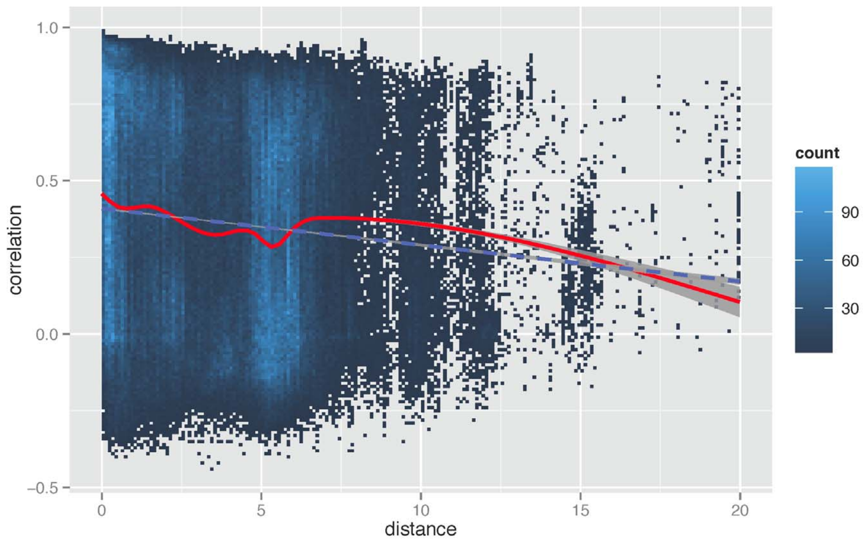


Figure 7. Distance affects pedestrian flow correlations in a city-scale. Correlation in pedestrian flows as affected by distance (in meters) between two locations. Approximately 300000 points contained in this figure, and contain all locations in our dataset. Blue dashed line indicates the regression between the two variables, while the red solid line shows the LOWESS smoother that uses locally-weighted polynomial regression. Both contain a confidence interval at p = 0.99. The results confirm a negative relationship between physical distance and pedestrian flow correlation amongst pairwise locations. The scatterplot contains some clustered sets of points at distance = 10, 12 and 15 kilometres, highlighting spatial clusters of WiFi access points and the polycentric nature of the region.
doi:10.1371/journal.pone.0063980.g007

fluctuations in data. For this reason we have followed a normalisation process described in the section "Data normalization" that does not rely on regression or opaque algorithms, but rather is based on a simple and explicit process that is justified by the nature of the data. Particularly, our normalization process looks at percentages of the whole population that visit a particular location, and also applies an annual min-max normalization step to account for incremental changes to the number of wireless devices in a city over time.

Conclusion

By utilizing the LAKE analysis presented in this study, we have shown a way to algorithmically identify online search queries that are semantically relevant to the location where pedestrian data is collected. In addition, we have shown that locations that are close to each other are likely to have correlating pedestrian flows. As the distance between a pair of locations increases, the correlation between the pedestrian flows remains high if the locations are semantically similar. Otherwise, distance has an inverse effect on pedestrian flow correlation.

Our work enables a new approach to investigating population mobility patterns between online and offline worlds by exploiting automated data collection methods and datasets. Our findings provide insight for modelling and understanding human behaviour as reflected by urban mobility, and our technique provides a mechanism for identifying appropriate search keywords to model arbitrary real-world collective behaviour. Finally, we expect that the LAKE analysis will enable researchers to reconsider the relationship between online and physical patterns, and further contribute to the analysis of community-level behaviour.

Acknowledgments

We would like to thank Vesa Kotila.

Author Contributions

Conceived and designed the experiments: VK SH JG. Performed the experiments: TJ VK JG SH TO. Analyzed the data: TJ VK JG SH. Contributed reagents/materials/analysis tools: TJ VK JG SH. Wrote the paper: VK TJ JG SH.

References

1. Domenico M, Lima A, Musolesi M (2012) Interdependence and predictability of human mobility and social interactions. Nokia Mobile Data Challenge Workshop.
2. Eagle N, Pentland A (2009) Eigenbehaviors: Identifying structure in routine. Behavioral Ecology and Sociology 63(7): 1057–1066.
3. Eagle N, Pentland A, Lazer D (2009) Inferring friendship network structure by using mobile phone data. Proceedings of the National Academy of Sciences 106(36): 15274–15278.
4. Hillier B, Penn A, Hanson J, Grajewski T, Xu J (1993) Natural movement: Or, configuration and attraction in urban pedestrian movement. Environment and Planning B: Planning and Design 20: 29–66.
5. Kostakos V (2009) Space syntax and pervasive systems. Geospatial Analysis and Modeling of Urban Structure and Dynamics : 31–52.
6. Dugas A, Hsieh Y, Levin S, Pines J, Mareiniss D, et al. (2011) Google flu trends: Correlation with emergency department influenza rates and crowding metrics. Clinical Infectious Diseases 54(4): 463–469.
7. Goel S, Hofman J, Lahaie S, Pennock D, Watts D (2010) Predicting consumer behavior with web search. Proceedings of the National Academy of Sciences 107(41): 17486–17490.
8. Ginsberg J, Mohebbi M, Patel R, Brammer L, Smolinski M, et al. (2008) Detecting influenza epidemics using search engine query data. Nature 457: 1012–1014.
9. Rui H, Liu Y, Whinston A (2011) Whose and what chatter matters? the impact of tweets on movie sales. NET Institute. In Press.
10. Bollen J, Mao H, Zeng X (2011) Twitter mood predicts the stock market. Journal of Computational Science 2(1): 1–8.
11. Polgreen P, Chen Y, Pennock D, Nelson D (2008) Using internet searches for influenza surveillance. Clinical Infectious Diseases 47: 1443–1448.
12. Mohebbi M, Vanderkam D, Kodysh J, Schonberger R, Choi H, et al. (2011) Google correlate whitepaper.
13. Official Statistics of Finland (OSF): Use of information and communications technology [e-publication]. Helsinki: Statistics Finland 2012.
14. Kostakos V, O'Neill E, Penn A, Roussos G, Papadongonas D (2010) Brief encounters: sensing, modelling and visualizing urban mobility and copresence networks. ACM Transactions on Computer Human Interaction, 17(1): 1–38.

Revisitation Analysis of Smartphone App Use

Simon L. Jones
University of Bath
Bath, UK
s.l.jones@bath.ac.uk

Denzil Ferreira, Simo Hosio, Jorge Goncalves, Vassilis Kostakos
University of Oulu
Oulu, Finland
{denzil.ferreira, simo.hosio, jgoncalv, vassilis}@ee.oulu.fi

ABSTRACT
We present a revisitation analysis of smartphone use to investigate the question: do smartphones induce usage habits? We analysed three months of application launch logs from 165 users in naturalistic settings. Our analysis reveals distinct clusters of applications and users which share similar revisitation patterns. However, we show that much of smartphone usage on a macro-level is very similar to web browsing on desktops, and thus argue that smartphone usage is driven by innate service needs rather than technology characteristics. On the other hand, on a micro-level we identify unique characteristics in smartphone usage, and we present a rudimentary model that accounts for 92% in the variability of our smartphone use.

Author Keywords
Revisitation, smartphone use, habits, user behaviour

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
We present a large-scale longitudinal analysis that seeks to quantify smartphone application use habits. A growing number of analyses in recent years have sought to investigate and model how individuals use their mobile phones. This has ranged from studying how individuals charge and manage battery [12], download, install and use different applications [11], and how use varies with context [10]. In terms of *habits* (i.e. patterns, routines) in smartphone application use, previous work has sought to model and predict which applications people are likely to install [37], use [31], and for how long [13].

While previous work has repeatedly shown that individuals exhibit well-structured and predictable patterns in their smartphone use, we are still far from developing a theoretical understanding of why these behaviours arise. This is important because it could provide guidance on the development of new mobile technology and interaction techniques, and on evaluating and comparing existing ones. For instance, while it is commonly accepted that mobile phones are "habit-forming" [25], it is not clear whether these habits are driven by the users themselves, their phones' affordances, or some combination of both. At the same time, little work has quantitatively compared how our smartphone use habits relate to our use habits for other technologies, particularly due to our lack of theoretical models of use.

We address both of these challenges simultaneously, by re-using an established methodology that has been extensively employed in prior literature, albeit in the context of web browsing: *revisitation analysis* [33]. Here we study how users revisit: (1) apps on their phone (i.e. how often do they return to a particular app?). This helps us quantitatively describe every individual app and the type of revisitation it attracts across users; (2) the phone itself (i.e. how often do they use apps on their phone?). This helps us to quantify individual users in terms of their overall revisitation habits; (3) applications in-session (i.e. how often they switch back and forth between active applications between unlocking and locking their phone?). This helps us to quantify backtracking and multitasking on a micro-level.

By re-using an established method to analyse revisitation patterns, we are able to compare revisitation behaviours that individuals exhibit on smartphones against those in desktop-based web-browsing. Our premise is that commonalities found across such different technologies are likely to reflect innate user habits, and therefore we are able to tease apart those habits that appear to be driven solely by smartphone technology.

Ultimately, we seek to make two contributions to the corpus on understanding mobile use. First, we attempt to investigate whether mobile phones are a "special" type of technology in the sense that they are habit-forming and shape our daily routines, or whether humans just exhibit the same habits of use across different technologies. The revisitation analysis method we employ focuses on detecting temporal patterns, hence in this paper we identify and compare habits that are temporal in nature (i.e. behaviours that are triggered after a certain time interval since the phone and its apps were last used).

UbiComp '15, September 7-11, 2015, Osaka, Japan.

http://dx.doi.org/10.1145/2750858.2807542

Second, we provide the tools and techniques for conducting a systematic analysis of revisitation behaviour patterns, and in our case apply it to smartphone use. An ongoing discussion and self-criticism within HCI and Ubicomp literature is that studies tend to avoid re-use of research methodologies [19,20], thus making it hard to reliably compare results with previously published work, and to consistently build theory. In other words, by developing tools that enable us to consistently build on previous literature we facilitate the articulation of testable theories of technology use, and in our case of smartphone revisitation habits.

RELATED WORK

Temporal patterns in smartphone use

Previous work has extensively focused on how individuals use their smartphones and their applications. For instance, Yan *et al.* [36] found that the majority of mobile device usage is brief: 50% of mobile phone engagement (the time period between the user unlocking and relocking the device) lasts less than 30 seconds. More recent work has shown similar results [13], further focusing on micro-usage habits that entail using apps in short bursts of less than 15 seconds, suggesting a strong temporal nature in smartphone use. Similarly, Falaki *et al.*'s [11] study on diversity in smartphone battery usage reported finding short-term application usage (10-250 seconds) among their participants, and Böhmer *et al.*'s [3] large-scale study on mobile application usage revealed that mobile phone owners use their device for an average of 59 minutes daily, with the average application session lasting 72 seconds. Oulasvirta *et al.* [26] demonstrated how attentional limitations emerge in the mobile contexts and result in short bursts of smartphone interaction.

Considering routine, and focusing on overall mobile phone users' habits, Oulasvirta *et al.* [25] suggest that mobile phones are "*habit-forming*" devices, highlighting the "checking habit: brief, repetitive inspection of dynamic content quickly accessible on the device." This habit was found to comprise a large part of mobile phone usage, and follow-up work [13] has argued that the checking habit is one of the behavioral characteristics that leads to mobile application micro-usage, which is subsequently manifested as short bursts of interaction with applications. However, it remains unclear whether other temporal patterns of behaviour exist, and whether they can be attributed to the technology, to humans, or a combination of both.

For instance, rapid switching between applications is not exclusively mobile: multitasking and the management of multiple tasks occur on desktop computers as well [8,15]. These studies found that desktop users were focused on goals but within these goals they embedded a set of secondary goals that were completed through multitasking, albeit not without impacting the users' attention and workflow [21]. This observation led to the technical advances within multitasking and the options that almost all multipurpose technical devices support: being able to seamlessly switch from one application (and often from one task) to another.

Further elaborating on cross-technology habits, previous research [2] has highlighted how smartphone users would 'control' their own use of communication by checking email and other communications on their smartphone, and then waiting to reply until they could access a larger device such as their laptop. This resonates well with findings reported in literature regarding brief bursts of micro-usage interaction: mobile devices seem to be often used for brief checking.

Modelling and prediction

Modelling mobile phone application use enables prediction of which applications will be launched next, or which likely actions will be taken next by the user. This could help in e.g. automatically promoting certain application shortcuts to the phone's home screen [31], or optimising the phone's internal memory management [36]. Several approaches have been trialed to create such application use models. For instance, Huang *et al*. [16] considers contextual information about last used application, time, and location and user profile to predict the application that will be used next. The results showed that a regression model works best by incorporating identified sequences of application use in predicting the next application. This suggests a strong sequential nature in application usage on smartphones.

Leroux *et al*. developed a mobile framework that creates *context profiles* by monitoring application use, day of week, and the user's speed and location [18]. Although using artificially created data, the resulting profiles match real life situations, such as "at work" or "commuting," and can be used to infer a set of applications the user is likely to use. In a similar study, a much wider set of context attributes, such as location, cellular activity (SMS, calls), personal schedule, battery status, screen and status, was used in addition to application use to provide a naturalistic and reliable model for making the predictions to create a dynamic home screen to reduce the time to find an application [31].

In summary, we found that an emerging theme in predicting applications is analysing the sequential and temporal nature of application use, and especially the suitability of using previous application launch history as the predictor of next applications to be launched. The analysis may focus on a micro-scale, i.e. which applications were used right before the predicted applications [16], consider longer periods of time during which an application is "trending" for the user [36], or establish clusters of applications that get used together often [36]. Our work aims to complement the existing approaches by analyzing temporal and sequential application usage in depth. Next we summarise an established methodology that can enable us to study the temporal nature of application usage on smartphones at varying timescales. It also provides flexibility in terms of making either users or applications the focus of the analysis.

Revisitation analysis

Previous research has shown that web page revisitation is one of the most frequent actions in computer use [1]. As a consequence, many researchers have sought to gain a deeper understanding of web revisitation behaviour and improve mechanisms for web navigation such as back buttons [24] and history mechanisms [32]. A significant number of studies in this area have collected logs of web activity to analyse browsing patterns.

A study on web revisitation by Tauscher & Greenberg [32] found a mean revisitation rate of 58%, i.e. on average, 58% of the web pages visited by a user had already been visited by that user at some point in the past. A subsequent study by Cockburn & McKenzie [7] reported web pages revisitation rates of 81%. More recent studies have shown that tabbed browsers – which provide the ability to visit multiple pages in parallel – have further altered browsing behaviour [38].

Obendorf *et al*. [24] developed a theoretical basis for revisitation analysis by distinguishing between different types of revisitation based on their inter-visit temporal duration: *Short-term*: these are instances when a user revisits a particular web page within a short period of time. This behaviour is akin to backtracking or undo. *Medium-term*: this behaviour reflects an intention to re-utilize or observe a particular page. *Long-term*: this behaviour is described as rediscovering or reusing a particular page. Therefore, Obendorf *et al*. argue that revisitation patterns are closely tied to purpose and intention. More relevant to our own work, Adar *et al.* [1] introduced the concept of the 'revisitation curve.' These are normalized histograms of inter-visit times, consisting of 15 exponentially-spaced bins, that can be used to characterise the revisitation pattern for a particular web page. Thus, a revisitation curve is effectively a smoothed histogram that provides a finer resolution than Obendorf *et al*.'s triadic categorisation.

Using revisitation curves it has been demonstrated that different categories of web sites invite different types of revisitation behaviour across all users [1]. For instance, search engines are revisited at extremely short intervals, webmail pages at medium intervals, and entertainment or hobby web pages at hybrid (combinations of short and long) intervals. Thus, revisitation curves are an ideal empirical lens for studying habits, as they are mathematically defined, empirically derived, and reliably comparable due to their quantitative nature.

Using revisitation curves has also made it possible to analyse the behaviour of individual users. Pushnyakov & Gusev [29] demonstrated that, much like web pages, web users have distinct revisitation curves which describe their web browsing behaviour. When considering web browsing on smartphones, Tossell *et al.* [33] showed that smartphone browsing sessions are three times shorter than desktop browsing sessions in terms of duration and pages visited, while web revisitation rates on smartphones are lower than desktops. They conclude that browsers on smartphones have largely given way to 'Native Internet Applications', suggesting that users increasingly access the Internet via native applications while using browsers for ad-hoc searching and medium-term revisitations. Finally, as in our own work, they propose that users' revisitation curves can be used to capture distinct user profiles and peculiarities.

The strong similarities between web browsing on desktops and native application use on smartphones [33] and the established methodology for analysis of web browsing behaviour [1], has lead us to consider revisitation analysis as a valuable approach for studying the use of smartphone applications. We hypothesize that different applications will have different revisitation patterns based on their purpose and content, and that users can be characterised based on their revisitation curves.

METHOD

We conducted a study using the Google Play application store as a platform for experimental data collection. We built a free application, Securacy, that allowed participants to monitor overall network activity on their mobile phones [14]. Securacy provides insight into the network connections that are established behind the scenes when apps are launched, including the use or otherwise of secure network connections and the geographical destination of the app data. Participants could choose to share the collected data with us, including the logs of which applications were launched and at what time. Participants were fully informed about the authors of the application and the scientific purpose of the experiment, and gave explicit consent when activating the software.

After deploying our software on the app store, we publicly promoted and advertised it on Facebook and Twitter in an attempt to reach a diverse population sample. We offered no additional compensation beyond the free use of the app, which offered users a useful function in itself. 165 participants downloaded the Securacy app and provided data using their Android smartphones in naturalistic settings, over a period of 3 months. Tablet devices were excluded from the analysis presented in this paper. 41% of participants were from European countries, 53% North America, and 6% others (e.g., Australia, South-America).

For our purposes, the application used the AWARE[1] framework to log a new record every time a user launched an application. The entry included a timestamp (in the time zone of the participant), a unique ID for the participant, a unique package name assigned to the application by the developer, and the localized application name.

[1] www.awareframework.com

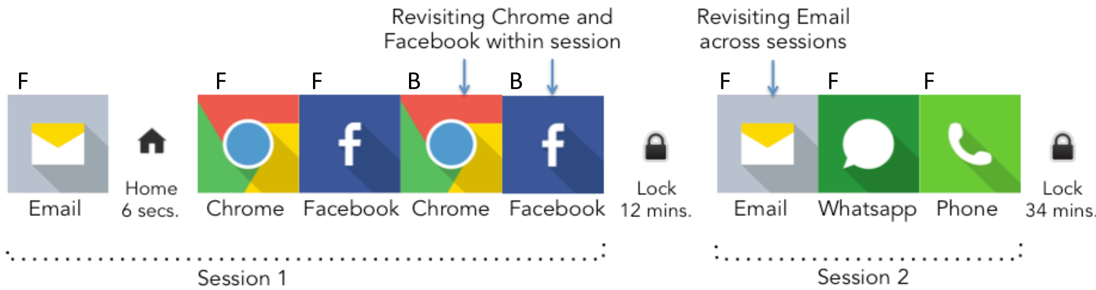


Figure 1. Example chain of application use across different sessions. An app may be visited for the first time in a session, tagged as a forward action (F), or revisited within a session (e.g. Chrome and Facebook), tagged as a backtracking action (B).

Our Securacy application interrupted use of other apps only under two conditions: when the user installed an application for the first time, and when the user removed an application. This meant that the application's functionality did not affect normal use of the phone, and that the app for data collection appeared infrequently in our app traces.

During the study we captured 199,859 application launch events for 1,527 unique applications across all users. In our analysis we ignore apps that are core to Android OS (e.g. the home screen and system UI), or a means to launch other applications. We found that many of the participants had installed skinned UIs and application launchers to replace parts of the standard Android OS. For these users we also ignore such system-wide applications.

Data Filtering

Since we are focusing on revisitation patterns, and to avoid outliers' bias in our data, we removed the users with fewer than 10 days of logged activity from our analysis. We also only analyse the revisitation patterns for apps that are installed by more than a single participant within our dataset (approximately 1400 apps within our dataset). Figure 2a shows the total number of app launches for each of our participants. Similarly, we constructed Figure 2b to determine the popularity of applications. Within the study there were few apps that were opened many times (approximately 100 apps being used more than 100 times each), and many apps that were opened only a few times (approximately 1000 apps opened less than 10 times each). The top 10 most frequently used apps (excluding filtered apps) were: Whatsapp, Chrome, Phone, Facebook, Google Hangouts, SMS/MMS, Gmail, Contacts, Play Store and Desk Clock.

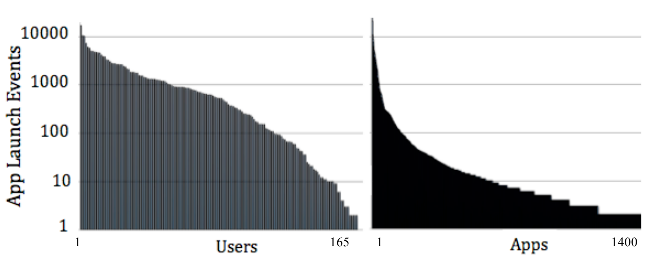


Figure 2a. Number of logged app launch events per user (left); Figure 2b. Number of logged launches per app (right)

While users greatly varied in how much they use their phone (Figure 2a), and apps greatly varied in how popular they are (Figure 2b), both distributions seem to follow a pattern that suggests that the wide range in frequencies is not due to outliers but to a skewed distribution.

Deriving Revisitation Curves

The key component of our analysis is the revisitation curve [1] representing the number of times that an app is revisited within a predefined time interval. We construct a revisitation curve for a certain application by considering the duration between revisits to that application by users (Figure 1). Similarly, a revisitation curve for a particular user is constructed by considering the duration in-between launching *any* app on their phone.

Following the methodology in [1], we use an exponential scale for revisit interval bins. The following 15 bins were used: 1, 2, 4, 8, 16, and 32 minutes; 1, 2, 4, 8, 16, 32 (i.e. 1.3days), 64 (i.e. 2.6days), 128 (i.e. 5.3days), 256 hours and above (i.e. >10.6days). A revisitation curve characterizes an app by its 15-dimensional vector, where each dimension corresponds to the frequency of revisits within the corresponding bin. These curves are like a "signature" of users' behaviour in launching a given smartphone application.

We found that our dataset contained instances of users being interrupted by 'pop-up' applications or processes within apps (e.g. invoking keyboards, cameras, barcode scanners etc.). To mitigate the effects of these pop-ups in our analysis, any return to an app from a recognized pop-up app within 30 seconds is not treated as a revisit when constructing the revisitation curve. Pop-ups that appear for longer than 30 seconds *are* treated as significant switches away from an application, given that 50% of smartphone engagement itself lasts less than 30 seconds [36].

Figure 3 illustrates the revisitation curve for the popular messaging application WhatsApp. The curve indicates the frequency or probability (y-axis) of observing a revisitation to WhatsApp within a particular time period (x-axis). In this case the data is aggregated across all users, since all devices share the same package name per application. For instance, we observe that there is a 16% chance that once an average user launches WhatsApp, they will launch it again within 2-4 minutes (Figure 3). Similar curves were calculated for all applications in our dataset. Additionally, we are able to calculate a revisitation curve for each user, as suggested in [29]. In this case the curve shows the probability the user, having just used an application, will use any other application within a particular time period.

For this analysis we use the same 15 temporal bins. Figure 3 also shows the revisitation curve for a single user from our study. In this example there is a 20% chance that the user will launch an application within 4-8mins following a previous application launch, and approximately 8% chance that they will wait 8-16hrs.

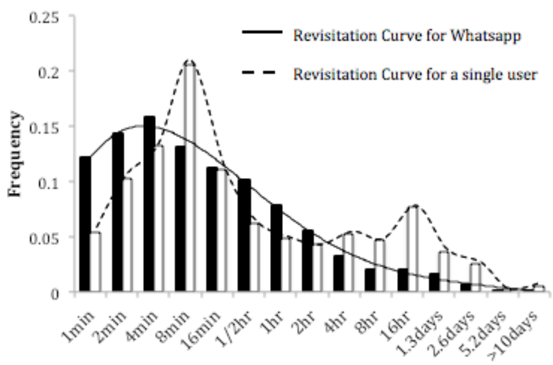


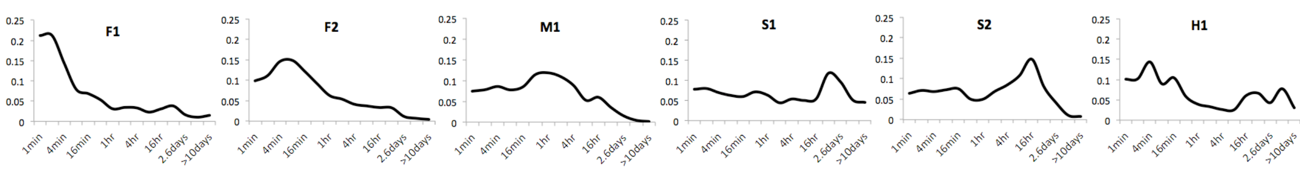
Figure 3. Revisitation histogram and curve for a single app (WhatsApp) and for a single user across all apps.

Clustering

The next step in our analysis consists of identifying clusters of similar applications or similar users. Each distinct application or user is described by a respective vector of 15 values corresponding to the 15 bins of their revisitation curve. By treating these values as features we are able to apply clustering algorithms to identify clusters with similar features. We used the k-means clustering algorithm to identify clusters of applications and clusters of users. Thus, we are able to identify applications and users that share similar revisitation curves. We iteratively apply k-means for a varying number of clusters and use within-groups sum of squares to build an elbow graph in order to identify the optimal number of clusters.

Revisitation analysis in-session

The final part of our analysis considers revisitation behaviour regarding applications that the user has recently launched since unlocking their phone. This analysis focuses on single "sessions" of use: for each session we only consider the set of applications that were used between unlocking and locking the phone. For every such session (performed by an individual user) we consider whether an application is launched for the first time (marked as "F" for forward), or whether it was launched previously within the session (marked as "B" for backtracking). For instance, consider a user who unlocks their phone and uses the following applications in sequence: Email, Chrome, Facebook, Chrome, Facebook (Figure 1), before re-locking their phone. This sequence of 5 application launches can be represented as: FFF-BB. Our encoding shows that the user used three distinct applications in succession (FFF), then used two applications that had already been used in the session (BB). For each string we can calculate its length and the ratio of B or F characters. Finally, we used regular expressions to group these strings into pre-determined categories. These categories emerged from qualitative analysis using two independent coders.



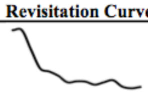
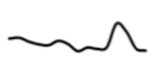
Cluster Label	Description	Centroid Revisitation Curve	Example Apps	Cluster Size (# Apps)
F1	Fast		Google Play Store, Facebook Messenger, InoReader, Chrome Beta, BlackBerry Messenger, Reddit, Okcupid	32 (13%)
F2	Fast		Chrome, Whatsapp, Facebook, Google Hangouts, SMS/MMS, Viber, Youtube, Contacts+, Google Maps, Firefox, Spotify, Skype, Snapchat, Xperia Conversations, Line, Reddit News, Telegram Messenger, Music, Falcon Pro	82 (33%)
M1	Medium		Phone, Gmail, Contacts, Email, Dialer, Clash of Clans, Instagram, Outlook, Yahoo Mail, Opera Browser	47 (19%)
S1	Slow		Gallery 3D, Calendar, Camera, Twitter, Calculator, Clean Master (Speed Booster) Runkeeper Pro, Flipboard, Google Play Services, Mobile Bank, Mobile Weather, Flickr, Google Doc Editor, Tumblr, Quick Office, Google Translate	30 (12%)
S2	Slow		Settings, Desk Clock, Organiser, Tinder, Plants vs. Zombies 2, Clash of Lords, Titanium Backup, Hot or Not, Control Panel, Candy Crush Saga, Castle Clash	40 (16%)
H1	Hybrid		Evernote, Google+, Google Docs, MusicBox, Adobe Reader, 9gag, Video Player, Meo Remote, Waze, Dictionary, Opera Mini	21 (8%)

Table 1. The 6 clusters of applications based on their revisitation curve. For each cluster we provide a label and description, a visualisation of its centroid revisitation curve, some representative apps from the cluster, and the popularity of the cluster

RESULTS

Revisitation Curves: Applications

Table 1 shows the 6 application clusters and the revisitation curve for each cluster centroid (i.e. the prototypical revisitation curve for that cluster). Since revisitation curves consist of 15 bins, we conceptually splits those bins into three 5-bin segments: fast, medium, and slow, as proposed in [24]. Each cluster is then characterised based on where its frequency modality lies (fast, medium, slow), or whether it has a bimodal distribution (hybrid).

We found that apps in cluster F1 had a peak revisitation at 1-2 minutes. The types of apps in F1 include *instant messaging* (Facebook Messenger, Blackberry Messenger), *app store*, *social media* and *dating*, and *web browser*. The F2 cluster had peak revisitation at 8-16 minutes and also included *instant messaging* (Whatsapp, SMS, Viber, Skype, Snapchat, Telegram Messenger, Google Hangouts, Line), *browsers* (Chrome, Firefox), and *social media* (Facebook, Reddit, Falcon Pro (Twitter client)).

On the other hand, applications in cluster M1 peaked their revisitations at 1-16 hours. The types of apps in this cluster include *email* (Android mail client, Gmail, Yahoo Mail, Outlook) and *phone communication* (Phone, Contacts, Dialer).

The applications with slow revisits were grouped into two clusters. In cluster S1 applications were revisited mostly at 2-3 days, and included *utilities* (e.g. Calculator, Camera, Speed Booster, Calendar, Mobile Banking, Weather, Translator, Document Editors), *multimedia* (Photo Gallery, Flickr, Magazines (Flipboard app)) and *health and fitness trackers*. Applications in cluster S2 peaked at 24 hour revisits, and included *games* (Castle Clash, Clash of Lords, Plants vs. Zombies, Candy Crush Saga), *dating* (Hot or Not, Tinder) and *phone settings* (Control Panel, Settings, Alarm Clock). Finally, in cluster H1 we found apps with peaks at 4 minutes, 16 minutes, 1 day, and 10 days. These are apps which that are typically used on a weekly or fortnightly basis, but the sessions in which they are used contain multiple shorter revisits. This cluster includes apps for *documents* and *notes* (Evernote, Google Docs, Adobe Reader), *GPS and Satnav* (Waze).

Revisitation Curves: Users

In our analysis we excluded 36 users because they did not have enough revisits (fewer than 10 app revisits in total) or did not participate in the study for longer than the maximum bin (10 days). For the remainder of users we

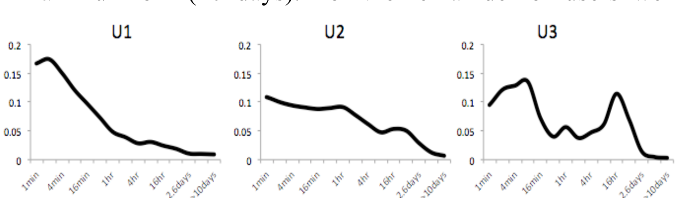


Figure 4. Centroid revisitation curves for the three clusters of users - Checkers (U1), Waiters (U2), and Responsives (U3)

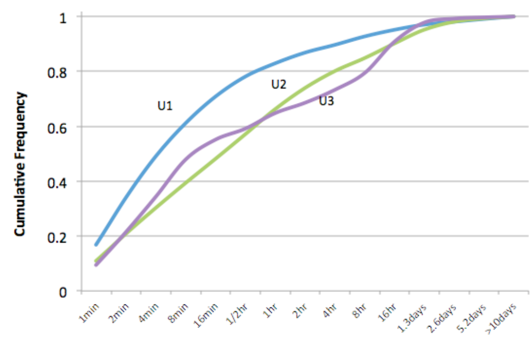


Figure 5. Cumulative probability distribution for each user type: checkers (U1), waiters (U2), responsives (U3).

identified a substantial trichotomy when clustering their user profiles (Figure 4).

Checkers (U1). These users exhibit brief revisit patterns heavily skewed towards fast revisitation (less than an hour). Checkers account for 44% of users.

Waiters (U2). These users exhibit longer revisit patterns uniformly distributed between short-medium revisitations (between 1min and 4hrs) and long revisitations (2hrs - 3 days). Waiters account for 46% of users.

Responsives (U3). These users exhibit sometimes brief and sometimes long revisit patterns. Responsives account for 10% of users.

Figure 5 shows the cumulative probability distribution for all user types. We observe that *Checkers* (U1) have a 60% probability of launching an application 4-8 minutes after launching a previous one – across the whole duration of the study. This ratio becomes 80% at the ½ hour mark. We also observe that the 1 hour and 8 hour marks are pivot points for *Waiters* (U2) and *Responsives* (U3): between those marks *Waiters* are more likely to launch a new application than *Responsives*.

Joint analysis of application and user clusters

For each user cluster we determined the popularity of different application clusters (Table 2). In other words, we measure the association between users' overall phone revisitation patterns and the types of applications that they use. A Chi-Square test showed a significant association between User cluster and App cluster ($\chi^2(10) = 3480$, $p < .01$). While this statistical test checks for a two-way association (instead of one-way causality) between the two factors, we assume that although apps can be targeted at particular types of users and for use in particular contexts, users ultimately exercise control over which applications

App Clusters	User Cluster 1 U1	User Cluster 2 U2	User Cluster 3 U3
F1	768	1591	61
F2	4402	1686	100
M1	7346	9467	400
S1	635	1340	32
S2	2328	3305	390
H1	22747	15648	1313

Table 2. Popularity of application clusters for each user cluster

Regular expression	Example	Strategy	Instances	Popularity	Mean length
^[F]+$	FFFFFF	Initiate	3762	57%	2.27
^[F]+[B]+$	FFFBBB	Initiate-Revisit	1762	26%	4.17
^[F]+[B]+[F]+$	FFBBBFF	Initiate-Revisit-Conclude	598	9%	6.36
n/a	FFBFFB	Other	510	8%	8.50

Table 3. Summary of the 3 strategies we have identified for in-session backtracking. Mean length refers to the average number of application launches within a session, and is further elaborated in Table 4

they use. Therefore, we interpret our results in a one-way manner, hence revealing user preferences.

Analysis of in-session revisitation

Finally, our analysis looked at in-session revisitation patterns, as reflected in the text strings of F's and B's we constructed and analysed (Table 3). We calculated the revisitation ratio (ratio of B's to the size of the session string) as a function of the length of the string itself. The results indicated that as the size of the session increases then the likelihood of revisiting an application increases too, but eventually levels off at approximately 55% (n.b. revisitation can account for more than 50% of a session where individual apps are re-visited more than once).

Our results reveal a much higher level of switching back and forth between applications than previous studies. For example, Leiva *et al.* [17] reported that app switching was uncommon – switches to other apps interrupted 8% of app interactions.

Next, we manually analysed all session strings looking for patterns and similarities across multiple users. A qualitative analysis with two independent reviewers identified 3 distinct patterns which we then coded as regular expressions. The results suggest that 92% of all sessions fall into one of these 3 "Strategies" shown in Table 3, while only 8% could not be classified using our pattern-matching approach. Details about the mean length of different phases (in terms of number of application launches) are shown in Table 4.

Strategy	Length of Initiate	Length of Revisit	Length of Conclude	Mean length
Initiate	2.27	-	-	2.27
Initiate-Revisit	2.36	1.81	-	4.17
Initiate-Revisit-Conclude	2.50	2.53	1.33	6.36
Mean length	2.32	1.99	1.33	3.63

Table 4. For each phase of each strategy we show the mean length (number of application launches)

DISCUSSION

Micro-level analysis of smartphone revisitation patterns

Our analysis of in-session backtracking behaviour has enabled us to develop descriptive statistics of micro-level app usage behaviour on smartphones. Based on our results in Tables 3 and 4, we are able to define IRC (Initiate-Revisit-Conclude), a rudimentary model of usage. The model suggests that smartphone use – on a micro level – begins with an Initiation phase. This consists of launching sequentially a number of distinct applications. This phase alone accounts for 57% of smartphone use (Table 3), and has a mean length of 2.27 apps (Table 4). Such an initiation strategy has been noted before [3], suggesting that certain apps are likely to encourage a user to engage with interaction with the smartphone.

However, we also find that there is a 35% chance that the Initiation phase will be followed by a Revisiting phase with mean length varying between 1.81 - 2.53 applications (Table 4). During this phase users will switch between applications they have launched moments ago since the session begun. This phase is indicative of users going back to the same applications possibly to check for new content, to confirm or verify some information, or for the sake of closure.

Finally, we observed that 9% of sessions contain a short (1.33 apps) Concluding phase. This consists of visiting a typically smaller number of new apps that were not launched during Initiation (and not visited during Revisitation). Our analysis suggests the existence of a user habit of checking one or two applications before turning off their phone. This mirrors the initiation habit reported previously [3], and is akin to a closing "ritual" for users. It is uncertain whether users choose to launch certain apps before switching off their smartphone, or whether launching certain apps leads them to switch off their smartphone.

Our rudimentary model can be used to make testable predictions. First, we expect that about 92% of all usage sessions can be described by this Initiate-Revisit-Conclude pattern. The remaining 8% may be attributed to users' inherent diversity. We have also identified certain interesting relationships between the different phases of our IRC model. The model predicts that in the presence of the Conclude phase we expect the Revisit phase to be longer than the Initiate phase (Table 4). However, in the absence of the Conclude phase then Revisit is actually shorter than Initiate. Thus, the model predicts that as Revisit grows in length it is increasingly likely that the user exhibits a Concluding phase before turning off the smartphone. Coupled with our result which poses an upper limit of approximately 55% on in-session backtracking – we can predict that as the length of Revisit approaches the length of Initiate, a Concluding phase is likely to take place and wrap-up the session.

Smartphone vs. web revisitation

Our analysis identified 6 clusters of applications, which we characterised as Fast, Medium, Slow, and Hybrid (Table 1). Surprisingly, our results bear interesting similarities and differences with previous work on web revisitation patterns. Specifically, our analysis reveals the same four high level revisitation patterns as those reported in [1], which reported the habits of more than 600,000 users. We highlight these in Table 5, where we show for each type of revisitation pattern which applications we found and which websites were reported by [1].

Fast Revisitation on smartphones and the web

The websites with fast revisitations are fast monitoring and auto refresh pages, whose content changes at a fast rate. These bear great resemblance to social media and instant messaging applications found on smartphones in our study. We note that the Adar *et al.* study [1] does not mention social media or social networking sites, since it was conducted in 2008, and such websites have significantly grown in popularity in subsequent years.

Nevertheless, it is interesting to also note that Hub & Spoke websites have similarly fast revisitation patterns as social media and browsers on smartphones. Adar *et al.* reported that many of the pages in this fast cluster appeared to exhibit a hub-and-spoke revisitation pattern. i.e. a page containing many links to other pages, which the user is likely to visit briefly before returning to the original page to explore. An example would be a shopping site with a list of many items. Our analysis suggests that social media applications (and their respective instant messengers) and browser applications are being used in much the same way. Social media hubs also tend to “link” to external applications, taking the user to the browser, Youtube, music player and so on. Similarly, browser applications can also trigger other applications on smartphone.

Finally, it is important to highlight a peculiar difference between our findings and those in [1]. The fast revisitations in the Adar *et al.*’s web analysis have almost no long-term revisitation. On the other hand, our clusters of fast revisitation have non-trivial long term revisitation. We attribute this difference to the fact that while there is a limited number of applications on one’s phone, there is practically an unlimited number of websites that one can visit. Thus, while users tend to revisit certain websites frequently until they completely forget about them; on smartphones users tend to use certain applications frequently, then forget them, and then come back to them eventually. This can be attributed to the visual presence of applications in a phone’s dashboard that act as a reminder and prominent anchor.

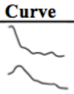
Cluster Group	Centroid Curve	Description	Corresponding cluster group descriptions from Adar *et al.* [1]
Fast (F1,F2)		Instant Messaging, Browser, Social Media	Hub & Spoke, Shopping & Reference, Auto refresh, Fast monitoring, Pornography & Spam.
Medium (M1)		Email and Phone Communication	Popular homepages, Communication, .edu domain, browser homepages.
Slow (S1,S2)		Utilities, Multimedia, Health and Fitness, Games, Dating, Phone Settings	Entry pages, Weekend activity, Search engines used for Revisitation, Child-oriented content, Software updates
Hybrid (H1)		Documents, Notes, Video, Satnav	Popular but infrequently used, Entertainment & Hobbies, Combined Fast & Slow.

Table 5. For different revisitation patterns we show which smartphone apps we report, and which websites are reported by Adar et al. [1].

Medium Revisitation on smartphones and the web

Across the web and smartphones, we find that communication-related pages and applications tend to exhibit medium revisitation patterns. On the web, we find that webmail and forums are in this cluster, while in our data this refers to email and phone communication. Adar *et al.* have suggested that such medium revisitation patterns (1-4 hrs, or daily) are likely due to the timescale of human-to-human interaction using web mediated communication. In other words, this is driven by humans’ tendency to communicate more on an hourly/daily basis rather than significantly faster or slower.

On smartphones we find a similar revisitation curve for phone and mobile-based email communication, suggesting that these communication technologies capture the same fundamental communication habits. We also note that communication with these technologies (asynchronous email and synchronous voice) differ from near-synchronous ‘instant messaging’ in our data, suggesting that these are appropriated differently by users, and serve a different purpose on smartphones.

An important difference we find between medium revisitation patterns on the web and on smartphones is attributed to home pages. Popular homepages appear in the medium revisitation cluster for web browsing, and they typically act as starting points for navigation in each browsing session. The equivalent in our dataset may be the UI or a popular application launcher, however we decided to filter these applications from our analysis due to their peculiar functionality. An informal assessment of their data has shown us that they would most likely fall in the fast clusters: the UI on smartphones acts both as a home page and a directory for finding applications, and therefore is more akin to a hub-and-spoke functionality.

Slow Revisitation on smartphones and the web

Once again, we find substantial similarities in slow revisitation patterns on the web and smartphones. In this cluster we find web pages with weekend activity and child-oriented content, which exhibit slow revisitation curves. These are conceptually very similar to the dating, health, fitness and game apps on smartphones which also exhibit slow revisitation. Thus, it can be argued that both categories reflect possibly individual or personal activities, which tend to follow a similarly slow periodicity on the web and smartphones. A further similarity we observe is that software update websites have a slow revisitation curve, as does the settings application and utilities on smartphones. These involve tweaking the configuration of one's system, and we find that both exhibit slow periodicity in both studies.

Do smartphones induce usage habits?

To investigate whether smartphones are a "special" technology that induces usage habits, we review the variety of evidence that our own study and previous studies offer. While we cannot explicitly test this assertion, we are able to interpret a variety of evidence across different studies. Our analysis has looked at temporal patterns of application usage at the macro level (revisitation patterns) and at the micro level (backtracking).

In our analysis we set out to investigate whether smartphones induce temporal usage habits that we do not see in other technologies. However, our analysis has shown that many of the macro-level behaviours we observed in our study bear close resemblance to web browsing habits. Hence our study provides some evidence to suggest that temporal patterns of usage behaviour on the web and on smartphones are driven by the nature of the service and information, and less so by the technology. The temporal behaviours that we have identified as common across the web and smartphones can be summarised as follows: (i) Users make quick revisits to applications and websites that contain fast-changing content or hub-and-spoke functionality. This is effectively the "checking" habit and "micro-usage" that has been reported on smartphones [13,25]. (ii) Websites and applications that facilitate asynchronous non-verbal or synchronous verbal human-to-human communication following medium revisitation. This is possibly driven by humans' tendency to communicate on an hourly/daily basis rather than significantly faster or slower [1]. (iii) Websites and applications that relate to personal activities follow a slow revisitation pattern. (iv) Websites and applications that involve technical tweaking and reconfiguration follow a slow revisitation pattern.

The first similarity we noted is the "checking" habit that leads to individuals checking and rechecking their phone. We argue that this habit is not unique to smartphones, as indeed users exhibit this habit on the web, and for similar purposes. Therefore, this habit is likely to be driven by the type of service or information that the use gets – irrespective of technology. In fact, the extent of similarities between user habits on the web and smartphones is such that we were unable to identify a temporal behaviour that is unique to smartphones. Most habits we identified on smartphones can be attributed to the service and not the technology: smartphones appear to be a conduit. Hence, we argue that, when viewed through the lens of revisitation analysis, smartphones do not induce usage habits at a macro-level.

However, we have found very interesting backtracking patterns which suggest that smartphones do induce usage habits at the micro level. Surprisingly, our very simplified model of Initiate-Revisit-Conclude accounts for 92% of all usage we observed. Such a model has not been reported in web browsing studies. It is likely that the behaviour predicted by our model is induced by a limitation of smartphone technology: we have a relatively small number of applications available on our smartphone at any given time. Unlike web browsing where an almost unlimited number of pages is available for visit, on smartphones we have a few applications to choose from – and we tend to use and re-use them within individual sessions.

User characteristics or preferences?

To maintain our assessment that smartphones do not induce macro-level habits, we need to explain the substantial differences we observed between the different user clusters. Our analysis identified 3 clusters of users, each with different habits of application usage. Analysis of the adjusted residuals for Table 2 shows some interesting differences: Checkers (U1) use a disproportionately large amount of F1 and F2 (fast apps), and use disproportionately small amount of other apps. Waiters (U2) use disproportionately fewer F1 and F2, and relatively more slow, medium and hybrid apps. Responsives (U3) seem to use a mixture of slow and fast apps, as opposed to using mainly hybrid apps.

We contend that these are not habits induced by smartphones, but are more likely to reflect user preferences. Our justification for this argument is that we note that every user cluster had available applications from every application cluster. Therefore, each user cluster had the *potential* to use each application cluster, but did not choose to do so. Given our one-way assumption regarding causality between users and apps, we can infer that the skewed results in Table 2 are due to user preferences and not applications' availability. In other words, user preferences are driving the installation of applications on smartphones, and subsequently the nature of these applications is giving rise to the revisitation patterns we have observed in our users.

The mobile nature of smartphones, the fact that we carry them constantly [9], and their always-on connectivity all point to a very intuitive assertion in our discipline: people are induced to check their phones all the time. Yet our analysis has revealed both medium and slow revisitation patterns on smartphones that contradict this intuition. In addition, we show that smartphone revisitation patterns are very similar to desktop-based web browsing, which is a technology with rather different affordances. Hence, our findings offer evidence of certain innate habits that we exhibit across many technologies we use.

There is a way to bridge our discipline's intuitive understanding about smartphones with the empirical findings we have presented. If we consider smartphones not as hardware devices, but as sets of functionalities and services (i.e. apps), then we can make this reconciliation.

We suggest shifting our assertion from “checking our smartphones because they are mobile and they are available” to “checking our smartphones because of their apps.” The mobility that our smartphones offer apparently makes no difference to our revisitation patterns. Whilst today we may be spending more time interacting with our smartphones than with our desktop computers a decade ago, in fact we are doing more of the same: we still exhibit fast, medium, and slow revisitation patterns. We still need to check for fast-changing content often, we still communicate with others at a medium pace, and we still engage in personal activities (and system tweaking) on a longer periodicity.

This brings us to the theoretical gap we pointed out in our motivation for this work, which in many ways remains elusive for our discipline. What are the fundamental forces that determine our use of mobile technology? Theoretical models have attempted to offer some explanation into how humans skew their behaviour when allocating time to applications, channels, searches, or web pages. A key purpose of technology use is to keeping informed, enhance social interaction, as well as simple entertainment [22,30], and people use technology that they believe will offer the gratifications they seek [28]. Therefore, as a ubiquitously available solution, mobile phones naturally compete against other information sources for users' attention. A useful analogy may be the use of the remote control, which has facilitated consumers to be more selective with TV channels so as to gain more gratification [28]. Do mobile phones make us more selective in our information needs? Our results suggest that different user clusters are more attracted to different types of services. However, additional investigation is necessary to address this question.

Further theoretical analysis of humans' skewed distribution of time allocated to activities has also been associated with people's ability to perceive their own past rate of activity [34] which subsequently leads them to accelerate or decelerate their rate of activity [4]. It has been hypothesized that this behaviour may be a fundamental evolutionary mechanism because it can synchronize populations of interacting species [5], stabilize them [23], and diversify gene pools [6]. In this sense, temporal skewing and backtracking may be far from confined to smartphones; rather it could be an instance of a broader human mechanism in how we allocate time to activities and how we regulate the rate of our activities.

LIMITATIONS

We note that our network monitoring application is likely to have attracted a specific type of user, which places some limitations on the generalizability of our findings. Due to our use of the Google Play App store to collect data we were unable to collect detailed information about the users included in our sample (e.g. gender, age, etc.).

We also note that although our study follows a “typical” [25] method to examine user habits, using quantitative data to identify frequent behaviours (e.g. as in [27] [35]), other work has noted a distinction between frequent and habitual behaviour which suggests that the latter is a subset of the former. Oulasvirta *et al*. [25] suggests that habitual behaviours are extremely rapidly executed, whilst non-habitual behaviours are slower due to decision-making, and that habitual behaviours are also consistently associated with a particular triggering context. In our analysis we only consider temporal intervals between app visits, and the act of unlocking the phone as triggers for habitual behaviour. Therefore our analysis is unable to isolate habits that may be the result of other contextual cues (e.g. location), which may be unique to smartphone technology.

Although our revisitation analysis reveals repetitive behaviour, which is important for identifying smartphone habits, our analysis does not address differences in content contained within apps or websites. While we indicate that desktop browsing and smartphone app use exhibit similar characteristics, further research is required to understand whether there are content differences (e.g. more personal content) that distinguish smartphone and desktop use.

A further limitation of the analysis presented within this paper relates to the major advancements in desktop-based web browsing technology and web content that have taken place since the study reported in [1] was conducted. The significant growth in social media, and the proliferation of features such as tabbed web browsing, for example, make it possible that the revisitation patterns associated with desktop-based web browsing may look quite different today.

CONCLUSION

Motivated by prior work on analysis of revisitation patterns on the web, our paper has presented an analysis of the diverse ways that people revisit smartphone applications. Our revisitation analysis highlights two important findings. On a micro-level, we propose a simplified model of backtracking which accounts for 92% of usage on smartphones. On a macro-level, we find that smartphone revisitation bears remarkable resemblance to web browsing on desktops. This indicates that much of our habitual use of smartphones is not driven by the technology's characteristics, but rather by the characteristics of the services and information needs we have. Additionally, we call for researchers studying smartphone use to consider addressing the more fundamental driving forces that shape our use of smartphones, and indeed of technology in general. As a first step in this direction, in this paper we propose using revisitation analysis as a methodology for studying technology use.

ACKNOWLEDGMENTS

This work is partially funded by the Academy of Finland (Grants 137736, 276786, 285062), TEKES (Grant 2932/31/2009), and the European Commission (Grants PCIG11-GA-2012-322138 and 645706-GRAGE).

REFERENCES

1. Eytan Adar, Jaime Teevan, and Susan T. Dumais. 2008. "Large scale analysis of web revisitation patterns." In Proc. SIGCHI conference on Human Factors in Computing Systems, pp. 1197-1206.
2. Genevieve Bell and Paul Dourish. 2007. "Yesterday's tomorrows: notes on ubiquitous computing's dominant vision." Personal and Ubiquitous Computing 11, no. 2: 133-143.
3. Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. 2011. "Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage." In Proc. 13th international conference on Human computer interaction with mobile devices and services, pp. 47-56.
4. Dirk Brockmann, Lars Hufnagel, and Theo Geisel. 2006. "The scaling laws of human travel." Nature 439, no. 7075: 462-465.
5. James M. Bullock, Robert E. Kenward, and Rosie S. Hails. 2002. Dispersal ecology: 42nd symposium of the British ecological society. No. 42. Cambridge University Press.
6. Jean Clobert, Danchin, Etienne, André A. Dhondt, and James D. Nichols. 2001. Dispersal. Oxford: Oxford University Press.
7. Andy Cockburn and Bruce McKenzie. 2001. "What do Web users do? An empirical analysis of Web use." International Journal of human-computer studies 54, no. 6: 903-922.
8. Laura Dabbish, Gloria Mark, and Víctor M. González. 2011. "Why do i keep interrupting myself?: environment, habit and self-interruption." In Proc. SIGCHI Conference on Human Factors in Computing Systems, pp. 3127-3130.
9. Anind K. Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin-Hyuk Hong, and Julian Ramos. 2011. "Getting closer: an empirical investigation of the proximity of user to their smart phones." In Proc. 13th international conference on Ubiquitous computing, pp. 163-172.
10. Trinh Minh Tri Do, Jan Blom, and Daniel Gatica-Perez. 2011. "Smartphone usage in the wild: a large-scale analysis of applications and context." In Proc. 13th international conference on multimodal interfaces, pp. 353-360.
11. Hossein Falaki, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. 2010. "Diversity in smartphone usage." In Proc. 8th international conference on Mobile systems, applications, and services, pp. 179-194.
12. Denzil Ferreira, Eija Ferreira, Jorge Goncalves, Vassilis Kostakos, and Anind K. Dey. 2013. "Revisiting human-battery interaction with an interactive battery interface." In Proc. 2013 ACM international joint conference on Pervasive and ubiquitous computing, pp. 563-572.
13. Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, Louise Barkhuus, and Anind K. Dey. 2014. "Contextual experience sampling of mobile application micro-usage." In Proc. 16th international conference on Human-computer interaction with mobile devices & services, pp. 91-100.
14. Denzil Ferreira, Vassilis Kostakos, Alastair R. Beresford, Janne Lindqvist, and Anind K. Dey. 2015. "Securacy: An Empirical Investigation of Android Applications' Network Usage, Privacy and Security." In Proc. 8th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec).
15. Victor M. González and Gloria Mark. 2004. "Constant, constant, multi-tasking craziness: managing multiple working spheres." In Proc. SIGCHI conference on Human factors in computing systems, pp. 113-120.
16. Ke Huang, Chunhui Zhang, Xiaoxiao Ma, and Guanling Chen. 2012. "Predicting mobile application usage using contextual information." In Proc. 2012 ACM Conference on Ubiquitous Computing, pp. 1059-1065.
17. Luis Leiva, Matthias Böhmer, Sven Gehring, and Antonio Krüger. 2012. "Back to the app: the costs of mobile application interruptions." In Proc. 14th international conference on Human-computer interaction with mobile devices and services, pp. 291-294.
18. Philip Leroux, Klaas Roobroeck, Bart Dhoedt, Piet Demeester, and Filip De Turck. 2013. "Mobile application usage prediction through context-based learning." Journal of Ambient Intelligence and Smart Environments 5, no. 2: 213-235.
19. Yong Liu, Jorge Goncalves, Denzil Ferreira, Simo Hosio, and Vassilis Kostakos. 2014. "Identity crisis of ubicomp?: mapping 15 years of the field's development and paradigm change." In Proc. 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 75-86.
20. Yong Liu, Jorge Goncalves, Denzil Ferreira, Bei Xiao, Simo Hosio, and Vassilis Kostakos. 2014. "CHI 1994-2013: Mapping two decades of intellectual progress through co-word analysis." In Proc. 32nd annual ACM conference on Human factors in computing systems, pp. 3553-3562.
21. Gloria Mark, Daniela Gudith, and Ulrich Klocke. 2008. "The cost of interrupted work: more speed and stress." In Proc. SIGCHI conference on Human Factors in Computing Systems, pp. 107-110.
22. Denis McQuail. 1987. "Mass Communication Theory: An Introduction, Beverly Hills." Calif: Sage.
23. J. D. Murray. 1993. Mathematical biology. New York.

24. Hartmut Obendorf, Harald Weinreich, Eelco Herder, and Matthias Mayer. 2007. "Web page revisitation revisited: implications of a long-term click-stream study of browser usage." In Proc. SIGCHI conference on Human factors in computing systems, pp. 597-606.
25. Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. 2012. "Habits make smartphone use more pervasive." Personal and Ubiquitous Computing 16, no. 1: 105-114.
26. Antti Oulasvirta, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. 2005. "Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI." In Proc. SIGCHI conference on Human factors in computing systems, pp. 919-928.
27. Kurt Partridge and Philippe Golle. 2008. "On using existing time-use study data for ubiquitous computing applications." In Proc. 10th international conference on Ubiquitous computing, pp. 144-153.
28. Elizabeth M. Perse, 1998. "Implications of cognitive and affective involvement for channel changing." Journal of Communication 48, no. 3: 49-68.
29. Philipp Pushnyakov and Gleb Gusev. 2014. "User profiles based on revisitation times." In Proc. companion publication of the 23rd international conference on World wide web companion, pp. 359-360. International World Wide Web Conferences Steering Committee.
30. Quint Randle. 2003. "Gratification Niches of Monthly Print Magazines and the World Wide Web Among a Group of Special Interest Magazine Subscribers." Journal of Computer Mediated Communication 8, no. 4.
31. Choonsung Shin, Jin-Hyuk Hong, and Anind K. Dey. 2012. "Understanding and prediction of mobile application usage for smart phones." In Proc. 2012 ACM Conference on Ubiquitous Computing, pp. 173-182.
32. Linda Tauscher and Saul Greenberg. 1997. "How people revisit web pages: Empirical findings and implications for the design of history systems." International Journal of Human-Computer Studies 47, no. 1: 97-137.
33. Chad Tossell, Philip Kortum, Ahmad Rahmati, Clayton Shepard, and Lin Zhong. 2012. "Characterizing web use on smartphones." In Proc. SIGCHI Conference on Human Factors in Computing Systems, pp. 2769-2778.
34. Alexei Vazquez. 2007. "Impact of memory on human dynamics." Physica A: Statistical Mechanics and its Applications 373: 747-752.
35. Hannu Verkasalo. 2009. "Contextual patterns in mobile service usage." Personal and Ubiquitous Computing 13, no. 5: 331-342.
36. Tingxin Yan, David Chu, Deepak Ganesan, Aman Kansal, and Jie Liu. 2012. "Fast app launching for mobile devices using predictive user context." In Proc. 10th international conference on Mobile systems, applications, and services, pp. 113-126.
37. Chunhui Zhang, Xiang Ding, Guanling Chen, Ke Huang, Xiaoxiao Ma, and Bo Yan. 2013. "Nihao: A predictive smartphone application launcher." In Mobile computing, applications, and services, pp. 294-313. Springer Berlin Heidelberg.
38. Zhang, Haimo, and Shengdong Zhao. 2011. "Measuring web page revisitation in tabbed browsing." In Proc. SIGCHI Conference on Human Factors in Computing Systems, pp. 1831-1834.

Traffic in the Smart City

Exploring City-Wide Sensing for Traffic Control Center Augmentation

Smart city technologies can provide substantial benefits that improve people's daily lives. Here, the authors investigate how ubiquitous traffic sensing technologies and techniques can be incorporated with conventional traffic control and monitoring practices in the city of Oulu, Finland. In collaboration with the city's traffic control center and traffic planners, they exploit inductive magnetic sensing and Wi-Fi scanning across the city center, and develop tools to assist traffic operators in their tasks.

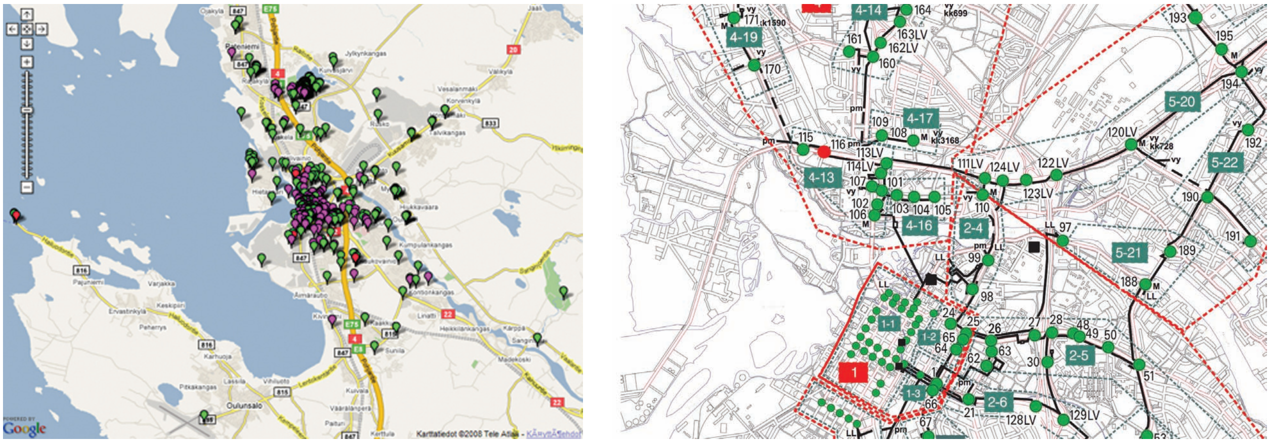
Vassilis Kostakos, Timo Ojala, and Tomi Juntunen
University of Oulu

In the fields of usability and interface design, control centers have long been the subject of study and redesign, often in the context of air traffic control.[1] Through observation and analysis,[2] researchers have typically sought to improve or extend the practices and systems that these operational centers employ,[3] with a particular focus on safety and efficiency. In parallel, new interaction and sensing technologies have arisen that let computing move "beyond the desktop." As these technologies matured, researchers explored ways to exploit them in the control center context. For instance, prior work has explored ways to integrate augmented reality in such a setting.[4]

Modern control centers are increasingly ad hoc,[5] yet still manage to achieve remarkable safety. Unlike the old days of nuclear power plants, in which operators received thorough training on a strict set of tools and instruments, today's control centers are much more dynamic: operators use an array of tools, including websites, in their day-to-day routines.[5] As part of a project at the University of Oulu, we wanted to identify ways in which we could contribute to a control center ecosystem using the technological capabilities of a smart city infrastructure.

Using infrastructure technology to capture city-scale data isn't new. Multiple smart city initiatives have explored city-scale instrumentation of major infrastructure, including energy networks (electricity, gas, and water), building monitoring (in terms of environmental and structural performance), and transport (urban mobility, energy costs, and environmental impact). Some notable initiatives include IBM's Smarter Cities (www.ibm.com/smarterplanet), the EU's Smart Cities and Communities (www.eu-smartcities.eu), the SmartSantander project in Spain (www.smartsantander.eu), and Korea's Songdo

 1089-7801/13/$31.00 © 2013 IEEE



(a) (b)

Figure 1. Technologies for developing traffic sensing and analysis tools. We can see (a) an overview map of the Wi-Fi network we use in this study, and (b) an excerpt map of traffic-light-controlled intersections using inductive loops.

city (www.songdo.com). In the context of urban mobility and transport, particularly, several previous projects have considered ways to estimate road traffic,[6] determine travel patterns using cellular[7] or proximity[8] technologies, or a combination of these, as with the EU's Instant Mobility project (http://instant-mobility.com).

Most urban mobility work has been demonstrated in the context of short-lived research projects and pilots. Monitoring city-scale traffic typically requires cooperation from local traffic operations and planners, and most cities already have policies and regulations in place for monitoring and controlling urban traffic. For example, many cities rely on dedicated traffic monitoring and control centers, either at the local or national scale, to manage traffic and data collection. Furthermore, the analysis of the data collected on urban traffic and mobility could be of interest to multiple stakeholders, including transport operators, environmental agencies, planners, and city councils. Thus, transitioning to smart city traffic monitoring isn't just a matter of technology, but involves multiple organizations and stakeholders.

Here, we describe our collaboration over the past two years with a traffic control center in Oulu, Finland. Rather than redesign the practices of traffic controllers and planners, our aim was to engage them and explore ways in which smart city sensing and analysis techniques could become part of existing practices. This approach has been fruitful in the past — for example, by incorporating GPS traces from taxi fleets and eventually end users to generate traffic estimations.[9]

Technical Setup

Figure 1 shows two main types of technology for developing novel sensing and analysis traffic tools: a free, citywide, open Wi-Fi network, and a network of inductive loop detectors for counting cars. The Wi-Fi network consists of almost 1,300 access points. Coverage is dense across the city center area and sparse in the outskirts located tens of kilometers from the center. These Wi-Fi access points record connecting devices' unique IDs and the time the connection starts and ends, then send this data to a central server for storage in a database. This database contains GPS coordinates for each Wi-Fi access point, letting us identify the patterns of movement for Wi-Fi devices throughout the city. On average, more than 30,000 unique devices use the network each month.

The second sensor type we use are inductive loops. These are typically installed under the road surface, and use magnetic fields to detect the presence of vehicles above them. They're configured to operate in a binary state (vehicle present/not present), and primarily function to dynamically adjust the traffic signal priorities at intersections. As such, counting vehicular traffic is a bonus. In our work, we have access to sensors at 123 traffic light intersections. Depending on the intersection's size and the

(a) (b) (c)

Figure 2. Screenshots from the traffic control center, showing (a) an overview of the work space; (b) the main screen wall, used to collaboratively discuss problems and maintain group awareness; and (c) individual workstations, where operators can focus on any camera feed and simultaneously work at a desktop environment to retrieve and analyze data.

total number of lanes, each intersection has up to 32 individual loop detectors, with each lane in a particular direction typically having two or more loop detectors. In this configuration, the sensors can identify stationary traffic waiting at traffic lights, and therefore can prioritize the traffic light changes based on volume. Due to technical restrictions, only aggregated data from these induction loops is available — that is, the data is an aggregate of vehicle detection events over certain time periods spanning from 2 to 15 minutes depending on the intersection. Overall, we obtained access to both downtown area sensors (which were dense) and main-artery sensors (which were quite sparse).

Our data collection system's architecture relies primarily on Internet technologies. All Wi-Fi device association events are logged by the network's Dynamic Host Configuration Protocol (DHCP) server, and data is uploaded to our data server in real time using http GET requests. Our server pulls traffic-light sensor data from the city's own servers using a persistent VPN connection. The city server that provides our data is in fact a replica server to minimize any potential side effects that our data collection might have on the primary traffic control server.

Traffic Control Center and Existing Practices

Early in the process, we conducted observations, interviews, and a focus group. We engaged with Oulu's traffic control center (see Figure 2), which manages day-to-day traffic, and with the environmental agency responsible for long-term transport planning in the city. The traffic center is a joint task force between multiple city and national agencies, including transport operators and the police. Its main responsibility is monitoring traffic conditions across the city in real time and operating all associated systems. Remote-controlled cameras installed at multiple junctions across the city provide video feeds through which operators monitor traffic. The street and road authorities own all software and hardware systems. Finally, the control center coordinates emergency response relating to traffic accidents and adjusts traffic-light priorities to deal with congestion.

As with other modern control centers,[5] the Oulu traffic control center environment is highly dynamic, and operators rely on various tools and information sources to do their work. We observed that although the city had already installed costly loop detectors, it didn't fully exploit the data they provided: although they can be used for near-real-time and historical traffic pattern analysis, the city initially used them to control traffic lights. The operators mostly analyzed data from the loop detectors using spreadsheets, both on a daily basis and for long-term strategic decisions. Operators relied on the visual camera feeds, as well as their own substantial experience and intuition, to manage day-to-day traffic. City planners similarly relied on spreadsheets to manually generate visual artifacts to help with their analysis.

An important and frequent task for the control center is to manipulate traffic light priorities to ease congestion. During major events such as ice-hockey games, as well as morning and evening rush hours, operators adjust lights to respond to the traffic volume. Their considerable experience with this kind of operation appears to come from two sources. First, because this activity is frequent, operators develop a tacit understanding and intuition about how traffic in our city responds to their adjustments. Second, operators will discuss and exchange ideas and experiences, thus fueling their practice. When new congestion "phenomena" appear — for example, due to a new road or football stadium — they will discuss ways in which it affects traffic and eventually develop strategies for managing it.

After our initial observations and focus groups, we identified two ways in which ubiquitous technology could contribute. The operators' reliance on visual feeds is shaped by those feeds' ephemeral nature, making it hard to discuss new phenomena and develop strategies. Hence, we expected that new tools that "replay" or "recreate" traffic using a representation that can be treated as an artifact would facilitate discussion, experimentation, and exchange of experiences. Similarly, planners' reliance on cumbersome copy/paste procedures using spreadsheet software for generating charts led to a disconnect between asking a question and obtaining an answer using traffic data. Hence, we believed that a direct manipulation tool for exploring historic and live traffic data would be a valuable addition to the existing ecosystem.

Tool Development

Using a co-design process guided by our technology's capabilities, and feedback from domain experts, we iteratively designed three tools. We developed them with both desktop and wall-sized displays in mind to account for the control center's varying screen sizes. First, we designed a tool for generating overview visualizations of traffic conditions across the city, both live and historic. The controllers' current practices are to visually inspect distinct feeds from numerous cameras. The sheer number of feeds makes this task challenging. While the cameras can provide very detailed information about situations at particular locations, a need also existed for a quick-glance overview of traffic conditions across the city center; the ability to explore historic data — particularly recorded during major events — was also necessary.

The heatmap in Figure 3a provides a quick-glance overview of traffic conditions across the city. This visualization can switch between data

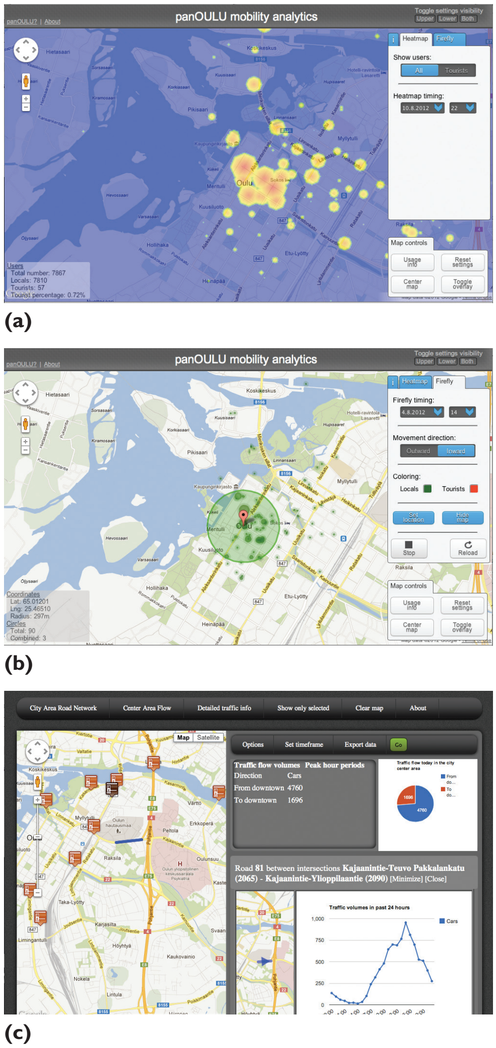


Figure 3. Screenshots of the tools we developed. (a) A heatmap generated from Wi-Fi mobility traces. The red areas indicate high activity, whereas yellow and blue indicate less activity. (b) A tool for defining catchment areas and identifying where individuals detected in a particular location go. Users can also employ color filters (here, red circles are visitors; green are locals). (c) A point-and-click tool enables operators to select roads on a map and view historical data for them.

from Wi-Fi mobility traces and inductive-loop data, as long as the data consists of geo-located volumes of "activity." When the visualized data is the number of clients connected to specific Wi-Fi access points during a particular period, then the heatmap's activity represents pedestrian mobility. Conversely, visualizing the data from the loop detectors represents vehicular traffic activity. The tool lets users define date ranges, so historic as well as real-time views are possible.

The second tool we designed focused on conducting live and historic catchment area analysis. This tool was inspired by a manual technique that urban architects use.[10] It defines locations of interest (such as the train station or the main square) and records all patterns of movement away from or toward that area. Often, planners record this data manually by shadowing pedestrians in situ, and use it to identify how inhabitants use their city — which paths they take, which roads they avoid, and in general how they move through the city. This data can give urban architects and developers insight, particularly at the planning stage, to help them identify patterns in inhabitants' behavior for consideration when redesigning or rebuilding large urban areas such as train stations or parks. We designed our tool to conduct this kind of analysis with pedestrians both in real time and using historic data; the aim is to explore how pedestrians arrive to or depart from major events and central locations. As the operators noted, although cameras can identify individual cars and pedestrians, they can't track an individual across multiple locations and camera feeds. Our tool, therefore, is a first attempt at capturing fine-grained travel information, and addresses an unfulfilled need: fine-grained replay of traffic flows across the city.

Figure 3b shows the catchment area tool, which works as follows. The operator defines a particular area by clicking on the map, and defines a radius for that area by dragging the cursor. Then, the system looks at the Wi-Fi mobility traces and identifies devices detected at that particular location during a selected time frame. The system then identifies for each device the ultimate location where it was seen during the time window. The system visualizes one or many devices as a circle, and then animates all the circles and moves them from the origin location (the one the operator defined) to the destination (the ultimate location where the device was detected). This animation effect gives the operator a visual representation of how individuals at a particular location move, but can also highlight large-scale patterns. For example, the tool can show how individuals detected at the train station move within one hour. In addition, we're developing various filters that operators can use to color individual circles. Our first filter has been used to differentiate between residents and visitors.[11] This requires an historic analysis on the mobility records for individual devices. Using a small set of heuristics, we essentially define visitors to be devices that have only been detected for a few days over a period of one year. On the other hand, devices that are more frequently detected over long time periods are labeled as locals.

Finally, we identified the need to easily explore historic and real-time traffic patterns. Currently, operators can access such data, but exploring it is cumbersome. Existing tools require that operators define queries to retrieve the data, and analysis relies on spreadsheet software. A critical shortcoming of this approach is that operators can't conduct geographic searches using direct manipulation, but must copy/paste their search results into a spreadsheet. This shortcoming also affects planners, who are more focused on long-term analysis and strategic long-term interventions to the traffic network.

The third tool we developed lets operators explore traffic data historically and geographically (Figure 3c). It's driven by a map containing markers for all potential data sources. Operators can click on a particular data source — either a road segment or a junction — and subsequently look at the data associated with it. The tool lets operators define a specific period, then plots dynamically the volume of activity the data source detects during that period. The tool provides a "results log" section at the right side of the screen. Users can temporarily store graphs and build up a series of them to facilitate exploration and comparison. This means that the tool's dynamic nature doesn't result in graphs disappearing when users click on new data sources.

Although cameras can identify individual cars and pedestrians, they can't track an individual across multiple locations and camera feeds.

Traffic Sensing: Lessons Learned

Smart city technologies have great potential to aid urban transport management and analysis. A crucial characteristic of these technologies is that they can sense individual entities, such as devices. Although many public transport authorities track their fleets in real time, this isn't the case for individuals' vehicles. Such sensing is an important advantage, and enables a range of potential applications and analytical tools, such as the catchment area tool. With their ability to differentiate between individual entities, such technologies should also be able to operate in a real- or near-real-time fashion. Operators mostly rely on camera feeds because of their live and direct nature. Smart city sensing can potentially match this realism and tacitness, so this research direction is worth exploring. Meanwhile, our technologies are showing operators that completely new tools are possible, given both existing and potentially new infrastructure.

Support for Dynamic and In-Depth Use

Operators and analysts work in a dynamic and ad hoc environment, with multiple tools competing for attention. As such, we didn't design our tools with any "official" procedure in mind, but rather to match the setting's dynamics. We found that operators still rely on video feeds because, in many ways, this technology is more trusted and much more familiar to them. Moreover, cameras are a favorite tool because they support quick-glance operation. Interestingly, of the three tools we developed, the data exploration tool (Figure 3c) has received the most encouraging response, despite our own expectations. We initially believed that the heatmap tool's support for quick-glance operation would make it a success, but it turns out that its low update frequency makes it less attractive. The data exploration tool, on the other hand, provided a capability that the existing tools didn't support well. Thus, our experience suggests that tools are likely to be winners if they address two distinct types of use: high-throughput quick-glance tools, and in-depth analytics tools. It's too early to judge our tools' success or failure, but we regularly receive suggestions and ideas for improving them, mostly as regards interface and interaction issues.

Privacy and Security

Operators' primary concern is safety — that is, avoiding accidents and casualties. With this perspective in mind, we were the only ones to raise the issue of privacy. Traffic control centers have a long tradition of using video cameras in public settings, an intrusive technology that can raise eyebrows in different contexts. However, operators have developed an appropriate ethos and attitude toward this technology, striking a balance whereby they use the technology but don't abuse it, for example, by stalking individuals. Also, the data sources we used had the "stamp of approval" from the local city council, further alleviating privacy concerns. Hence, our experience suggests that privacy isn't a primary concern; security is, however.

Security protocols limited which software we could install on operators' and analysts' workstations, so we decided to rely on Web applications that only require a Web browser. Furthermore, we had to set up an elaborate security mechanism to obtain data from the inductive traffic loops. This involved creating an identical replica

Smart city technologies have great potential to aid urban transport management and analysis.

of the database system, which we could access in read-only mode through virtual private networking. This setup ensured that our research in no way compromised the control center's operation. At the same time, it was a nontrivial investment for Oulu that could be perceived as a goodwill gesture. Moreover, the city viewed it as an investment to further utilize existing infrastructures. Specifically, the city has spent substantial resources deploying city-wide Wi-Fi and traffic sensor networks, so our project was an opportunity to further use these infrastructures with relatively small additional costs.

Smart Cities Are Still on Paper, So Let Demos Shine

The smart city concept is still maturing, and for now its principles aren't widely understood among the public. Explaining to traffic engineers what smart city computing is, and how it can help traffic management, was challenging. Once we developed our initial demos, however, all stakeholders quickly saw the potential. Hence, an important point we wish to make is that demos are crucial. Discussing the possibilities for a tool is very different from demonstrating it. In our case, we experienced renewed engagement when the first versions of the demos came out, and we showed them to stakeholders. We feel that demos provide a concrete artifact to discuss and drive collaboration. So, we suggest that researchers approach a potential collaboration with traffic officials with a demo in hand. Sometimes, providing a demo without real data might be impossible, so using data stubs could be appropriate.

The New "Smart Home"?

Intelligent computing has greatly benefited from the focus on specific use cases and scenarios to drive research. For example, the focus on the "smart home" let researchers identify multiple challenges and develop numerous prototypes that addressed them. This includes indoor positioning, session transfer between devices and screens, context inference, and intelligibility concerns. One reason the smart home setting is appealing is that it entails a set of well-defined requirements and can easily be replicated.

Traffic sensing has the potential to be the next major playing field for intelligent computing researchers, for two reasons: relative uniformity and real-world impact. First, traffic engineering and traffic management is a mature discipline; vehicle manufacturing is highly standardized, and, for the most part, road networks are rather uniform across countries. These conditions suggest that technological solutions that work in one setting are likely to be applicable to other situations. Of course, many regions are unique and pose interesting case studies.

Second, the intelligent computing field ethos takes a "your noise is my data"[12] approach, and the traffic sensing domain seems to be a fertile ground in which this approach can flourish. From our own experience, we were surprised by the rich but underutilized data available to the city. This was the case in our project with the inductive loop detectors, which aren't fully exploited. This might occur for several reasons, including lack of resources or expertise. Thus, we expect that other cities might be in a similar situation, offering researchers opportunities to engage with officials. Smart city computing in a traffic sensing context can genuinely affect people's daily lives by moving beyond toy demonstrations: traffic management systems and organizations are vehicles for ubiquitous computing technology to have real-world impact.

We feel that it's crucial for traffic operators and other stakeholders to experience firsthand the benefits that smart city technologies can bring — for example, by using tools and demos similar to those we present here. The road to deploying citywide ubiquitous sensing infrastructure can be challenging: it involves a substantial investment, so some stakeholders might be resistant. Having the support of traffic controllers and analysts is a big first step in enabling wide technological deployment.

References

1. T. Prevot, "Exploring the Many Perspectives of Distributed Air Traffic Management: The Multi Aircraft Control System MACS," *Proc. Int'l Conf. Human-Computer Interaction in Aerospace* (HCI-Aero), AAAI, 2002, pp. 149–154.
2. W.E. MacKay, "Is Paper Safer? The Role of Paper Flight Strips in Air Traffic Control," *ACM Trans. Computer-Human Interaction*, vol. 6, 1999, pp. 311–340.
3. R. Bentley et al., "Ethnographically-Informed Systems Design for Air Traffic Control," *Proc. 1992 ACM Conf. Computer-Supported Cooperative Work*, ACM, 1992, pp. 123–129.
4. W.E. Mackay et al., "Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control," *Proc. SIGCHI Conf. Human Factors in Computing Systems*, ACM, 1998, pp. 558–565.
5. M. Wahlström et al., "Resolving Safety-Critical Incidents in a Rally Control Center," *Human-Computer Interaction*, vol. 26, 2011, pp. 9–37.
6. R. Bolla and F. Davoli, "Road Traffic Estimation from Location Tracking Data in the Mobile Cellular Network," *Proc. Wireless Communications and Networking Conf.*, IEEE, vol. 3, 2000, pp. 1107–1112.
7. F. Calabrese et al., "Estimating Origin-Destination Flows Using Mobile Phone Location Data," *IEEE Pervasive Computing*, Oct–Dec. 2011, pp. 36–44.
8. V. Kostakos et al., "Brief Encounters: Sensing, Modeling, and Visualizing Urban Mobility and Copresence Networks," *ACM Trans. Computer-Human Interaction*, vol. 17, 2010, p. 2.
9. S. Lorkowski et al., "Towards Area-Wide Traffic Monitoring Applications Derived from Probe Vehicle Data," *Proc. 8th Int'l Conf. Applications of Advanced Technologies in Transportation Eng.*, ACSE, 2004, pp. 389–394.
10. M. Schlossberg and N. Brown, "Comparing Transit-Oriented Development Sites by Walkability Indicators,"

Transportation Research Record, vol. 1887, 2004, pp. 34–42.

11. F. Girardin et al., “Digital Footprinting: Uncovering Tourists with User-Generated Content,” *IEEE Pervasive Computing*, Oct.–Dec. 2008, pp. 36–43.
12. G.D. Abowd, “What Next, Ubicomp? Celebrating an Intellectual Disappearing Act,” *Proc. 2012 ACM Conf. Ubiquitous Computing*, ACM, 2012, pp. 31–40.

Vassilis Kostakos is a professor of computer engineering in the Department of Computer Science and Engineering at the University of Oulu, Finland. His research interests include ubiquitous computing, human-computer interaction, and social systems. Kostakos has a PhD in computer science from the University of Bath, UK. He’s a member of ACM. Contact him at vassilis@ee.oulu.fi.

Timo Ojala is a professor of computer science and engineering in the Department of Computer Science and Engineering at the University of Oulu, Finland. His research interests include ubiquitous computing and urban informatics. Ojala has a DrTech in electrical engineering from the University of Oulu. He’s a member of IEEE, ACM, and the Pattern Recognition Society of Finland. Contact him at timo.ojala@ee.oulu.fi.

Tomi Juntunen is a research assistant in the Department of Computer Science and Engineering at the University of Oulu, Finland. His research interests include security and human-computer interaction. Juntunen has a BSc in computer science and engineering from the University of Oulu. Contact him at tomi.juntunen@ee.oulu.fi.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.

Technological Forecasting & Social Change 84 (2014) 29–42

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

From cyberpunk to calm urban computing: Exploring the role of technology in the future cityscape

Hannu Kukka [a,*], Anna Luusua [b], Johanna Ylipulli [a], Tiina Suopajärvi [a], Vassilis Kostakos [a], Timo Ojala [a]

[a] Department of Computer Science and Engineering, University of Oulu, Oulu, Finland
[b] Department of Architecture, University of Oulu, Oulu, Finland

ARTICLE INFO

Article history:
Received 2 November 2012
Received in revised form 10 June 2013
Accepted 27 July 2013
Available online 3 September 2013

Keywords:
Prototyping
Rapid ethnography
Storytelling
In-the-wild studies
Ubiquitous computing
Urban computing

ABSTRACT

In this paper we will present studies aimed at uncovering the attitudes, needs, and expectations people have towards novel types of urban computing technologies deployed in a city. First, we conducted a storytelling competition to elicit future visions of how people imagined the role of technology. Second, we conducted a rapid ethnographic study using a mock-up prototype device in various public locations to gain a deeper understanding of how people would appropriate a specific technology, namely interactive public displays. Lastly, we collected ethnographic material through a diary study and interviews where people recorded their use of existing technology, and through these experiences, imagine how future technologies might affect their lives. We found that these methods proved useful in engaging a city's community to imagine the city's future. Consequently, we were able to explore the current use of technologies in the city and project their possible future use. Contrary to previous speculation in academic and cyberpunk literature, we conclude that digital technologies will not necessarily induce an abandonment of physical urban spaces. Rather, we project an increased sophistication in the sociable uses of urban spaces and technologies, where people blend their online and offline worlds into a single lived reality.

1. Introduction

In much of science fiction literature, and especially in the cyberpunk genre, epitomized in the works of the authors featured in Sterling's “Mirrorshades: The Cyberpunk Anthology” [58], future cities are often depicted as dangerous, alienating, media-saturated mega-cities dominated by large corporations where our lives (and minds) are more integrated with technology than ever before. In addition to cybernetic limbs, characters in this genre are often augmented with cranial implants that enable direct brain–computer interaction and offer access to virtual worlds, which are often depicted as semi-physical places where the protagonist, through an avatar, ventures to find information or to gain access to places that are inaccessible in the physical world. These virtual worlds are dangerous places, where computer firewalls and viruses are also represented by beings that can harm the character, and often the result of injury in these worlds also means death in the physical world — a prime example of such a world can be found in the Matrix films [42], where sentient machines have enslaved all of humanity and keep them under control by feeding a virtual representation of the world as it was before the revolution directly to their brain. Of course, a small group of resistance fighters have managed to escape from the illusion and can traverse between the real world and the virtual one. This dichotomy between the physical and the virtual worlds is very common in much of the cyberpunk literature — people are either in the physical world, or in the virtual, but these two are seen as

* Corresponding author. Tel.: +358 40 7027653.
E-mail addresses: hannu.kukka@ee.oulu.fi (H. Kukka), anna.luusua@oulu.fi (A. Luusua), johanna.ylipulli@oulu.fi (J. Ylipulli), tiina.suopajarvi@ee.oulu.fi (T. Suopajärvi), vassilis@ee.oulu.fi (V. Kostakos), timo.ojala@ee.oulu.fi (T. Ojala).

0040-1625/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.techfore.2013.07.015

separate, detached from each other. These science fiction narratives are also widely used by researchers as materials to study human-technology relations through its various themes, e.g. globalization, embodiment, feminism, and urbanism [64].

Whereas the image of a city in cyberpunk novels and movies is without exception very bleak, a growing body of literature in the domain of *ubiquitous* and especially *urban computing* paints a completely different picture. Starting with the writings of Mark Weiser [63], the father of ubiquitous computing, these visions of future always depict technology as benevolent, friendly, useable, and always there when needed, never failing, never disturbing. Ubiquitous computing (from the Latin word *ubique* meaning "everywhere") is the name given to the third major era of computing, following the mainframe era and the personal computer era.

In a world of ubiquitous computing, technology is built into the physical environment. Computers are everywhere, invisible, watching our every move in order to know exactly when to step out and provide whichever service is required. Mark Weiser called this calm technology [63], to highlight the fact that computers blend into the background when not needed, and only ask for our attention when appropriate. In this sense ubiquitous computing is the exact opposite of the vision put forth in science fiction literature — instead of making humans live inside the computer in a virtual world, the vision of ubiquitous computing is to make the computer live in our world. As William Gibson, the author of well-known cyberpunk novels such as Neuromancer [18] and Mona Lisa Overdrive [19], explains in an interview with the Rolling Stone magazine:

> *"Totally ubiquitous computing. One of the things our grandchildren will find quaintest about us is that we distinguish the digital from the real, the virtual from the real. In the future, that will become literally impossible. The distinction between cyberspace and that, which isn't cyberspace is going to be unimaginable. When I wrote Neuromancer in 1984, cyberspace already existed for some people, but they didn't spend all their time there. So cyberspace was there, and we were here. Now cyberspace is here for a lot of us, and there has become any state of relative non-connectivity. There is where they don't have Wi-Fi…*
> *In a world of super-ubiquitous computing, you're not gonna know when you're on or when you're off. You're always going to be on, in some sort of blended reality state. You only think about it when something goes wrong, and it goes off. And then it's a drag"* [36]

There is no doubt that technology is increasingly pervading our shared urban spaces, and the rate at which new digital services are made available is continuously increasing. Accelerated by the digitalization and miniaturization of electronics and the explosion of communication networks, new technologies have pervaded the society in many ways. There are already over 5 billion mobile subscribers in the world and over two billion people use the Internet [6,25,48]. Over 30 billion RFID tags have been embedded in our world and a billion transistors per human, each costing one ten-millionth of a cent [6,48]. Over one billion Bluetooth chips are shipped every year and a new WiFi access point is deployed every four seconds [6,48]. The convergence of smaller, cheaper and faster computers and ubiquitous communication technologies have made it easier to control systems and to empower people, to make cities smarter.

In this paper we will discuss the concept of urban computing and its various facets from the point of view of creative prototyping. In our case, the prototype in question is a large-scale technology deployment built into the existing urban fabric of the city of Oulu, in northern Finland. This deployment of various technologies constitutes a large, geographically distributed and completely public civic laboratory for studying ubiquitous computing technologies "in the wild", and is technologically the largest such research deployment in the world. Here, our main focus will be on discussing shared urban technologies, i.e. technological artifacts deployed as an integral part of the public urban fabric, and thus useable by the whole community. Specifically, we will look at the early stages of planning and design process of the technology deployment, and at the various studies conducted to understand the needs, attitudes, and even fears people have towards new urban technology.

The contribution of this paper is threefold: first, we will offer a theoretical look into cities as settings for new technology, and identify challenges and opportunities that arise from the introduction of a new digital layer into the existing physical urban environment. Next, we will introduce work done in designing the "Ubiquitous Oulu" (Fig. 1), a prototype of a future smart city augmented with new technology and services, and discuss the steps we took to harness the imagination of the community in the design process. Using storytelling and rapid ethnography we first identify ways in which citizens envisioned that such technology would change their lives once installed across a city center. By transitioning through a complete design process including elicitation using storytelling, interviews and focus groups, and drawing from a rich theoretical background from the fields of urban design, architecture, anthropology, and computer science, we will then attempt to forecast how such urban technologies might induce societal change in the following decades.

The rest of this article is structured as follows. In Section 2 we will present an extensive theoretical literature review and attempt to place our own research in this context using concepts of *spatialization*, *temporalization*, and *embodiment*. In Section 3 we will then move on to describe the studies we conducted to elicit feedback on the planned technology injection from people. These studies include storytelling, rapid ethnography using a mock-up prototype, and technology diaries and interviews meant to uncover so-called *stories of everyday life* on how people use existing technology, and how they see technology in the future. In Section 4 we will then discuss the findings of our studies, and reflect on those using the concepts introduced in the theoretical framework. Through this reflection we will attempt to project on the role of public urban computing technologies in the upcoming decades. Section 5 concludes the paper.

2. Urban computing artifacts: design challenges and opportunities

In order to take a look forward and speculate on how urban computing artifacts might affect society in the future, we first take a look back and reflect on activities in urban

Fig. 1. Conceptual illustration of a 'smart city'.

contexts on a conceptual level. This is necessary in order to address the varied and multifaceted underpinnings of what it means for technology to function in such an environment. First, we will briefly introduce urban computing as a field of research. Then, we will provide a theoretical look into past research on cities and attempt to place our work within this rather broad context using the concepts of spatialization, temporalization, and embodiment.

2.1. Urban computing

Urban computing, defined as "the integration of computing, sensing, and actuation technologies into everyday urban settings and lifestyles" [55], is an emerging research field, which considers the use of pervasive computing technologies in spaces not traditionally associated with computing needs — urban environments. These urban places are often seen as "third places" — the spaces between home and work [53], where many vitally important social activities happen. Space and place are key concepts in urban computing [21,62]; however, as these terms are often used interchangeably in everyday circumstances, we must take care to specify what we mean by them. Borrowing from architectural and urban design sources, Harrison and Dourish [21] have expressed the difference between the terms in the following manner: "*Space* refers to the structural and geometrical qualities of a physical environment, while *place* includes dimensions of lived experience, interaction and use of space." Urban computing, then, views spaces as settings for social protocols, conventions, and values, as well as means for shaping our own individual concept of identity, community, and self [31,55].

However, there is no fundamental theory, knowledge base, principled methods nor tools for designing and building computing systems as integral elements of the urban landscape [31]. While urban computing is focused on understanding technological effects on the urban landscape, it is important to reflect on urban life itself [54]. As urban computing is a field situated at the intersection of a multitude of different disciplines, the themes and research topics can easily seem fragmented and divided. This is partly due to the inherent nature of urbanity and densely populated urban areas, which in effect mirror all the complexities and varied aspects of human life, and brings with it the barely controlled chaos that is the construct we identify as a city. It is this omnipresence of places and people that excludes urban computing from the family of purely technological subjects, and requires researchers to focus on themes that are technology-driven but not technology-focused.

2.2. On cities, technologies, and people

Public and shared urban places outside of the home and the workplace – sometimes referred to as 3rd places [53] – are places which enable people to engage in crucial social activities. Understanding the often competing needs, requirements, hopes and even fears people have towards new technology in such places is a crucial component in designing urban computing systems that are bound to alter and disrupt the everyday flows and functions that occur in these places. Furthermore, shared urban places arguably encompass the whole spectrum of human characteristics, and it is important yet very challenging for designers to navigate the maze of competing needs, attitudes, skills and expectations of people from very different demographics, and to propose designs that would serve the needs of the many instead of being exclusive to a certain niche subgroup of the population such as young, technologically savvy students [60].

The seat of urban computing, after all, is the city in all its forms, the locus of great many activities where people come together. If we are to create new forms of urban experiences and services through urban computing, it is crucial to develop ways to understand the role of the built environment in urban life. To understand different kinds of outdoor activities, we refer to architect Jan Gehl's [16] categorization, according to which there are three main types of outdoor activities. *Necessary activities* include those actions, which we must accomplish, such as commuting and shopping. These activities occur no matter how unfavorable the surrounding environment. *Optional activities*, like sitting, playing or eating outside, however, require a certain level of enjoyment of the environment. These pursuits only occur if we wish to participate in them, and the time and place allow us to do so. Finally, *resultant* or *social activities* often occur during these necessary and optional activities. These include chance meetings with friends on the street, children playing outside, and simply seeing and hearing other people. These activities occur spontaneously, Gehl [16] argues, as a "direct consequence of people moving about and being in the same spaces", for example, while we are running errands or just spending time in a park. This is, of course, a greatly simplified categorization of urban life, but it does create a more serious urgency to explore the role of urban computing in the city. How can we create situations that not only serve us well, but enable us to willingly enjoy our urban spaces more? Furthermore, it advises us that any deployment of urban computing systems that aims to succeed needs to consider these modes of urban activity.

We could say, then, that cities "deal in physicality and proximity. They are like machines for facilitating physical meetings" [7]. As this is the essence of cities, what the urban computing community must also address is the perceived threat that information technologies impose on urban life. For example, Mitchell [46] has argued that urban computing technologies will bring about a wholesale change of urban life and the built urban form. What does it mean that we do not have to use and traverse public spaces to pursue necessary activities (to use Gehl's categorization), or that we do not even have to live in cities to be connected to the rest of humanity? This contradicts, as Castells [4] has pointed out, with the fact that the world is urbanizing at an ever-increasing rate [13]. Similarly, even Mitchell argues that the "twentieth century still needs agoras — maybe more than ever" [46].

The question, then, is how can urban computing support and create situations for a larger variety of urban activities? As our social lives increasingly move towards the digital – today, we learn more about what our friends are doing through social networking services such as Facebook than through face-to-face conversations – can urban technology be used to facilitate real-world social encounters? Can public displays, for instance, be viewed as modern day meeting points where people come together to collaborate, share content, and discuss everyday happenings either physically or virtually?

Related to the blending of physical and virtual worlds through new technology, we must understand how temporal rhythms and the variation of seasons, weekdays, and even the time of the day affect the use of urban computing artifacts. Kevin Lynch has argued that time and space "are the great framework within which we order our experience. We live in time-places" [38]. Furthermore, he goes on to state that we experience time in the city in two ways: through *rhythmic repetition* and through *progressive or irreversible change*. Thus the city is not the same in night as it is in day, and neither is it the same in the summer as it is in the winter, nor does it remain the same throughout the years. The inhabited environment is under relentless change through an endless iteration of daily rhythms and yearly cycles [47,49,50]. These rhythms depend crucially on the kinds of spaces and places Oldenburg and Gehl describe: the urban environment should foster "casual encounters and should facilitate solitary as well as informal social activities" [30]. Moreover, progressive change must also be managed; metal rusts, technologies become obsolete, places might become abandoned or their very nature might change. To these, Carmona et al. [3] have added a third aspect, namely *continuity and stability*, which takes into consideration those aspects that do not change over a period of time, either by their nature or by our will. Consequently, we arrive at a temporal framework, which factors in rhythmic change, progressive change, and continuity or stability. If pervasive technologies are truly here to stay, any successful application of urban computing must take these aspects into consideration. Environments should be designed "to reflect and enhance the changing day and season add to the richness of the urban experience" [3]. These rhythms, however, have been disrupted by the introduction of the so-called '24-hour society' [32], due to which our patterns of use and activity have changed drastically, reducing the likelihood of impromptu coincidences in time and space. Fortunately, digital technologies are, by their very nature, much more adaptable than the traditional urban places; and through this virtue, they harbor a potential to support these complex ongoing cycles. In other words, digital technologies that operate in the digital counterpart of the physical world are much less susceptible to the real-world factors that affect how people function in the city — weather conditions, for instance, can limit or encourage social behavior in outdoor public spaces. People are far less likely to stop and socialize when it's pouring with rain, and conversely far more likely to do so in fair summer weather. Interaction through digital technologies is not affected by such real-world conditions, however, and thus people will have a choice to move between the digital, the physical, and the blended worlds according to which ever fits their current conditions best.

Urban spaces are also quintessentially physical spaces where people move around, interact with technology, structure cognition and experience the world through their bodies and senses in a particular social and physical context. This is process, which is fundamental to the human experience, is referred to as *embodiment*, a concept whose major proponent in Western philosophy was Merleau-Ponty [44]. In ubiquitous computing and interaction design, Dourish [11] has built upon theories of embodiment, introducing the concept of *embodied interaction* to the field. It has been pointed out that powerful designs intimately incorporate artifacts into bodily practice, to the point where people perceive artifacts as an extension of themselves; they act through it rather than on it [29].

Further, as urban computing systems reside in the public domain (i.e. are not personal devices but rather owned by somebody else), it is challenging to make people understand that they are allowed to physically interact with the devices. This is because typically you are not supposed to touch surveillance cameras, advertisement displays, or any number of other technical devices in public spaces. Thus, we need to consider also the power dynamics and hierarchies that define people's behavior in public urban spaces [20]. Sociocultural, sometimes very local, rules determine how we are supposed to behave in relation to public installations but also in relation to other people; what kind of actions we are allowed to take; who is welcomed in public places and who is not; this is something David Harvey called 'the right to the city'. He writes: "…but new rights can also be defined: like the right to the city which is not merely a right of access to what the property speculators and state planners define, but an active right to make the city different, to shape it more in accord with our heart's desire, and to re-make ourselves thereby in a different image" [22]. New technologies can – and will – alter the existing hierarchies and rules, and therefore their possible effects, both good and bad, should be considered carefully. If access to the city and urban life is partially gained only through the digital, we must be ever cautious of the many ways this might cause the already existing and identified *digital divide* [51] widen between the inhabitants of the city. The term, first popularized by the National Telecommunications and Information Administration, describes the gap between those who benefit from and have access to new technologies, and those who do not [56].

2.3. Designing for the city: Challenges, possibilities, and threats

The research community is still in the process of transitioning from the traditional software engineering for workplace environments to the messy and iterative processes used in urban computing design. Designing urban computing systems and services is technologically challenging. These systems often span multiple devices and platforms, and interaction may begin on one device and, as better-suited resources become available in the environment, may continue on a completely different device. An example of such multi-device interaction is browsing digital photographs on a mobile device, and when passing by an interactive public display, continuing browsing on a larger display surface [26]. Numerous projects have looked at building toolkits to aid developers in designing multi-device off-the-desktop systems (see for example [8,9,35,26,37,40]), but these toolkits are by default aimed at technologists, i.e. programmers, and therefore are not suitable tools in eliciting early-stage requirements from non-technologists [12].

In general, the cultural and social aspects of urban computing remain insufficiently explored and represented in the design process [14]. While technical innovation has been considered as a cultural practice and explored within social studies of science and technology (see e.g. [34,59]), the problem of relating these theories to technology and everyday life persists [14]. Mackenzie [39] introduced the term 'transduction' to describe a way of thinking about technologies processually, i.e. as events rather than objects, and to refocus investigations towards non-representational understandings of technological practice — in other words, to understand technology not just as a tool but rather as an integral part of everyday processes and activities. Applied to urban computing, this concept allows us to "shift our focus from computers as networked objects or artifacts, to computing as diverse procedures or performances in which socio-technical assemblages take shape", and thus "identify precise moments and locations in which we may possibly intervene and alter the course of events" [14]. The concept of directly intervening in the everyday flow of urban life was already an integral part of the Urban Probes methodology used by Paulos and Jenkins [55]. They draw an analogy to archeologists excavating and altering a dig site to gain a better view of the space, and insightfully remark that in order for researchers to understand urban spaces, they must directly intervene to alter/disrupt the usage, actions, or flow within the urban focus of attention.

However, the potential for urban computing systems to alter and disrupt everyday life in shared public spaces is something that needs to be considered carefully when designing such deployments. As outlined above, the complex social, cultural, political and technological underpinnings of such potentially invasive technologies may in fact turn against their intended purpose, and rather than enhancing the experience of the city cause a backlash from the community they were supposed to serve. To take an example, let us consider an application that lets people post photographs from their mobile devices to a large public display with the purpose of pointing out problem spots in the neighborhood. What if, instead of pictures of uncovered manholes and missing traffic signs someone posted, say, invasive pictures of their neighbors? The technology, intended to help make the neighborhood a better place, would have done potentially irrevocable damage and thus had turned against its original intended purpose. Similarly, we must be sensitive to the existence of the digital divide and carefully consider what demographics our designed technologies serve and, even more importantly, which ones they do not, as discussed next.

3. Creative techniques for designing the public spaces of tomorrow

Our research is based on the premise that instead of running "artificial" studies in laboratories or other controlled environments built to merely convey the idea of a public setting, we actually develop a unique civic laboratory at the very heart of our city to study the multifaceted field of urban computing. Technologically, our installation is today the largest of its kind, aimed at studying urban computing in an authentic city environment. The most visible part of our installation is a network of large interactive public displays (Fig. 2) [33,52], made available to the whole community on a 24/7 basis. The displays support several modalities of use, including direct input through a touchscreen interface, and distributed user interfaces between personal mobile devices carried by the users, and the displays. The displays also provide multiple applications and services, and enable collaborative use for several co-located users using a single shared display, or several users using dislocated displays.

An important point to note is that our work focuses on interactive, as opposed to one-way broadcast displays. While public displays have been utilized in one-way broadcasting of advertisement for decades, recent technological advances have enabled interactivity on these displays. However, the majority of experiments with interactive public displays focus on a single application running on a single display, which is not meant to be a permanent addition to the space it is deployed in — in other words, a display is deployed in a space for the duration of the experiment, and is then used by either a pre-selected group of people or made available to the general public, usually under researcher supervision. After the study is over, the display is removed. Therefore, given the ephemeral nature of such deployments, and the fact that the display in question is single-purpose in nature, i.e. only running a single application

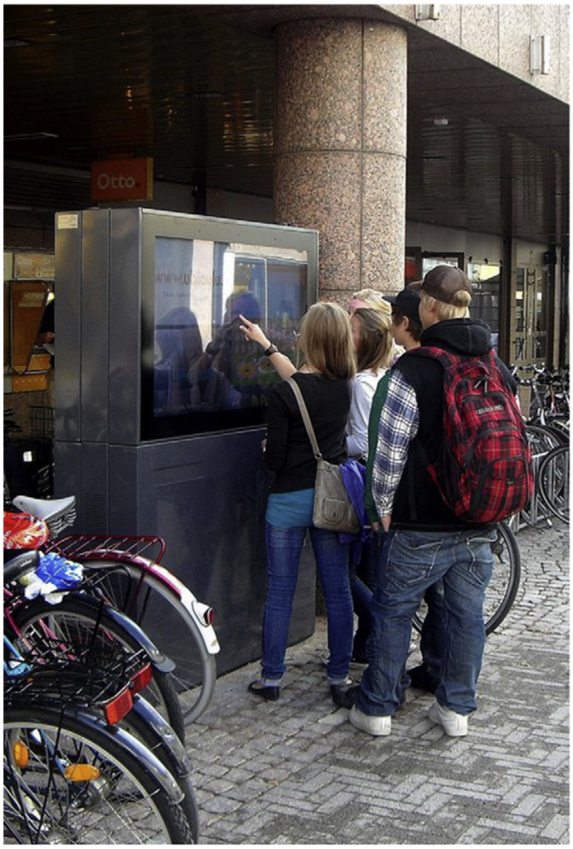


Fig. 2. Interactive public display in Oulu.

or service, it is very hard to identify longitudinal effects of the technology, how it becomes appropriated, and what impact it has on the society once the novelty effect wears off.

Prior to actually deploying technology into downtown Oulu, we first conducted a series of studies to understand how the planned technology deployment would best serve the community. These studies will be the focal point of this article. First, a storytelling competition was launched to engage the local community in the design of new urban technologies. Despite the high media visibility around the upcoming deployments and valuable prizes to winners, the storytelling competition gathered a relatively small amount of entries. However, these stories provided us multiple design ideas and helped us to understand how citizens perceive technologically augmented cityscapes.

In order to involve the local community and gain a more in-depth understanding of how people perceive the addition of new technology into their everyday surroundings, we carried out a "rapid ethnography" study where we made the concept of interactive public displays more concrete through a lo-fi prototype device designed to conjure the idea of an interactive public display. By using the mock-up as a demonstration device, we interviewed people to learn what type of content and services they felt would be valuable and useful through this infrastructure.

A third set of data consists of technology diaries and semi-structured theme interviews collected from the young adult citizens of Oulu (ages 20–30). The technology diary study combines traditional ethnography and methods from design studies, and is partially based on cultural probes methodology [15,43]. Cultural or design probes are usually based on self-documentation and aim at uncovering people's personal perspectives and experiences. The participants first kept a diary for a couple of weeks; it was supposed to map their everyday practices and encounters with information technology. After finishing the diary, participants were invited to take part in a group or couple interview where they could share their thoughts about the themes of the diary.

We discuss these studies next.

3.1. Storytelling

Storytelling is considered a valuable method of Participatory Design, since stories are both tools for remembering and contain clues to future making [1]. In order to engage the local community, we launched a storytelling competition with the theme 'Oulu in 2020'. The competition was advertised in an extensive newspaper article detailing the upcoming plans for building Oulu of the future. The competition was promoted with handsome prizes, including a laptop for the best story selected by a jury of researchers and industrial experts, and a smartphone raffled amongst all participants. To help guide the storytelling process, we also wrote and illustrated a set of our own 'Oulu in 2020' stories (Fig. 3) and posted these on the project website.

The Oulu in 2020 stories written by us consisted of 30 short snapshots into the imaginary lives of 15 personas of different demographics ranging from young children to elderly retired people. The stories were set in varying environments including home, school, café, restaurant, downtown, etc., and depicted everyday situations where new technology helped the characters in different ways.

Our stories naturally introduced technologies that we were planning to include in the upcoming deployment: interactive public displays, sensor networks, short-range wireless communication through RFID and Bluetooth, and personal mobile devices. Adhering to the vision of ubiquitous computing, our stories were all very *optimistic* — technology was there when needed, never failing, and never used for malicious purposes. Of course, as the stories were meant to create a positive vibe for the upcoming developments in Oulu, we did not want to paint a bleak picture of a future where technology would work against its users and affect urban life negatively, but rather encourage positive discourse around the topic. Further, as we wanted to engage the whole community and not just those who had an interest in new technology, our stories featured scenarios that were easily understandable and approachable even with superficial understanding of technology.

> *"Valtteri and Julia are in high school. They decide to have coffee after school. Walking on a pedestrian street towards their favorite café, they notice an aquarium, which is actually wall-mounted display, but the animated fish are so realistic that most passers-by mistake them for actual living fish.*
>
> *As Valtteri and Julia approach the aquarium the fish become more active, and as they stop in front of the display the fish scatter in all directions, revealing a localized bulletin board. Valtteri takes out his personal mobile device and starts waving it in the air. He thinks that gesture control is just*

Fig. 3. "Oulu in 2020" story illustration.

perfect for these kinds of situations — controlling a wall-mounted display, for instance. With a few gestures he pulls out the specials for nearby cafés, and together they pick one that has two cappuccinos for the price of one". (Excerpt translated from Finnish by authors).

Despite the high media visibility around the upcoming deployments and valuable prizes to winners, our storytelling competition gathered only 8 entries. The stories featured scenarios ranging from the ordinary (controlling home appliances through smart phones) to the extraordinary (holographic guide dog). Popular settings for the stories were public and shared indoor locations such as cafés, pubs, and bars, where technology was used to perform various functions including ordering beverages, paying the bill, using place-based messaging services, etc. From the stories we were able to identify issues related to the three themes presented previously, namely *spatialization*, *temporalization*, and *embodiment*.

Spatialization was discussed in several stories where the characters moved about in urban space, navigating unfamiliar paths with the help of technology. Unfamiliar cities were seen as confusing places where the help of technology was sought out in several narratives.

"An hour later the man attempted to walk in a straight line in the hotel lobby.' Where the hell am I?', Mustonen cursed to himself.' Just take it easy, don't worry you bastard', he muttered while his headache continued to grow worse. Suddenly a small dog materialized beneath the chandelier. It stared at Mustonen, looking somehow amused, and appeared to be asking him to follow.' Damn, too much Jack Daniels', Mustonen growled.
The greasy receptionist waved his hand. 'Go ahead, the Writer will show you the way'. Mustonen stared at the young man. Lunatic. The dog barked and started dancing towards the large revolving doors of the hotel. 'What the hell, let's go', Mustonen said and followed the dog." (Excerpt translated from Finnish by authors).

Temporalization also featured in several narratives, mainly in the form of characters needing to know when given activities were available, or when certain locations such as shops were open. In the stories technology was seen as close to all-knowing, able to provide information on any topic upon request.

"The young couple make their way towards a nearby restaurant in the hopes of dinner, but at the door are told that the restaurant is full of fans celebrating the success of a local figure skating team, and the overall atmosphere seems too rowdy for a romantic candlelit dinner anyways. Timi-Kullervo checks his smart device, and finds the reservation status of all nearby restaurants. He makes a reservation to an idyllic bistro that's only a block away. At the restaurant door their reservation is easily found by touching Timi-Kullervo's device to that of the doorman, and the couple is shown to their table." (Excerpt translated from Finnish by authors).

Due to the nature of the stories and their settings, embodiment was found in every story — characters interacted with their environment through their bodies, by physically manipulating digital interfaces or simply by moving about in an augmented space. A common denominator was that bodily interacting with public technology was seen as very natural and effortless, and the characters were intuitively aware of where these interfaces were located. Further, the characters were not worried about things related to privacy, but were rather eager to share even their most personal information with the imaginary systems.

"I finished my coffee and began planning my schedule for the rest of the afternoon. My wardrobe was in dire need of updating, and the shops were open for a few more hours. I knew of a public display a few blocks away, where I could have my strategic measurements up to my shoe size digitally taken. After this the computer would search all nearby stores for clothes that would fit me — I could browse the selection on the screen, and after that march on into the store to purchase the items I had picked out." (Excerpt translated from Finnish by authors).

The winning story was titled "*By golly, where am I?*" (translated from Finnish). The humorous story depicts events during a one-day period, revolving around an apparently demented elder female character, who navigates various everyday situations with the help of her intelligent wrist computer (story translated from Finnish):

"Sit here for a moment, while I go to the restroom. How about you read the train company's magazine?" Mirjami Jyrä-Suonio, a nurse from Jyväskylä if I remember correctly told someone, possibly me. Her eyes darted to my left wrist, to the screen of a gadget with an animated smiling face. "Seems you have enough battery, 'spose you'll be fine without me for a bit", she said and left me by myself at a small table with the paint flaking off. "Hi Sanelma, we are at the Oulu railway station", the smooth masculine face on my wrist gadget said, appearing to be looking around.
"How on earth can this thing speak", I said half to myself. "These modern times are just full of wonder". I looked at the device, then at my hand, which looked dry and wrinkled. "Could use some lotion, to be sure…"
The face lit up. "Would you like some lotion?""Yes please", I said, stroking the freckled back of my hand with the other, equally dry hand.
"By golly, where am I?" I looked around. In a shop.

A rotund woman dressed in a red shirt and an indecent black leather corset handed me a small black plastic bag. "Here's your lotion. That'll be 15 euros." The woman extended her arm, and the device, which looked like a wristwatch beeped almost simultaneously with my own.
"The Erotic Shop at Asemakatu just charged your account for 15 euros", my wrist device stated.
"The Erotic Shop at Asemakatu? By golly, where have I gotten to now? I have got to get out from here, my oh my". I limped to the door, but didn't have to courage to go out. "I can't leave, there's too many people! They'll see where I'm coming from!" I exclaimed to myself. The lady behind the counter tittered.
"There is one passer-by outside. Not very close", my wrist gadget stated and started showing a map of the nearby streets with a green and a red dot.
"Well, that's a relief, but where am I going to?"
"The next bus to the Raatti Island leaves in seven minutes. Showing route now." The map changed to showing a route in red, and a compass.
"By golly, where am I? In a car?" The gadget, secured to my wrist with wide leather straps, blinked to life. "Hi Sanelma, you are on a bus headed towards the Raatti island. We are almost there. I will signal the bus to stop". The apparatus vibrated shortly, and the stop-sign towards the front of the bus lit up. The bus slowed down, eventually coming to a full stop.
I stepped carefully over the curb and onto hot asphalt. I was followed by several young men and women, some of who were very loud. One was carrying a can of beer — in public, oh dear! The bus doors shut with a whoosh, and the car left. I remained *at the stop, looking at the peculiarly dressed youngsters.*
"Public Transportation Services of Oulu just charged your account for 1 euro and 23 cents. Distance travelled was approximately 0.9 kilometers", my gadget stated in its usual monotonic drawl. The screen flickered from showing a face to map and compass. "Your destination is less than 100 meters away".
The hands of a young, bald man groped my body"By golly, young man, keep your hands to yourself!" I screamed, again wondering where I was. My memory was playing tricks.
"Might I check your bag?" The same rude young man asked and took the plastic bag from my hand, as if I'd given my consent. "Okay, thank you very much", he said returning the bag to me. "Grandma's got a wild night planned I see, ha ha. Anyways, we'll charge you automatically based on your location, feel free to roam the festival area".
"What are you laughing at young man? My poor old skin is dry", I replied, suddenly remembering what was in the bag. The man exploded with laughter, doubling up on himself with mirth. How very, very rude.I looked around. A huge crowd of people, mostly youngsters, colorful tents and music. A pervading scent of food. "Where am I?"
"You are at the Qstock festival. The program is as follows…"
"Oh, there you are!" A distantly familiar voice interrupted my wrist apparatus. Mirjami. "I only went around the corner and you were gone. Thank goodness for that device." Mirjami looked at her own wrist apparatus. "How would demented people get by without it?"

This story was selected as the winner because it depicted scenarios and technologies that were imaginative yet plausible, and while the story was centered around an elderly woman, it also featured personas from various demographic groups as supporting characters. In other stories the users of technology were mostly depicted as young citizens spending time in restaurants, bars and cafes. The winning story was constructed around temporal and spatial rhythms that divided the story into smaller sections creating transitions from one situation and space to another but creating also dark forgotten temporal spots due to the dementia of the main character. The urban rhythms of the story were constituted together with other pedestrians, as well as with cars and buses. Space and technology were experienced through bodily practices when the talking wristband both gave the user the information she was asking for, and gave orders about her movements as an independent actor. Her moving in the city space shifted between Gehl's [16] necessary activities, like shopping for body lotion, and resultant/social activities, like spontaneous visit to the rock festival, though this was not dependent on the user herself. Like other stories, the winning narrative happened during a period of one day, and thus the temporal rhythms were not connected to seasonal or even weekday changes. In addition, all stories were situated in the near future, and new ubiquitous technology was appropriated in the situations that already were part of people's everyday practices [10].

Overall, the stories painted a picture of a society very much adhering to the vision of *calm computing*. Technology was there when needed, would stay in the periphery of attention until requested, and could provide answers to all questions. However, the competitive nature of the storytelling task most likely influenced the way people envisioned future technology. People likely wrote positive stories because they felt that this was what the competition organizers wanted to read, and thus attempted to improve their chances of winning. While this very optimistic view was of course an encouraging sign that people were willing to welcome new technology into their everyday lives, the low number of stories led us to design another study that would allow us to better engage with a larger group of people in the actual shared urban spaces we were planning to bring new technology into.

3.2. Rapid ethnography with situated mock-up devices

Rapid ethnography [45], sometimes also referred to as "quick-and-dirty ethnography" [23], is a field study method where researchers, instead of spending at best months in the field documenting a certain phenomenon as in traditional ethnography, do short focused studies to rapidly gain an understanding of the target setting. Especially in the field of HCI (Human–Computer Interaction), with quick-paced product development cycles and deadlines, it simply is not possible to spend months or even weeks in the field gathering data, and a similar amount of time understanding the field data. The benefits of examining field situated user activity, however, remains crucial [45]. The well established benefits of doing field studies combined with the aforementioned problems with the storytelling competition led us to device another study where we would make the concept of interactive public displays more concrete through a lo-fi prototype device designed to conjure the idea of such a public display. The apparatus in question was built by rigging a large whiteboard on a wheeled stand (Fig. 4), and researchers took the mock-up to various public locations around downtown Oulu. The locations were selected as candidates for the upcoming deployments, and we also wanted to prompt people to give their opinion on how such public displays would fit in these locations. In a sense this mock-up study can be seen as an appropriation of the generative technique introduced by Sanders [57], where users are given a toolkit with simple components and are then asked to make artifacts such as collages or mock-ups to describe their past or ideal experience for the domain under consideration [2]. In this technique, the main interest is the story that the person tells while s/he is describing the artifact. Similarly, we were interested in the stories of everyday life that people described when drawing their service suggestions and ideas on the whiteboard.

The study ran for a period of 2 days (~8 h per day), during which 74 interviews were conducted. Interview subjects included both individuals and groups from a wide range of demographics. Researchers first explained the concept of interactive public displays to the interviewees using the mock-up as a demonstration device, and then asked the participants to describe how they would want to use such technology by sketching concrete application ideas on the whiteboard. Most interviews were videotaped for later reference (some participants did not agree to be videotaped), and researchers also collected extensive field notes and photographs.

Overall, participants reacted positively to this type of municipal service. They specifically named non-commercial content that would benefit their everyday lives as more desirable than commercial or sponsored content. Participants also felt that reliability and consistency was more important than application sophistication and the "wow-factor" — people wanted services that *work for them*, *are reliable*, and *easy to use*. Possibly due to the ease-of-use factor, people also predominantly named touchscreen technology and services as more desirable than distributed interfaces using device ensembles such as mobile phones as private input devices to public displays. Several participants also brought up another recent (commercial) technology deployment made into the shared public spaces of Oulu, namely displays at bus stops showing timetables and arrival times of buses. As the displays had suffered from severe technical problems, interviewees often remarked that they do not "need more black screens" in the city.

The mock-up display proved successful in capturing the curiosity and imagination of passers-by, and the familiar nature of the whiteboard, combined with the "life-sized drawing" technique clearly enticed people to approach the researchers and give present their ideas. However, a caveat that should always accompany low-tech prototyping should be that the given measurements, materials, look and feel of the prototype necessarily affect and possibly limit the participants' imaginations. Nevertheless, the participants proposed a wide variety of application ideas, most of which were very practical and down-to-earth such as event information, news items, local services, public transportation, etc. Also, as a hot topic around the time was the planned construction of a large underground parking facility beneath the downtown area, several respondents suggested that the displays could be used to provide information on free parking spaces. Overall, the ideas of the participants were tightly connected to the public urban space where the mock-up study took place, and were connected to all three types of outdoor activities categorized by Gehl [16]. Participants felt that new ubicomp technology could assist them in their *necessary* activities by providing for instance information on opening hours of stores, bus timetables, and available parking spaces; in their *optional* activities with real-time information on local events, like concerts, movies and plays, as well as on restaurant menus; and in *resultant/social* activities with leisure applications such as games. Participants seemed split in their opinion about using the displays for social entertainment activities, with some being strongly in favor of them and others claiming that the displays should be used for "serious" purposes only. Temporally, the ideas were tied to both momentary activities, like parking spaces and queues in nearby restaurants; more long-term information on, for example, municipal decision making; and relatively permanent information such as maps of the city center. The mock-up study was conducted in autumn before snow season, and consequently, seasonal changes were not apparent in the imagined future of new ubiquitous city space. There were also participants who claimed that the whole idea of deploying public displays into the city was ludicrous, and that the money required for such an installation would be better used for other purposes such as elderly care. These participants were a small minority, though, as only two people expressed this opinion.

Fig. 4. Participant interacting with the mock-up device.

Further, the importance of spatialization became very apparent. As the mock-up device was taken to several different locations during the 2-day period, we were able to observe the effect of location on the willingness of people to interact with us, and the mock-up. For instance, in the main square of downtown Oulu people were in a very leisurely mood, happy to stop to chat and interact with the device. Interviews in this location lasted markedly longer than those in other locations. Conversely, when we moved the mock-up just one block away from the square, to a busy intersection where four pedestrian streets meet, people did not even glance at us or the mock-up, but rather passed by in a hurry and even went as far as to change their route in order not to pass by the mock-up too close. Even though the two locations are geographically very close to each other, the nature of the spaces is completely different. In the square people come to meet, spend time, and just look at other people — in other words, perform *social* activities to again use Gehl's classification. In the intersection, however, people were passing through in a hurry, coming out of one shop and going into the next, or looking for a restaurant to have dinner in — performing *necessary* activities, which do not leave room for other distracting actions such as stopping to talk with random people.

The stories people told while explaining their ideas were of particular interest. These stories mirrored the complexities of everyday urban life, with rushed schedules, competing tasks, the need to coordinate activities with several other people, and the never-ending need for everyday information items to help mentally and bodily navigate the urban messiness. People also had a very omnipotent view of technology, assuming that whatever they needed could be built, though their ideas were simultaneously entwined with their everyday practices of that moment. We encouraged these wild ideas rather than attempted to suppress them, as we felt it was interesting to see how, in a perfect world, people would utilize such urban technology.

3.3. Technology diaries and interviews

To continue probing the idea of stories of everyday life introduced in the previous section, we began systematically collecting a large amount of ethnographic material consisting, among others, of technology diaries and semi-structured group interviews with 20 to 30 years old citizens of Oulu. The study involved 48 participants between ages 20 and 30 years. 37 of them were women and eleven men. Participants were recruited mostly through mailing lists of different academies in Oulu. A majority of participants were studying at the University of Oulu or at the Oulu University of Applied Sciences or had already graduated from either. Most were still students; a few were working full-time; two were unemployed; one was on maternity leave and one a stay-at-home mother. Participants came from highly different areas of expertise, from communication to industrial engineering and midwifery. Only sixteen participants were originally from Oulu, while 25 had lived in the city less than five years and four of them even less than a year. This reflects the role of Oulu as the most popular student city in northern Finland, which attracts young people, especially from northern Finland and nearby small townships.

In their diaries and interviews, participants were asked to contemplate broadly their own technology usage in the city. Through this material we hoped to gain a more nuanced understanding of everyday practices connected to information and communication technologies in this particular northern city; the goal of this material was not to function outright as a design tool, but it could also be used to such purposes.

Further, the collected data gives us a more in-depth view into the lives of the citizens. Given Mackenzie's [39] thoughts on understanding technology not just as a tool but rather as an integral part of everyday processes and activities, we need somehow to get a grasp on those everyday processes, and understand how they are connected to spatialization, temporalization, and embodiment. We have adopted an approach called "thick description" [17], which usually refers to an (anthropological) study that does not just explain the behavior but the *context* as well, so that the meaning of the behavior becomes understandable also for an outsider. In other words, we do not only aim to describe the action, but the socially constructed meaning behind it as well.

The collected data complements findings from the rapid ethnography and storytelling. While these provided us with a broad understanding of how people view technology in their everyday life shared urban environments, the view remained somewhat *fragmentary*. The main weakness with rapid ethnography is that it does not allow us to contextualize the gathered information; the needs presented by citizens remain detached from the larger flows, rhythms and connections of everyday life. Storytelling, on the other hand, was found challenging by the citizens, possibly because it was so future-oriented and detached from the present. As Buskermolen and Ozcelik [2] state, storytelling methods in participatory design produce more useful outcomes if participants are first allowed to contemplate their existing experiences and build links between the past and the future. After this it is easier for them to dream what kind of services or technologies could be useful or desired in the future.

Thus, we found it useful to ask people to tell "stories" about their present, real lives with the devices they already have and then continue with some hypothetical questions. This approach emphasizes the utility value technology can bring into the everyday lives of the people. However, the "too slow" *traditional* ethnography and "superficial" *rapid* ethnography methods are not contradictory. Rather, we are constructing new methods to study urban public technologies that mingle different ways of collecting data and exploit our experiences on different types of studies.

Results from the diary study and follow-up interviews show that for the clear majority of young adults, a mobile phone or a smartphone was the most essential technological tool in public places. First and foremost it was used for finding and getting in touch with important people; we call this practice "social navigation". In addition, the possibility to check any information anywhere was valued; it was obvious that especially smartphones had reduced the participants' need for careful pre-planning, and increased the flexibility of everyday life. On the other hand, some mobile phone users claimed they did not need a smartphone for online information access because they could always call a friend who had access to the internet. The phone itself acted as a mnemonic for many; it also enabled making written notes or photographing things to remember. Many previous studies have shown how mobile or smartphones collapse the restrictions posed by time and place and turn users into cybernetic creatures who are dependent on the digital "layer" offered by easy-to-access social networks and information [5,28,27]. Thus, the mobile phone has been a powerful device in moving urban life towards visions put forth in science fiction.

Further, from this material we can interpret both optimistic views towards new technology, and pessimistic attitudes that resonate with the cyberpunk dystopias discussed in the beginning of this article. These attitudes are apparent throughout the material, but are especially often stated in the answers concerning dreams about the future technology; what participants would wish the future would be like. The storytelling competition did not reveal the anxiety and distress that at least one fourth of the young adults participating in this study expressed. In the diary study and follow-up interviews, some participants repeatedly complained how computer usage is consuming all of their free time; some felt that the demand of constant availability was stressful; technology was also blamed for the feeling of restlessness or being constantly in a hurry. These participants often wished that in the future the role of technology would diminish. In addition, it is worth noting that quite a few of the young adults expressed that they cannot follow the rapid development of the new technologies anymore. They felt they were not competent users of current technologies, like smartphones. This caused anxiety for some and made them skeptical towards technological innovations.

When discussing about possible threats posed by future technologies, a couple of participants referred directly to bleak future visions of science fiction films like *WALL-E* [61]. In this particular film the earth has been completely covered in waste due to mass consumerism and people are totally dependent on machines. This indicates how media feeds people's imagination and can also color their perceptions about new technologies.

On the other hand, a majority of the participants expressed either neutral or positive attitudes towards new technology. Perhaps not surprisingly, speed and efficiency were recurrent adjectives connected to future technologies. Some of the participants clearly just mirrored the current development rather than expressed radically new views. However, some of the dreams differed greatly from the usual future visions of the world of efficient, fast machines and included rather organic aspects; also the word "natural" was used relatively often to describe the ideal human-technology relation. Many participants also wished that technology would become more ecological. One group of technology dreams was surprisingly similar with the original vision of ubicomp (see [63]): a hope that technology would settle in the background of daily life, become more invisible and adjust to peoples' routines, and not vice versa.

4. Discussion: Towards blended reality in shared urban spaces

In the discussion part of the article we will focus on issues raised both in the theoretical section, and those uncovered during the various studies. By discussing these issues, we will attempt to take a look forward and speculate on the role of urban computing technologies in the following decades.

4.1. Change in the use of public spaces

It is an already established everyday reality that many activities have been virtualized. The number of local branches of banks and government offices has been reduced to an absolute minimum, and even some educational facilities have moved to a primarily online existence [46]. Reflecting on Gehl's [16] categorization of activities in urban spaces, we can observe a decline in necessary activities. Building upon this, many have hypothesized that the resultant social activities in the form of impromptu encounters are reduced due to virtualization of many necessary activities. This could lead to at least a partial abandonment of urban spaces, causing an impoverishment of urban life. In the stories submitted to the storytelling competition, however, this divide between physical and virtual services was not present; rather, online and the offline were seen as complementary. This was apparent in the case of clothes shopping, in which the superior search functions of online shops were enhanced by the instant physical accessibility of high street shops and their products.

Consequently, an argument can be made for the likelihood that these blended commercial services continue to evolve and become more of a function of everyday life. However, we also argue that public services could be made more attractive and efficient with the use of urban technologies and embodied interaction. Libraries have already learned to blend physical and electronic books and materials, and powerful databases with their existing networks of buildings and logistics. One might argue, then, that it would be of crucial importance that clinics and hospitals, which are notoriously hard to navigate, and where up-to-date information plays an enormous role in the overall well-being of patients and their loved ones, could be made more friendly, inviting and efficient with the aid of urban technologies.

Additionally, participants in the diary study clearly claimed that technological tools (in this case, a smartphone) improve the possibility of arranging impromptu meetings while out on the town. This was expressed as "social navigation". Taken a step further, we can assume that when the formerly pressing need to work or run errands in a specific place is reduced, social navigation in urban spaces through technological devices will become much more important [cf. []]. After all, when outside forces do not mandate our presence in a certain place, we are also able to escape some place-specific social norms; in effect, we can use technology to pick and choose, which norms we wish to adhere to. Consequently, it is prudent to argue that we wish to engage in more elaborate social activities when given the opportunity to do so. This runs counter to the bleak, anti-social visions so often present in cyberpunk literature, and also in more academic visions, such as Mitchell's idea of a "city of bits" [46].

The prevailing view that arises from our analyses is that activities in public spaces are not necessarily diminished, but rather likely to take new form. Extrapolating, we foresee further sophistication in the use of public spaces for an innumerable array of social uses, such as activism, socializing and artistic purposes, i.e. for optional and resultant/social activities. A good example of this would be the flash-mobs fad, wherein a group of people organizes a real-world performance, often with activist overtones, through online means. It is reasonable to suggest a further development of similar social uses of a blended reality also in the future.

4.2. The perceived threat of information technologies on urban life

Looking back to the bleak dystopian depictions of future found in cyberpunk literature, we must also consider the flipside of urban technology. While in most of our studies participants adopted a highly optimistic view of new technology, the material collected in the diary study also revealed contradicting attitudes. Many participants were rather skeptical towards current technologies and expressed distress; they felt the role of technology has grown too broad, and that they cannot follow the fast cycle of development anymore. The willingness and capabilities of adopting new technology differed greatly even among this rather homogenous (age, level of education) group. Further, given that even young people felt they were getting left behind in technological advances, we can only speculate how far from the "technological focal point" more marginalized groups such as elderly people are drifting.

Additionally, the voluntary opting-out of some younger people could lead to a widening digital divide among people within a generation. If the city and the access to urban life becomes primarily a matter of being within a blended reality what, then, will become of the right to city from the point of view of 21st century Luddites? The urban computing community must make sure that access to urban life is not barred from those who lack personal technological devices. Our studies show a need for more public forms of computing; public in the sense of physical and socioeconomic accessibility as well as a locus for sociable activities. Continuing in this vein, the lo-tech mock-ups of urban screens served as a prototype of such a public technology. Similarly, the constant availability in time demanded by personal devices and the subsequent distress and anxiety reported by our interviewees during the diary study could be diminished by urban technologies that are time-agnostic and work primarily in space, rather than vice versa.

With public services, however, the issue of control remains ever presents. Who controls the technology, and the space in which the technology operates (cf. [41])? In cyberpunk visions public spaces are often controlled by mega-corporations, who use technology to further their sinister and selfish agendas. Public screens blare brainwashing messages to the unwilling masses, and urban spaces are continuously monitored by multitudes of security cameras and sensor networks that report the movement of civilians in excruciating detail, depriving people of any kind of privacy. This Big Brother view on technology is scary, and if not addressed, can lead to people rejecting the technology and even avoiding spaces they know are augmented. In a sense we are in a fortunate position to make such a technology injection into shared urban space, as universities and research institutes are seen as 'less evil' than large corporations or even governmental agencies. On the other hand, universities often lack funding and resources to truly build a large-scale deployment, whereas corporations have no lack thereof. However, in the future, the issue of control and ownership will take an even larger role in dictating the power dynamics and hierarchies in shared spaces, as it is highly likely that future large-scale urban technology deployments will be privately owned rather than communal — that is, owned and controlled by corporations instead of communities [24].

Following the privatization of urban technology, this technology would be geared for commercial gain — either through advertising, as is the case with modern day one-way public screens that are solely used for this purpose, or through direct consumer marketing of goods, services, or other consumables. In a worst-case scenario public urban technology would no longer be accessible to everyone, but rather only to those with the financial means access it. This development would again only serve to widen the digital divide. Additionally this would further catalyze the process in which public outdoor and indoor places are becoming ever more subsumed into private ownership. In the very city that our living laboratory is located, bus stops and the area they occupy are controlled by an international advertising company; furthermore, the very pedestrian area is owned by a collective of businessmen.

4.3. Effect of temporal rhythms on the use of urban computing artifacts

Our participants' stories were closely linked with their everyday lives in their city, using them as a backdrop from which they drew their inspiration. Similarly, they based their stories on their personal history in that location. This shows the importance of the past in imagining the future, the progressive and irreversible changes [38] that add to the historical and cultural richness of the city, layer upon layer, and the personal histories of those who inhabit it.

As discussed before, shared urban spaces are highly susceptible to temporal variations and natural rhythms of days, weeks, months, and seasons of the year. Cities both operate in time, where some parts change permanently and some stay constant for longer periods of time, and also force time onto others — deliveries must be made before rush hour, people must commute to and from work at certain times, and the in-between time is then dedicated to leisure activities. Similarly, urban computing systems must respect and respond to these rhythms, progressions and constancies.

Due to the increasing digitization of services, people are no longer tied to the operating hours of commercial or municipal service points such as stores or banks, but are rather able to conduct their business regardless of place or time. Similarly, urban computing technologies that are tied to certain physical contexts will have to both support activities that are sensitive to the temporal dimension by, for example, providing up-to-date information on local events, traffic, opening hours, or scheduled meetings with friends and colleagues, but also support time-agnostic activities that occur in that specific context, such as place-based asynchronic messaging, navigating the physical space, or even entertaining oneself to 'kill time'.

Further, on a more concrete level, the variation of seasons greatly affects the activities people do in shared urban spaces, and thus also their usage of public technology. Especially in countries with much variation in the conditions around the year, such as the arctic climate found in northern Finland, the turning of seasons must be taken into consideration when designing urban technology. Such harsh conditions do not invite people to linger in outdoor settings, and thus optional and social activities in places that are teeming with activity during the warmer months is greatly diminished. However, the need to perform necessary everyday tasks and actions, socialize, and obtain everyday information pertains despite the changing seasons. In the future, technology should have the capability to adjust according to varying conditions, and proffer different ways of interacting based on these factors — to take a simple example, a public display might offer a sophisticated touchscreen interface and foster shared use during the warmer months, and switch to fast-to-use gesture control when weather conditions do not permit the user to linger.

5. Conclusion

Fictional storytelling, rapid ethnography with situated mock-up devices, and technology diaries and interviews proved to be useful methods in engaging a city's community for the purposes of imagining the city's future. Each method was found to offer different, complementary views. The storytelling method enabled us to gain a good understanding of the role that technologies play in the lives of the inhabitants of a northerly city. However, as storytelling is highly future-oriented and in our case deals with a topic people have little frame of reference with – novel urban technology – we did not reach a very wide audience through this study.

Rapid ethnography through the mock-up study deepened our knowledge by allowing us to have meaningful conversations with people about how they perceived new technology, and how they felt about including such technology in public spaces. However, as the encounters on streets were very fast and fleeting in nature, we could not really contextualize the information properly, or probe the participants' feelings very deeply. Furthermore, a study of this nature necessarily harbors some restrictions on the participants' imaginations due to the given material properties of the mock-up; and the carnivalistic, highly social nature of the event does not enable the participants to envision the future scenario where they necessarily meet the designed technological artifact after implementation all by themselves, sans researchers and the honey-pot effect. However, it can be very desirable from the designers' point of view to engage the community in such a positive way, establishing a respectful, dynamic relationship between the community and the implementers of new technology.

To counter these issues, we conducted a diary study to better contextualize how people utilize technology they currently have, and through these diaries then probe how they would see technology in the future. This last study was highly valuable in that it also uncovered negative thoughts and feelings related to technology use, something that did not become apparent in the previous studies. Thus, each method offered a different viewpoint into the shared topic of urban technology. Data gathered from each study was different, but put together these studies nicely complement each other and provide a rather holistic view on the topic.

Resulting from the use of these methods, we were able to explore the current use of technologies in the city and project their possible future use. The perceived threat posed by digital technologies towards the use and sociability of urban places is not a necessary evil; rather, based on our participants' stories, which highlight their hopes and dreams, we conclude that the way people use and imagine using technology does not warrant the extreme dystopian views presented in many cyberpunk stories. Rather, we are convinced that people will continue to develop ever more sophisticated ways to interact with each other in urban spaces, and that we must explore the possible uses of public technologies in this, bearing in mind the various divides that exist among different users of technology. For instance, the use of mobile phones in what our participants called "social navigation" was interesting, and extrapolating from that we conclude that social activities are among the most important aspects of urban life that need to be supported through the introduction of new technology and services.

By harnessing the imagination of the inhabitants of the city, we can begin to remold our deeply held ideas about urban life and project them into possible, desirable futures. Our ethnographic material pointed towards an undercurrent of negative feelings towards existing technology, and this is something that needs careful consideration in building future systems. We, as designers, must be careful not to widen the digital divide by further excluding certain demographic/social groups from the technology-oriented society of tomorrow. However, a majority of people participating in our studies had positive views towards future technology; ethnographic material also revealed an interesting set of visions where organic images derived from nature blended with technology.

Finally, we can state that contrary to previous speculation in academic and cyberpunk literature, digital technologies will not necessarily induce an abandonment of physical urban spaces. Rather, we project an increased sophistication in the sociable uses of urban spaces and technologies, where people blend their online and offline worlds into a single lived reality.

Acknowledgments

The authors gratefully acknowledge financial support from the Academy of Finland, the City of Oulu, and the UBI (Urban Interactions) consortium.

References

[1] E. Brandt, T. Binder, E.B.-N. Sanders, Tools and techniques: ways to engage telling, making and enacting, in: J. Simonsen, T. Robertson (Eds.), Routledge International Handbook of Participatory Design, Routledge, London & New York, 2012.
[2] D. Buskermolen, J. Terken, Co-constructing stories: a participatory design technique to elicit in-depth user feedback and suggestions about design concepts, Proc. PDC'12, 2012, pp. 33–36.
[3] M. Carmona, T. Heath, T. Oc, S. Tiesdell, Public Places — Urban Spaces: The Dimensions of Urban Design, Architectural Press, Oxford UK, 2003.
[4] M. Castells, Space of flows, space of places: materials for a theory of urbanism in the information age, in: S. Graham (Ed.), The Cybercities Reader, 2004.
[5] M. Chayko, Portable communities, The Social Dynamics of Online and Mobile Connectedness, State University of New York, Albany, 2008.
[6] Cisco Visual Networking Index: Forecast and Methodology, 2012–2017. Retrieved on Jun 8, 2013 from http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf 2013.
[7] C. Davies, Colin, Thinking about Architecture: An Introduction to Architectural Theory, Laurence King Publishers, London, 2011.
[8] A.K. Dey, R. Hamid, C. Beckmann, I. Li, D. Hsu, aCAPpella: programming by demonstration of context-aware applications, Proc. CHI'04, 2004, pp. 33–40.
[9] A.K. Dey, D. Salber, G. Abowd, A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications, Hum. Comput. Interact. 16 (2–4) (2001) 97–166.
[10] P. Dourish, G. Bell, Divining a Digital Future: Mess and Mythology in Ubiquitous Computing, MIT Press, Cambridge & London, 2011.
[11] P. Dourish, Where the Action Is: The Foundations of Embodied Interaction, MIT Press, Cambridge, USA, 2001.
[12] S. Dow, T. Saponas, Y. Li, J.A. Landay, External representations in ubiquitous computing design and the implications for design tools, Proc. DIS '06, 2006, pp. 241–250.
[13] In: M. Freire, R. Stren (Eds.), The Challenge of Urban Government: Policies and Practices, The World Bank Institute, Washington, DC, 2001.
[14] A. Galloway, Intimations of everyday life: ubiquitous computing and the city, Cult. Stud. 18 (2–3) (2004) 384–408.
[15] W. Gaver, T. Dunn, E. Pacenti, Cultural probes, Interactions 6 (1) (1999) 21–29.
[16] J. Gehl, Life Between Buildings: Using Public Space, Van Nostrand Reinhold, New York, 1987.
[17] C. Geertz, Thick description: toward an interpretive theory of culture, in: C. Jencks (Ed.), Culture: Critical Concepts in Sociology, Routledge, London UK, 1973.
[18] W. Gibson, Neuromancer, Ace Books, New York, 1984.
[19] W. Gibson, Mona Lisa Overdrive, Victor Gallancz, Ltd., 1988.
[20] N. Green, H. Leslie, Mobile Communications: An Introduction to New Media, Berg, Oxford, 2009.
[21] S. Harrison, P. Dourish, Re-Place-ing space: the roles of space and place in collaborative systems, Proc Computer Supported Cooperative Work '96, 1996, pp. 67–76.
[22] D. Harvey, The right to the city, Int. J. Urban Reg. Res. 27 (4) (2003) 939–941.
[23] J. Hughes, V. King, T. Rodden, H. Anderson, The role of ethnography in interactive systems design, Interactions 2 (2) (1995) 56–65.
[24] L. Humphreys, Mobile social networks and public space, New Media Soc. 12 (5) (2010) 763–778.
[25] Internet world stats. Retrieved Jun 8, 2013 from http://www.internetworldstats.com/stats.htm&gt 2013.
[26] M. Jurmu, S. Boring, J. Riekki, ScreenSpot resource discovery for smart spaces and mobilevue media sharing application, Proc MobiQuitous '08, 2008.
[27] J.E. Katz, Magic in the Air: Mobile Communication and the Transformation of Social Life, Transaction Publishers, London, UK, 2006.

[28] T. Kindberg, M. Chalmers, E. Paulos, Guest Editors' Introduction: Urban Computing. Pervasive Computing, IEEE Pervasive Comput. 6(3) (2007) 18-20.
[29] S. Klemmer, B. Hartmann, L. Takayama, How bodies matter: five themes for interaction design, Proc DIS '06, 2006, pp. 140–149.
[30] P. Knox, Creating ordinary places: slow cities in a fast world, J. Urban Des. 10 (1) (2005) 1–11.
[31] V. Kostakos, E. O'Neill, A. Penn, Designing urban pervasive systems, Computer 39 (9) (2006) 52–59.
[32] L. Kreitzman, The 24 Hour Society, Profile Books, London, 1999.
[33] H. Kukka, V. Kostakos, T. Ojala, J. Ylipulli, T. Suopajärvi, M. Jurmu, S. Hosio, This is not classified: everyday information seeking and encountering in smart urban spaces, Personal and Ubiquitous Computing, Online First, 2011.
[34] B. Latour, Pandora's Hope: Essays on the Reality of Science Studies, Harvard University Press, Cambridge, MA, 1999.
[35] J.C. Lee, D. Avrahami, S. Hudson, J. Forlizzi, P. Dietz, D. Leigh, The Calder Toolkit: wired and wireless components for rapidly prototyping interactive devices, Proceedings of the 5th Conference on Designing Interactive Systems (DIS '04), 2004, pp. 167–175.
[36] A. Leonard, William Gibson (cover story), Rolling Stone 1039 (2007) 162.
[37] Y. Li, H.I. Hong, J.A. Landay, Topiary: a tool for prototyping location-enhanced applications, Proceedings of the 17th annual ACM symposium on User interface software and technology (UIST '04), ACM, New York, NY, USA, 2004, pp. 217–226.
[38] K. Lynch, What Time is This Place? MIT Press, Cambridge MA, London UK, 1972.
[39] A. Mackenzie, Transduction: invention, innovation and collective life. Retrieved Jun 8, 2013, from http://www.lancs.ac.uk/staff/mackenza/papers/transduction.pdf.
[40] B. MacIntyre, M. Gandy, S. Dow, J.D. Bolter, DART: A toolkit for rapid design exploration of augmented reality experiences, Proceedings of the 17th annual ACM symposium on User interface software and technology (UIST '04), ACM, New York, NY, USA, 2004, pp. 197–206.
[41] D. Massey, Power-geometry and a progressive sense of place, in: J. Bird, et al., (Eds.), Mapping the Futures: Local Cultures Global Change, Routledge, London, 1993, pp. 59–70.
[42] Matrix, Directed by Wachowski, L and Wachowski, A. Village Roadshow Pictures, 1999.
[43] T. Mattelmäki, Design Probes. (Dissertation) University of Art and Design, Helsinki, 2006.
[44] M. Merleau-Ponty, Phenomenology of Perception, Routledge, London, 2005.
[45] D. Millen, Rapid ethnography: time deepening strategies for HCI field research, in: Daniel Boyarski, Wendy A. Kellogg (Eds.), Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS '00), ACM, New York, NY, USA, 2000, pp. 280–286.
[46] W.J. Mitchell, City of Bits: Space, Place, and the Infobahn: Space, Place and Infobahn, MIT Press, Cambridge, MA, London UK, 1995.
[47] J. Montgomery, Animation: a plea for activity in urban places, Urban Des. Q. (53b) (1995) 15–17.
[48] Wireless, Wi-Fi, RFID & Cellular Industry Overview. Retrieved Jun 8, 2013, from http://www.plunkettresearch.com/wireless-cellphone-rfid-market-research/industry-statistics&gt 2013.
[49] J. Montgomery, Café culture and the city: the role of pavement cafés in urban public social life, J. Urban Des. 2 (1997) 83–102.
[50] J. Montgomery, Making a city: urbanity, vitality and urban design, J. Urban Des. 3 (1998) 93–116.
[51] NTIA, Falling Through the Net: A Survey of the "Have Nots" in Rural and Urban America, National Telecommunications and Information Administration, U.S. Department of Commerce, Washington, 1995. (Retrieved from http://www.ntia.doc.gov/ntiahome/fallingthru.html).
[52] T. Ojala, V. Kostakos, H. Kukka, T. Heikkinen, T. Lindén, M. Jurmu, S. Hosio, F. Kruger, D. Zanni, Multipurpose interactive public displays in the wild: three years later, Computer 45 (5) (2012) 42–49.
[53] R. Oldenburg, D. Brissett, The third place, Qual. Sociol. 5 (4) (1982) 265–284.
[54] E. Paulos, E. Goodman, The familiar stranger: anxiety, comfort, and play in public places, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, 2004, pp. 223–230.
[55] E. Paulos, T. Jenkins, Urban probes: encountering our emerging urban atmospheres, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, 2005, pp. 341–350.
[56] R.D. Pinkett, Digital divide, the. In Editor-in-Chief: Hossein Bidgoli (Ed.), Encyclopedia of information systems, Elsevier, New York, 2003, pp. 615–633. http://dx.doi.org/10.1016/B0-12-227240-4/00042-3.
[57] B.N. Sanders, E.B.N, Generative tools for co-designing, Proc. CoDesigning, 2000, pp. 3–12.
[58] In: B. Sterling (Ed.), Mirrorshades: The Cyberpunk Anthology, Ace Books, New York, 1988.
[59] I. Stengers, The Invention of Modern Science, University of Minnesota Press, Minneapolis, 2000.
[60] T. Suopajärvi, J. Ylipulli, T. Kinnunen, Realities behind ICT dreams designing a ubiquitous city in a living lab environment, Int. J. Gend. Sci. Technol. 4 (2) (2012) 231–252.
[61] WALL-E, Directed by Stanton, A. Walt Disney Pictures, 2008.
[62] A. Williams, P. Dourish, Imagining the city: the cultural dimensions of urban computing, Computer 39 (9) (2006) 38–43.
[63] M. Weiser, The Computer for the 21st Century, Scientific American, September 1991.
[64] In: G. Westfahl, W.K. Yuen, A.K.-E. Chan (Eds.), Science Fiction and the Prediction of the Future: Essays on Foresight and Fallacy, McFarland & Co. Inc. Publishers, Jefferson, NC, 2011.

Dr. Hannu Kukka (Ph.D. in Computer Science) is a post-doctoral researcher at the Department of Computer Science and Engineering at the University of Oulu. He got his M.Sc from the Department of Information Processing Science at the University of Oulu in 2006, and his PhD from the Department of Computer Science and Engineering in 2012. Dr. Kukka's research interest lie in the area of ubiquitous and urban computing, and specifically in understanding how human information behavior theories can be applied to the evaluation of novel pervasive computing services in public urban spaces. Contact him at hannu.kukka@ee.oulu.fi.

Anna Luusua, MSc. (Architecture) is a researcher of digitally augmented urban spaces. The focal areas of her research include sensory experience, esthetics and embodiment. She is employed as a doctoral student in the UBI Metrics research project, which is funded by the Finnish Academy. She is conducting her PhD studies at the Department of Architecture at the University of Oulu. Contact her at anna.luusua@oulu.fi.

Johanna Ylipulli (M.A. in Cultural Anthropology) is currently doing her PhD thesis in a project which studies the design and implementation of new ubiquitous technology in the city center of Oulu from a cultural point of view. The research is based on ethnographic methods and it takes into account views of the designers as well as experiences of the citizens. In her studies, Ylipulli is specialized in media anthropology, new communication technologies and body and gender theories. Contact her at johanna.ylipulli@oulu.fi.

Dr. Tiina Suopajärvi's (Ph.D. in Cultural Anthropology) research focus has shifted from forest and forestry to urban environment and ubiquitous computing. In both settings she discusses the gender and spatial relations connected with expertise, knowledge and agency. In social technology studies she is specialized in feminist, participatory and multidisciplinary approaches. Contact her at tiina.suopajarvi@oulu.fi.

Vassilis Kostakos is a professor of computer engineering and in the Department of Computer Science and Engineering at the University of Oulu. His research interests include ubiquitous computing, human–computer interaction, and social systems. Kostakos received a PhD in computer science from the University of Bath, UK. Contact him at vassilis@ee.oulu.fi.

Timo Ojala is a professor of computer science and engineering and the director of the MediaTeam Oulu research group in the Department of Computer Science and Engineering at the University of Oulu, Finland. His research interests include ubiquitous computing and urban informatics. Ojala received his Dr. Tech. degree in electrical engineering from the University of Oulu in 1997. He is a member of IEEE, ACM, and the Pattern Recognition Society of Finland. Contact him at timo.ojala@ee.oulu.fi.

frontiers in ICT

OPINION ARTICLE
published: 23 February 2015
doi: 10.3389/fict.2015.00003

The rise of ubiquitous instrumentation

Vassilis Kostakos* and Denzil Ferreira

University of Oulu, Oulu, Finland
**Correspondence: vassilis@ee.oulu.fi*

Edited by:
Alessandro Vinciarelli, University of Glasgow, UK

Reviewed by:
Yolanda Vazquez-Alvarez, University of Glasgow, UK

Keywords: smartphones, sensing, community, instrumentation, humans, ubiquitous computing

Mankind's greatest scientific instrument is arguably in our pockets. In this short article, we point out that their *popularity*, *flexibility*, and *behavioral sensing* capabilities make them ideal for studying humanity at various levels of complexity, in ways that were previously not possible.

POPULARITY AND ECONOMIES OF SCALE

Mobile phones overtook personal computers in popularity in 2013. Today, an estimated 22% of the world's population owns a mobile phone, while 20% owns a personal computer. In western societies, mobile phones are used by more than 90% of the population, and developing nations are quickly catching up. The economies of scale involved in manufacturing *billions* of smartphones every year have reduced their cost and dramatically increased the scientific potential of their hardware.

These economic, engineering, and scientific advances are rapidly turning smartphones into compelling scientific instruments. For instance, in the last decade smartphones have repeatedly been used to study the spread of disease (Wang et al., 2009; Wesolowski et al., 2012). Increasingly, however, this type of *post hoc* modeling overlooks many of the scientific opportunities afforded by modern devices and their sensing capabilities. Because smartphones have become ubiquitous, much more than personal computers, their instrumentation can now provide scientists unparalleled insight into humanity. Therefore, it is now possible to run studies and experiments that previously were simply impossible.

FLEXIBILITY

Researchers are using smartphones to study humanity at the *micro*, *meso*, and *macro* scales. At the microscale, scientists use smartphones to instrument everyday environments of individuals and monitor behavior or physiology over time. For example, a smartphone-controlled digital pancreas is helping fight type 1 diabetes (Clery, 2014), smartphone "labs-on-a-chip" can provide cancer diagnosis within an hour (Haun et al., 2011), and tetraplegics are using smartphones to control their wheelchair (Kim et al., 2013). At the mesoscale, scientists are leveraging the capacity of smartphones to collect data, and seek to reveal patterns of group behavior in communication (Malmgren et al., 2008) and social interaction (Lazer et al., 2009). At the macroscale, smartphones can be used as a comprehensive proxy for societal-level behavior, for example, in terms of HIV spreading (Wesolowski et al., 2012) and social diversity (Eagle et al., 2010).

These three approaches resonate with visions for computational behavioral science (Lightfoot, 2007), computational social science (Lazer et al., 2009), and engineering of social systems (Eagle and Greene, 2013), respectively. Common to these approaches is the use of personal technology as a lens through which to study humanity. They mostly rely on rich hardware sensors: smartphones now come equipped with GPS, magnetometers, barometers, accelerometers, and even photometers. Yet, it is the recent emergence of two new *types* of behavior sensing on smartphones – using *software* sensors and *human* sensors – that have the potential to revolutionize science since they allow us to conduct behavioral sensing.

BEHAVIORAL SENSING

Software sensors on smartphones conceptualize human actions through data. These sensors do not measure physical properties, but behavioral properties. For example, one's calendar activities indicate routine and personal-work balance. Other software sensors include the amount of messages one sends or receives, or how many friends one has in their address book, as an indicator of one's social engagement and support. There is a potential myriad of software sensors due to (i) our increasing use of smartphones to manage everyday activities, (ii) the maturity and diversity of application programing interfaces (APIs) available on smartphones and online, and (iii) ongoing efforts at harmonizing sensing on smartphones such as the AWARE framework. Thus, while much ongoing research focuses on the analysis of big data collected by the likes of Facebook or Google, software sensors are shifting the spotlight toward a decentralized model whereby personal devices can sense and react to smaller yet richer data in real-time.

Human sensors take advantage of smartphones' ubiquity and interactive capabilities to collect data by querying humans. This resembles crowdsourcing, but with a rather real-time and localized perspective. Smartphones have been used to collect text and photo reports of earthquake aftermath or other crises (Rogstadius et al., 2013), but at the same time they can be used to collect qualitative data on a large scale, for example, comfort levels during heat waves and happiness levels at the workplace. Human sensors can be triggered in response to hardware and software sensor readings. They encapsulate our ability to collect in real-time data from large numbers of humans, and prompt a shift away from static data *sets* to live data *sources* in the study of humanity.

OPPORTUNITIES AND CHALLENGES

What makes smartphones ideal scientific instruments? We argue that their ever-increasing sensing and computational capabilities, coupled with their popularity, makes them ideal for "peaking" into

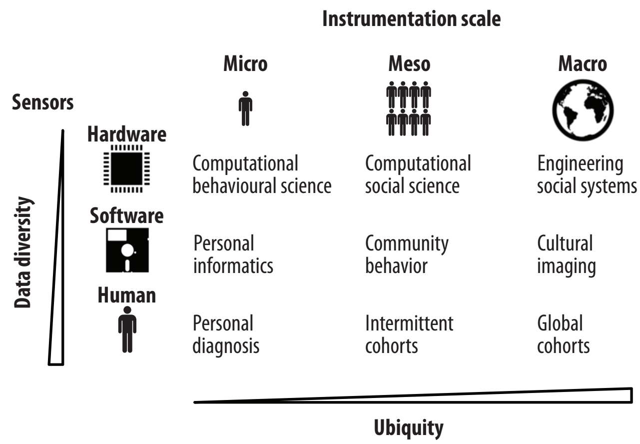


FIGURE 1 | Smartphones enable instrumentation at the micro, meso, and macroscale. Orthogonally, there are three types of sensors that smartphones can provide: hardware sensors (to measure the environment), software sensors (to measure behavior), and human sensors (to query humans). Researchers are already using *hardware* sensors to address visions for computational behavioral science (Lightfoot, 2007), computational social science (Lazer et al., 2009), and engineering of social systems (Eagle and Greene, 2013). *Software* sensors allow for personal informatics (reflecting on one's own behavior), community behavior (measuring the behavior of groups), and cultural imaging (global studies contrasting behavior). *Human* sensors allow for personal diagnosis (targeting queries to an individual's needs), intermittent cohorts (querying multiple individuals based on their context), and global cohorts (engaging million-size groups across continents).

humanity in ways that existing scientific protocols cannot. They provide a means to complement lab and clinical studies, and ground scientific progress in naturalistic settings. Global application stores (i.e., appstores) now enable us to link smartphones into one collective, widely adopted scientific instrument, that records and reacts to data streams and enables different disciplines to collaborate on global challenges (Lang, 2011). Fundamentally, each one of us is personally motivated to maintain and protect their device, fix it when broken, and keep it alive with enough battery power. Unlike conventional scientific instruments that often require substantial investment by governments or institutions, smartphone maintenance is democratized.

These developments highlight the rise of what we call "ubiquitous instrumentation." This trend is likely to see smartphones and other personal devices become increasingly valuable scientific instruments for studying humanity, enabling real-time public health studies, large-scale behavioral monitoring, and million-size cohort studies to name a few (**Figure 1**). Ironically, we argue that this trend will shift the focus from "big data" to "small data for the big picture," and from data *sets* to data *sources*. For science, this means that methods dealing with big data sets will have to be complemented with methods for dealing with small, diverse, and intermittent data sources, while experiments that take data snapshots for analysis will require novel data dashboards for real-time monitoring. And maybe someday, a visit to the doctor may even entail a prescription for your phone.

Of course, many challenges lie along the way. Crucially, not everyone has a mobile phone at the moment, with certain vulnerable populations (elders, children, poor) not having equal access to this technology. This can greatly affect the validity and reliability of any study, and care needs to be taken to ensure that representative samples are used. Furthermore, privacy is a major obstacle to this kind of work. To address this, we need ways to ensure that sensitive data remains on individuals' devices, and only sanitized data are ever shared. Finally, the question of value is crucial: what is in it for users? We need to ensure that any type of ubiquitous instrumentation provides fair value to its users, either through the provision of useful information and services, empowerment, personalized advice, and the ability to self-reflect on ones' own behavior.

REFERENCES

Clery, D. (2014). A pancreas in a box. *Science* 343, 133–135. doi:10.1126/science.343.6167.133

Eagle, N., and Greene, K. (2013). *Reality Mining: Using Big Data to Engineer a Better World*. Cambridge, MA: MIT Press.

Eagle, N., Macy, M., and Claxton, R. (2010). Network diversity and economic development. *Science* 328, 1029–1031. doi:10.1126/science.1186605

Haun, J. B., Castro, C. M., Wang, R., Peterson, V. M., Marinelli, B. S., Lee, H., et al. (2011). Micro-NMR for rapid molecular analysis of human tumor samples. *Sci. Transl. Med.* 3, ra16–ra71. doi:10.1126/scitranslmed.3002048

Kim, J., Park, H., Bruce, J., Sutton, E., Rowles, D., Pucci, D., et al. (2013). The tongue enables computer and wheelchair control for people with spinal cord injury. *Sci. Transl. Med.* 5, ra166–ra213. doi:10.1126/scitranslmed.3006296

Lang, T. (2011). Advancing global health research through digital technology and sharing data. *Science* 331, 714–717. doi:10.1126/science.1199349

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., et al. (2009). Computational social science. *Science* 323, 721–723. doi:10.1126/science.1167742

Lightfoot, D. W. (2007). Social and behavioral scientists building cyber infrastructure. *First Monday* 12. doi:10.5210/fm.v12i6.1907

Malmgren, R. D., Stouffer, D. B., Motter, A. E., and Amaral, L. A. N. (2008). A Poissonian explanation for heavy tails in e-mail communication. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18153–18158. doi:10.1073/pnas.0800332105

Rogstadius, J., Teixeira, C., Vukovic, M., Kostakos, V., Karapanos, E., and Laredo, J. (2013). CrisisTracker: crowdsourced social media curation for disaster awareness. *IBM J. Res. Dev.* 57, 4: 1–4: 13. doi:10.1147/JRD.2013.2260692

Wang, P., González, M. C., Hidalgo, C. A., and Barabási, A.-L. (2009). Understanding the spreading patterns of mobile phone viruses. *Science* 324, 1071–1076. doi:10.1126/science.1167053

Wesolowski, A., Eagle, N., Tatem, A. J., and Smith, D. L. (2012). Quantifying the impact of human mobility on malaria. *Science* 338, 267–270. doi:10.1126/science.1223467

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 January 2015; accepted: 10 February 2015; published online: 23 February 2015.

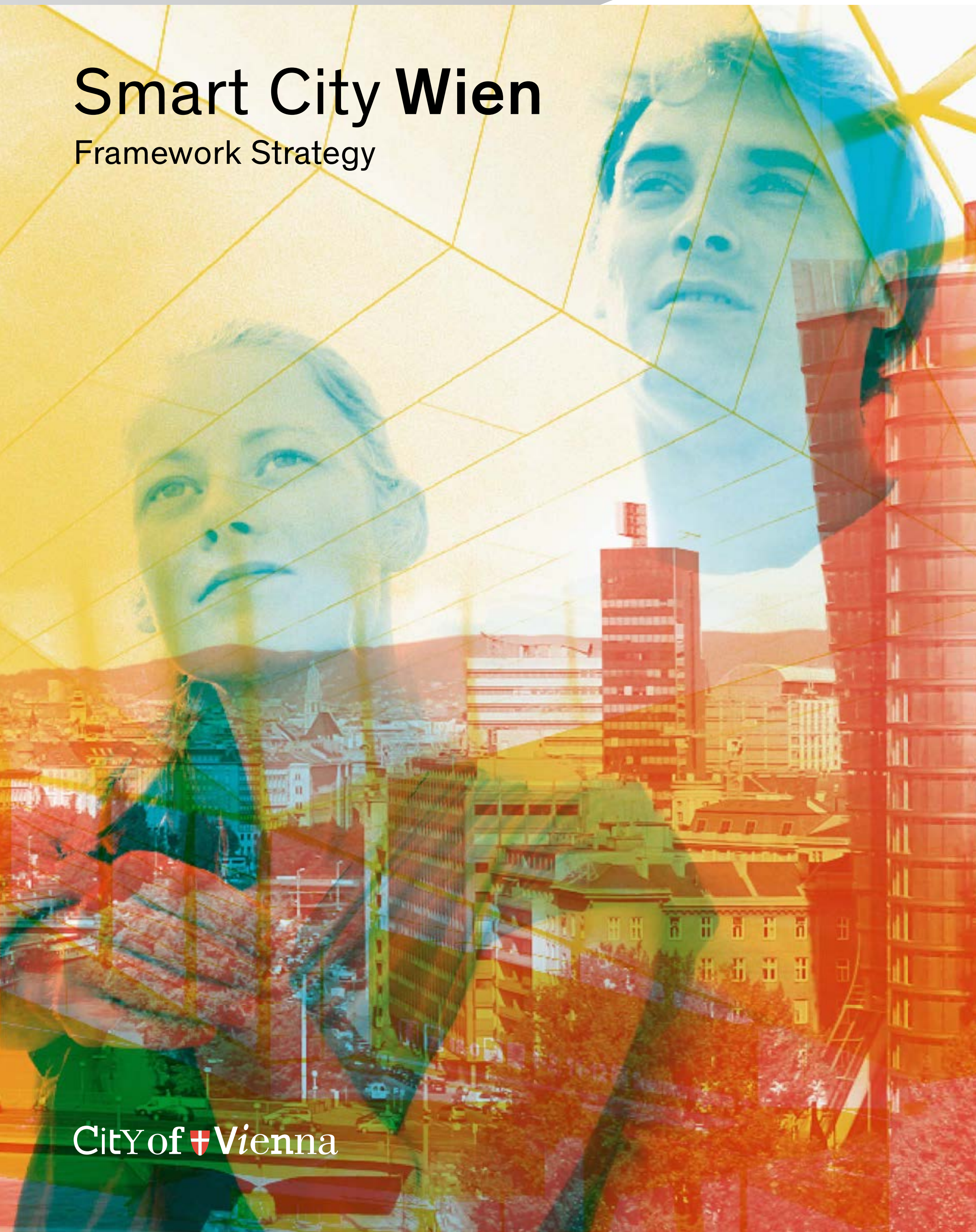
Citation: Kostakos V and Ferreira D (2015) The rise of ubiquitous instrumentation. Front. ICT ***2****:3. doi: 10.3389/fict.2015.00003*

This article was submitted to Human-Media Interaction, a section of the journal Frontiers in ICT.

SMART CITY WIEN

Smart City **Wien**

Framework Strategy



City of Vienna

Smart City **Wien**

Framework Strategy


Our city has been smart for several generations. Far-sighted, intelligent solutions for its daily life have made Vienna the city with the highest quality of living worldwide. This is clearly borne out by its drinking water supply or social housing construction activities. However, to maintain this high quality of living against the background of restricting conditions, it is necessary to strive for constant self-analysis and the development of new and innovative solutions – in brief: to reinvent oneself continuously, especially as climate change and increasingly scarce resources call for novel global approaches.

With the Smart City Wien framework strategy, Vienna is charting its course towards becoming a “smart city”. This is a course that differs from the strategies of other cities in one key respect: Vienna will not let anybody down. For Vienna, the integration of the social component into all areas is an essential element of its framework strategy. Climate-related and ecological objectives and the improvement of the everyday realities of its citizens are assigned the same importance in Vienna. Cities are smart if all people living in them have access to the same degree of participation.

The Smart City Wien framework strategy constitutes a milestone in the future development of the Austrian capital – a strategy designed to ensure that all Viennese will continue living in the world’s most liveable city even in the coming decades.

Dr. Michael Häupl
Mayor

Mag.ª Maria Vassilakou
Executive City Counsillor for Urban Planning, Traffic & Transport, Climate Protection, Energy and Public Participation

Photo Häupl: Michelle Pauty/PID, Photo Vassilakou: Lukas Beck

Foreword
The Smart City Wien Initiative

Vienna is a fantastic place to live and work in. The city is growing, and so are its opportunities. This growth is based on several strong factors, beginning with the city's company structure and educational sector and including an intact environment and ample green spaces. On the international scene, Vienna moreover scores with its public transport network, extensive social housing activities and social services that are available and affordable for everyone. All these things are to be further developed, both in quality and to meet the needs of a growing city.

However, we do not want to attain these objectives by further raising the consumption of the resources needed in the future as well. This concerns first of all fossil fuels, which our current lifestyles and economies still manifestly rely on. They are not infinitely available, entail dependencies and contribute significantly to climate change as well as to its immense consequential costs, which we all will have to bear.

Our future will be designed in the cities. Traditionally, cities have been places of major changes and social innovations; they are home to the majority of the world's population and offer great opportunities for a novel way to deal with resources. A smart city is a city that faces the challenges in the wake of decreasing resource consumption combined with rising demands. However, a smart city will also strive for a high degree of social inclusion. In our opinion, a smart city needs to opt for resource preservation while ensuring high quality of living combined with innovation in all fields.

In Europe as well as worldwide, Vienna is considered a forerunner smart city. For many years, we have done many things right – in transport, housing, urban development, environmental protection, supply and waste management services. Across the world, Vienna is respected for its success in the field of social participation and its high standard of services for the public.

The big Smart City Wien Initiative was launched in 2011 under the aegis of Mayor Michael Häupl. Based on a broad stakeholder process and many approaches derived from different areas of action of the Vienna City Administration, the development of the present strategy was begun in 2013. All Executive Policy Groups as well as numerous experts have contributed to this document. At the same time, there is an intense exchange of experience with other European cities that likewise aspire to meet smart city goals.

Thus Vienna's Smart City strategy is characterised by both an internal effect to render existing plans even more ambitious and to inspire new ideas. At the same time, its external effect is to create an international frame of reference for what is happening here and to generate publicity for Vienna's aims.

The present document is a framework strategy: its time horizon extends to 2050, since the necessary and often fundamental changes in the fields of energy, mobility or construction cannot happen overnight. The thematic arc stretches from the future of Vienna as a hub of research and business to the preservation of all-important social achievements. Concrete methods of application must still be developed in many areas – but the direction is clear: Vienna wants to reduce its resource consumption notably. At the same time, the city intends to continue offering all citizens maximum quality of living, safety and security. These challenges can be met if we tackle change actively und make Vienna a place that fosters innovation even more than it does today.

Table of contents

Vienna's strength stems from far-reaching infrastructure-related decisions

These developments have marked Vienna and permit us today to move into the future from the very high level of the present. A few outstanding examples bear ample proof of this:

Social housing construction – both municipal and non-profit – has produced more than 400,000 high-quality dwellings distributed all over the city. This contributes significantly to a good social mix and affordable housing costs for all.

Public transport is highly developed and makes it possible to quickly reach almost all parts of the city. Inexpensive fares and excellent reliability as well as quality ensure high acceptance levels.

Vienna's water is of supreme quality for a metropolis. The outstanding security of supply and high efficiency of drinking water distribution are outcomes of hard work and sustainable investments made by the City of Vienna for over one century. The amount of Euro 30 million is invested annually in Vienna's water pipe network – a boon for all Viennese.

The waste disposal structures of the city, including wastewater purification, waste treatment, waste separation or cogeneration, i.e. combined waste incineration and heat generation, are considered models of good practice by many other cities.

Vienna is a dense city – and at the same time manages to keep its share of green spaces at 50%. Large-scale structural decisions such as the preservation of the Vienna Woods or the creation of Danube Island have combined environmental quality with attractive leisure options and – in the case of the training of the Danube – have protected the city against flooding.

This shows clearly that a strong municipal policy and far-reaching investment decisions are certainly worthwhile.

Introduction: **Smart City Wien** – Guidelines for the future

Cities have always been the engines of processes and changes in society. Likewise, they have always been at the origin of social innovation. Cities provide space and inspiration for new and different developments for individuals of diverse backgrounds and educational attainments and with varying needs. This is a task that requires constant redefinition.

To safeguard a socially and environmentally acceptable development for the future and to protect the national and international competitiveness of the Austrian capital, the long-term Smart City Wien framework strategy lays down ambitious objectives. The strategy focuses on the intention of preserving and further evolving the city as a liveable, socially inclusive and dynamic space for future generations. The Viennese smart city approach is based on sparing resource use in order to massively reduce CO_2 emissions[1] and dependencies in connection with scarce and finite resources. At the same time, Smart City Wien means to uphold and further increase Vienna's high quality of living and social participation. Ultimately, Smart City Wien stands for change based on innovation, active organisation and, where necessary, the development of new forms of public and private service delivery.

The present Smart City Wien framework strategy is directed at all target groups of the city: Vienna's citizens, enterprises, non-profit institutions and, last but not least, the public sector itself. The strategy and the actions deriving from it are moreover to deploy a strong external effect: Vienna positions itself as a responsible and impulse-generating metropolis in Europe and in the world at large.

1 In the following, the terms "CO_2" and "CO_2 equivalents" are used synonymously for reasons of simplicity. The CO_2 equivalent value describes the greenhouse gas potential of a gas as compared to CO_2 over an observation period of (mostly) 100 years. In 2009, 94% of all greenhouse gas emissions (expressed in CO_2 equivalents) in Vienna were due to CO_2; only 6% resulted from methane, nitrous oxide (laughing gas) and fluorinated (F) gases.

Vienna has chosen the right track and pursues this approach systematically.

Vienna is an extremely liveable city. In the opinion of many, no city boasts higher quality of living. Corresponding international studies[2] underpin this with evidence taken from all areas of life – from infrastructure or green spaces to educational facilities, from services of general interest to good neighbourly relations, gender equality, safety and security in the city.

This superlative quality of living is largely due to the fact that the city has always known how to change and reinvent itself. Examples of this are the big Gründerzeit construction push with its infrastructure facilities and buildings that characterise Vienna to this day, the municipal housing projects of the interwar period, the gentle urban renewal approach since the 1970s and the renewed flourishing of the city after the dismantling of the Iron Curtain in 1989. Vienna has taken far-reaching "smart" decisions already several times in the past (page 10).

Today the federal capital of Austria is about to take the next big step towards change. This concerns the city's contribution to the most far-reaching (and anthropogenic) challenge we are faced with in our time: climate protection is a task for everyone. To attain it, we must largely abandon fossil fuels by means of a long-term, gradual process and substitute them by other, more sustainable forms of energy, conversion technologies and services.

Smart City Wien comprises first and foremost the aim of **resource preservation**. Development and modification processes in the sectors of energy, mobility, infrastructure and building management are to dramatically reduce CO_2 emissions by 2050. For this purpose, it is essential to make much more efficient use of the required input energy. It is the first big task of the present Smart City framework strategy of the City of Vienna to highlight ways and means in which Vienna can contribute to the attainment of the major European climate and energy objectives (for 2020, 2030 and 2050). Potentials towards these goals still exist to a high degree, although Vienna, due to ambitious plans and decades of consistent action, has already achieved much in the fields of climate and environmental protection. However, the responsibility for future generations above all motivates us to continue our pursuit of ambitious goals in the context of the EU energy and climate objectives for the period until 2050.

2 Inter alia: Mercer "Worldwide Quality of Living Survey", Smart City ranking by Boyd Cohen, UN-HABITAT – United Nations Human Settlements Programme "State of the World's Cities 2012/2013".

European energy and climate objectives

20-20-20 targets of the European Council for 2020 (2007)

The European Union has set itself ambitious goals in the fields of energy and climate policy: the “20-20-20 targets” obligate EU Member States to reduce greenhouse gas emissions by at least 20% from 1990 levels by 2020, to improve energy efficiency by 20% and to attain a 20%-share of EU energy consumption produced from renewable resources.

Proposed 2030 goals of the EU

In January 2014, the EU Commission submitted its proposal for the energy and climate goals for 2030, which was seconded by the European Parliament. Accordingly, the EU-wide CO_2 emission volume is to be reduced by 40% from 1990 levels by 2030. It is likely that national goals will be proposed to define the CO_2 emissions reduction for each EU Member State. At the same time, the share of renewable energy is to be increased to 27% (proposal by the EP: 30%) across the EU. For energy efficiency, which is to be improved by 20% by 2020, the new goal of 30% by 2030 has been proposed.

2050 objective of the EU heads of state and government

In late October 2009, the EU heads of state and government – together with all other industrial nations – unanimously supported an EU objective to reduce greenhouse gas emissions in the EU by 80 to 95% as compared to 1990. According to scientific findings, this reduction is essential to limit global warming to less than 2 degrees centigrade (»two-degree objective«), as it is assumed that any value above this threshold will result in uncontrollable consequences of climate change.

If correctly implemented, any success relating to climate protection in the fields of transport, housing and production has direct effects: the cost of energy is reduced, while energy security is improved. The results are more green spaces, cleaner air, more liveable neighbourhoods, shorter distances and easier accessibility plus a more varied and affordable range of public spaces and public services. However, this also means greater co-determination options for Vienna's population. The further increase of the **quality of living** is a second objective that is given as much importance as sparing resource use: in environmental protection or healthcare, Vienna can build on already very high standards. The relevant political decisions were and are significantly derived from the principle of social inclusion. The creation of affordable and attractive housing, the provision of low-cost and resource-conserving mobility and the financing of services of general interest are only a few examples of the implementation of this principle in reality.

But the City of Vienna is also taking very intense efforts to further equal opportunities. The city belongs to women as much as to men, and all social groups are thus called upon to participate in shaping our society. For Vienna, the three central strategies to attain this goal are the advancement of women and their rights, gender mainstreaming and gender budgeting. A key precondition for safeguarding the same opportunities of self-realisation for both women and men lies in their equal participation in social and political decision-making processes. Another special feature of Vienna stems from the fact that gender mainstreaming is made part and parcel of the city's development as a cross-cutting principle.

Many changes can be implemented quickly, while other processes may require decades: what are our means of transport, how do we communicate, what heating systems do we need? We can certainly only imagine part of the possibilities that will be available to us in coming decades. However, we do know that Vienna is able to develop technically, organisationally and socially exemplary solutions. In view of the huge challenges, it is better to advance the necessary changes proactively and benefit from them both economically and as a society. To reduce resource use as planned while maintaining or even improving the quality of living, **innovation** is the third major approach pursued by Vienna's Smart City framework strategy. As a smart city, Vienna boasts a dynamic economy, boosts information and communication technologies, assigns very high priority to education and, last but not least, defines itself as a first-rate research hub. It is a central concern of Vienna to transform challenges into opportunities – for Vienna's enterprises, for its residents' career prospects and for good neighbourly relations between all citizens.

A "smart city" also means social inclusion: Vienna takes account of the needs of all residents!

Any city that utilises the smart city concept as a metaphor for processes of change defines it in its very own way. While some cities focus above all on the implementation of technological possibilities, others aim mainly at the reduction of emission levels. Conversely, Vienna continues its chosen and successful road by following several objectives concomitantly and favours social inclusion even more than in the past. Vienna can only be smart if …

… the needs of many different population groups can be met: Smart City Wien means recognising this diversity. The city and its institutions will continue to make sure that processes of change will be socially balanced, that disadvantages will be compensated as far as possible, and that the high level of social security will be maintained.

… high quality of living is possible also for persons with lower incomes: Smart City Wien means a superlative standard of public services, affordable housing and public transport, spacious and publicly accessible green and leisure spaces, a highly developed healthcare system and many other things. The further development of Vienna equals development for all and is to be perceived as such by all citizens.

… codetermination and participation shape the development of the city: Smart City Wien means creating space for locally fine-tuned solutions and self-initiatives as well as the possibility for citizens of having a say in the development of their city.

… innovations and progress have a social component: finally, Smart City Wien means fostering what is new. While this often involves technological novelties, social innovations too, are increasingly coming to the fore. Contrary to purely technological developments, these are more strongly inspired by the needs of citizens, rest on a broader basis and take particular account of the interaction of different individuals and organisations.

Our ability to master the future can only be safeguarded if the necessary processes of change enjoy wide support. All cities today are facing major challenges. For Vienna, the crucial point lies in the fact that these changes will entail further improvement of the living conditions of all parts of the urban population. The leitmotif of this is Vienna as a social city that supports all persons in their effort to live a good life.

As a smart city, Vienna must also be resilient and hence robust, flexible, adaptive and able to react quickly and in keeping with the challenges when faced with internal and external influences. In this, resilience is strongly dependent on the availability of room to manoeuvre, on the possibilities for self-organisation or for re-organisation of economic and social systems, on social coherence, on the competencies of residents and on a flexible and innovative administration.

The three major sets of goals – resource preservation, quality of living and innovation – are closely interlinked. Vienna wants to maintain its excellent position in the international competition of cities, although it is not enough to hold a top position regarding only one of these sets. Vienna maintains a close dialogue with leading cities in Europe and worldwide on promising approaches. The Viennese approach will be very special!

It is thus the key goal for 2050 of Smart City Wien to offer optimum **quality of living**, combined with highest possible **resource preservation**, for all citizens. This can be achieved through comprehensive **innovations**.

The present framework strategy describes the key goals and principal approaches chosen to attain them. It represents guidelines for the numerous important specialised strategies of the city that define concrete multiyear plans for such areas as urban planning, climate protection, the future of energy supply or Vienna as an innovation hub. In this, the framework strategy poses a twofold challenge: first of all, how can the goals be gradually rendered more and more ambitious despite the demanding practical and financial frame conditions? And, secondly: how can policy and change processes be designed in an even more cross-cutting, multi-sectoral manner?

A cross-cutting concept also underpins the following Smart City Wien 2050 (page 19) vision embraced by the city.

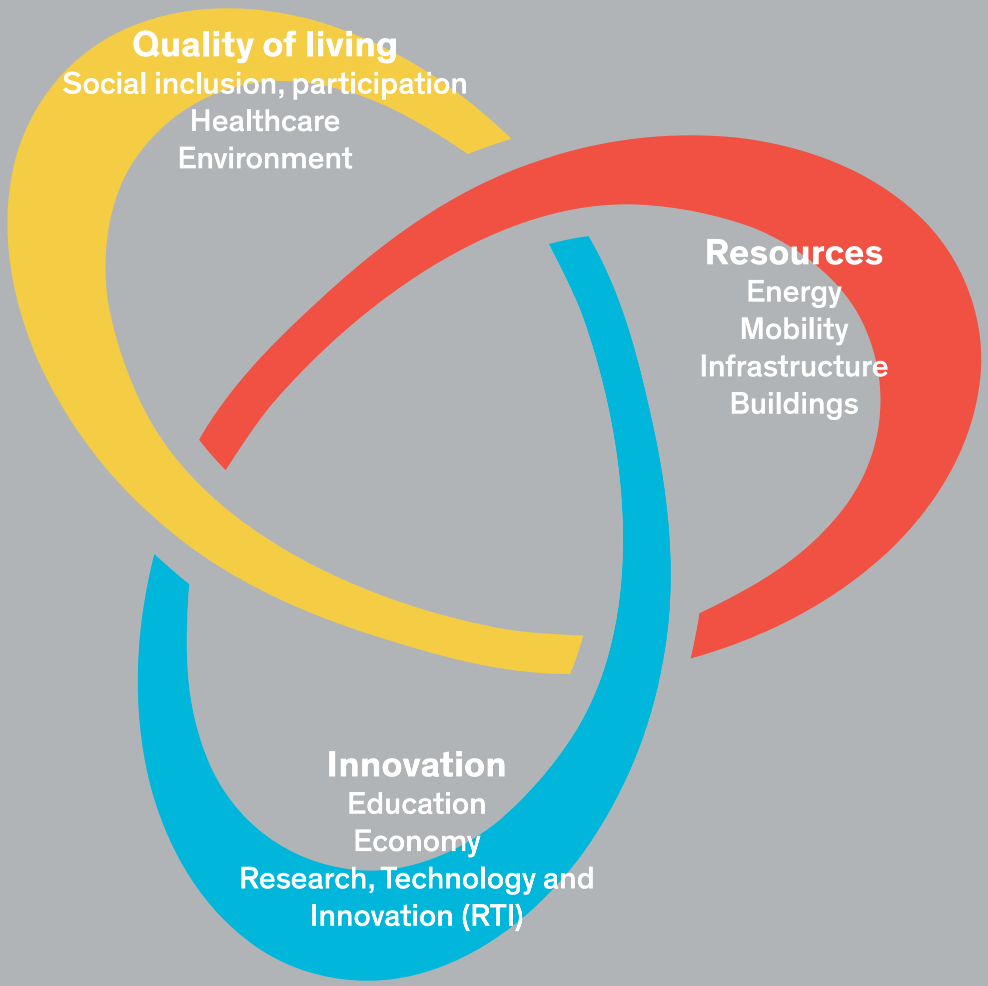


Fig. 1 “The Smart City Wien principle”

"In a city of short distances, I can travel on foot or by bicycle. Moreover, this helps me to save time in my travels through the city and hence to combine my work and my private life more effectively."

"The preservation and enlargement of green spaces contribute to absorbing CO_2 and reducing particulate matter emission. Due to manifold, attractive leisure options and natural recreational areas, I can relax without leaving the city."

"Thanks to special apps, I can move through the city without barriers even with a fully loaded pram."

2 The Smart City Wien 2050 vision

In 2050, Vienna is a vibrant metropolis and one of Europe's most attractive cities. This position is based on strategically planned, long-term measures of the city, which over the first half of the 21st century have led to a noticeable improvement in all fields of life: quality of living, sustainability, prosperity as well as quality and quantity of educational options and workplaces. Together with other leading cities of Europe, Vienna generates impulses and impacts European policy.

Vienna is a liveable city for children, young people, women and men, elderly persons, families, entrepreneurs, artists, researchers, persons with special needs – in short: a city that is open to all, no matter how long they have been living here.

Vienna is recognised worldwide for the deeply entrenched yet uncomplicated way in which it accords ample **possibilities of participation and codetermination** to all parts of the population. Citizens take active part in developing their city. There are many ways of participating: everyone has the possibility of voicing, discussing and implementing their own ideas and opinions regarding the city.

In a unique manner, Vienna offers **affordable quality of living** as well as spacious, easily accessible leisure and green spaces "around the corner" and allows for individual recreational activities. Social justice is a key principle and lays the basis for comprehensive services of general interest in Vienna. The urban structure and municipal services of Vienna generate a feeling of safety and security in the city.

Art and culture are crucial drivers of society and constitute an integral element of Smart City Wien 2050.

At any moment, the Viennese population may draw on numerous combined offerings and joint riding options for time-tried and **innovative means of transport** to meet their mobility requirements; these means of transport also open up new economic opportunities and leave ample leeway for creative development. Movement in the city is characterised by resource preservation and respect for public space, which has been gradually recovered by the residents. As a result, Vienna's inhabitants experience quality of living based on low noise levels and clean air in the city – day after day.

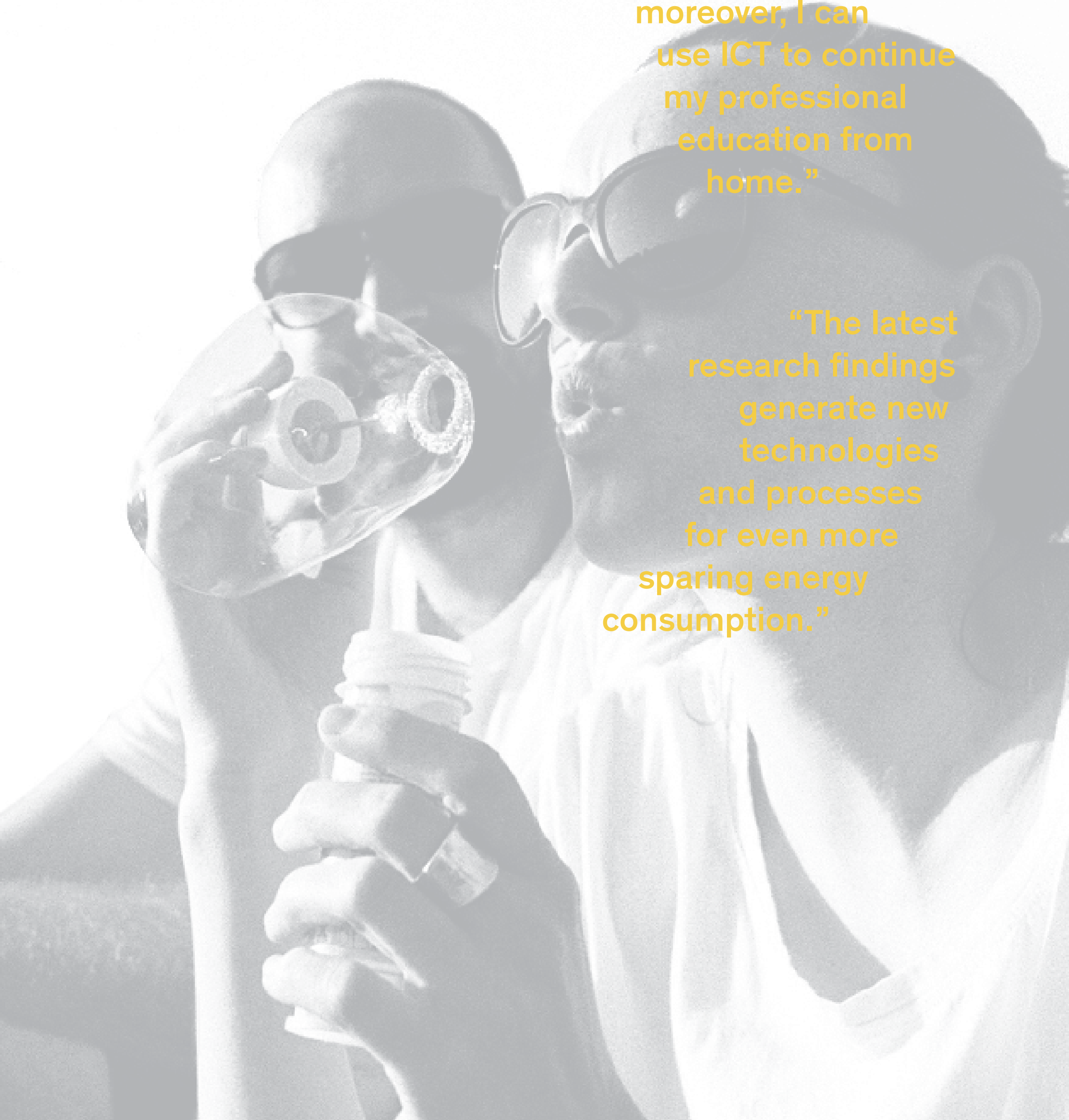
The conscious and sparing use of resources coupled with innovative solutions allows for maximum security of supply. For this reason, the further development of district heating, which in Vienna is largely produced from waste heat, is given particular importance. Renewable sources – e.g. from geothermal energy and low-temperature waste heat – must be developed to step up district heating. In addition, Vienna's energy requirements are met to a large degree by renewable sources. The use of state-of-the-art technologies and high environmental and energy awareness underpin the actions of private households, the municipal administration and enterprises in Vienna, with decentralised renewable energy supply in urban quarters suitable for this purpose playing an important role in this context.

Efficient, intelligent networks and useful information systems facilitate life in the city all around.

Vienna combines history and tradition with a modern, redesigned urban environment. The "second Gründerzeit" period in the 21st century provides impulses for **striking new developments in the fields of architecture,** design and sustainability. At the same time, the Viennese are proud of their multifaceted neighbourhoods scattered all over the city. These urban subcentres make for short distances regarding supply and services. People enjoy the wide range of available offerings and appreciate the excellent range of goods and services as well as the manifold possibilities to shape their own city, which actively supports their everyday life in multiple ways, thereby safeguarding equal opportunities for all.

“Information and communication technology simplifies my life and enables me to handle transactions anywhere and anytime at my leisure; moreover, I can use ICT to continue my professional education from home.”

“The latest research findings generate new technologies and processes for even more sparing energy consumption.”

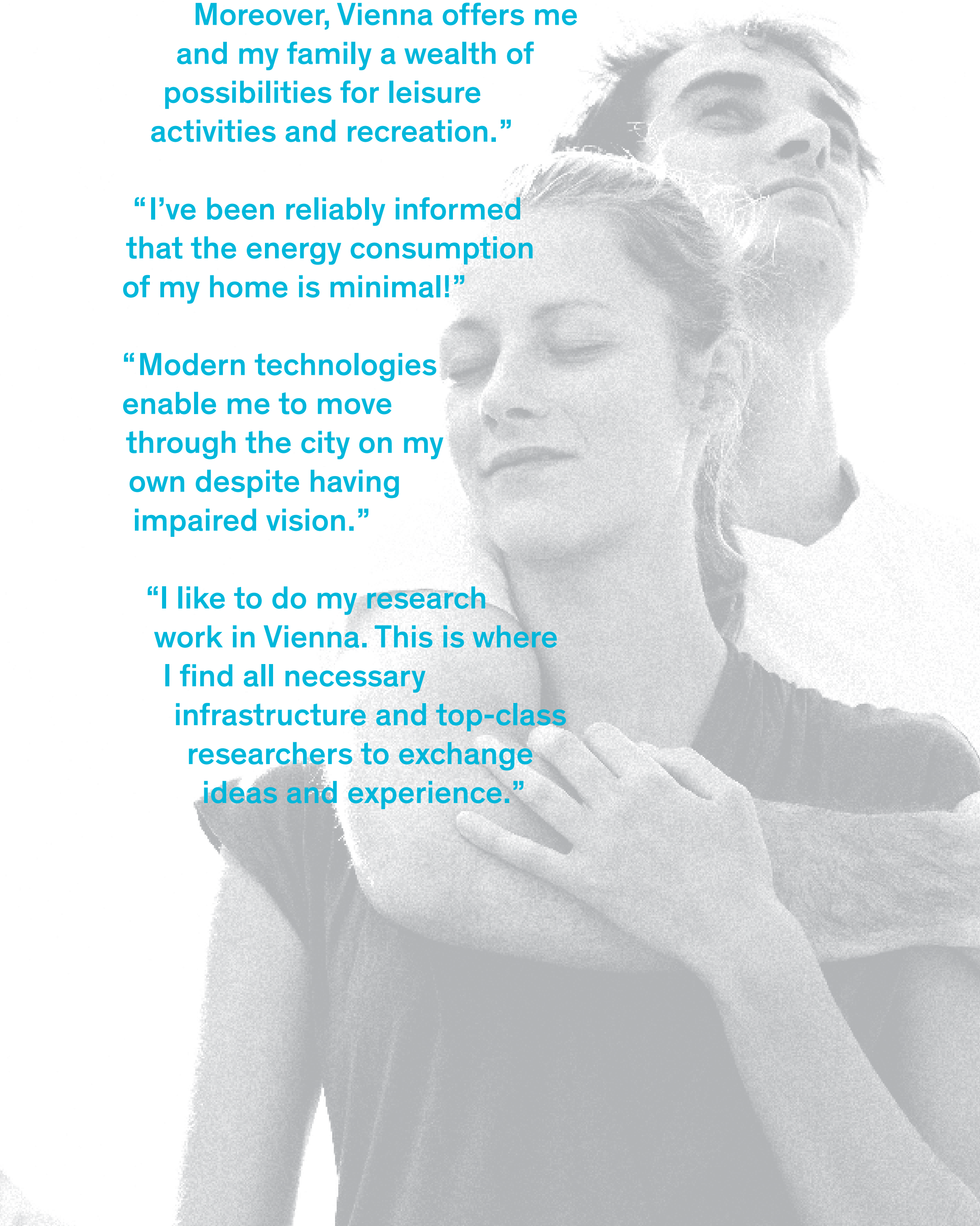


“The city is the centre of our life. I like to live and work in Vienna. Moreover, Vienna offers me and my family a wealth of possibilities for leisure activities and recreation.”

“I’ve been reliably informed that the energy consumption of my home is minimal!”

“Modern technologies enable me to move through the city on my own despite having impaired vision.”

“I like to do my research work in Vienna. This is where I find all necessary infrastructure and top-class researchers to exchange ideas and experience.”



For students, teachers and researchers, Vienna is a **Central European capital of research and hence** an attractive location and hub. Vienna maintains a lively exchange of knowledge and thought with other important international centres of research. All these developments are based on outstanding educational possibilities.

Vienna's prosperity stems from a **strong economy** that is steeped in the efficiency of the producers of goods and services and their workers. Entrepreneurs develop creative ideas and implement them successfully. Partly also due to the use of cutting-edge information and communication technologies, this opens up a multifaceted work environment that reacts to numerous interests and skills of both women and men and generates a sufficient number of workplaces designed to enable workers to combine career and family. Vienna's economy is thriving and generates a wealth of innovations. Viennese know-how, products and services in such areas as energy, mobility, sustainability, healthcare and many other segments are exported all over the world.

Vienna's population lives in a **smoothly functioning metropolitan region.** This is possible because the environs of the Austrian capital are actively involved and because new forms of co-operation, e.g. in mobility, housing, spatial development and energy supply, are made use of.

Vienna embodies quality of living at the very highest level.

This image of Vienna in 2050 can indeed be attained. The basis for this must be laid today and through decisions that are taken in a spirit of responsibility, day after day, by everyone in Vienna.

3 Vienna: status quo

Vienna enjoys a very good starting point for becoming a smart city. The Austrian capital differs from most other metropolises through its good performance in so many areas: housing, public transport and other infrastructure services (e.g. waste separation, Spring Water Mains), education and universities as well as vast urban green spaces. All this contributes towards high quality of living. In 2011, Vienna took the top place in the first international smart city index[3]. In 2012, Vienna kept a very good place (4th), followed by third place in 2013. Other studies[4] document Vienna's strong and enviable position on the international scene, as the Austrian capital continues to keep its worldwide top rank.

At the end of 2011, it was possible to avoid the production of 3.7 million tonnes of CO_2 since 1990, partly due to the Climate Protection Programme of the City of Vienna (KliP). Internationally, this means a very good position for Vienna. Between 1990 and 2010, the CO_2 emission level[5] in Vienna decreased by 21% per capita and by 10% in absolute figures. This result was inter alia attained by upgrading the district heating network, improving thermal building standards (both in building rehabilitation and new structures; Vienna is the city with the highest number of passive houses) and doubling the share of energy from renewables from 5% in 1993 to 11% in 2011. Above all, however, greenhouse gas reduction was supported by the strong growth of public transport (from 29% in 1993 to 39% in 2013) and bicycle traffic (from 3% in 1993 to 6% in 2011).

Vienna may rely on a dense network of scientific institutions, centres of excellence and university facilities, which include ten state-owned universities with very different profiles, several universities of applied sciences plus a number of private universities and, in particular, many non-university research institutions. Vienna's universities alone generate a value added of Euro 2.3 billion annually.[6] Close to 35% of Austria's R&D spending is invested in Vienna.[7] This trend, which has progressed very dynamically especially over the past decade, has resulted in a specific status assigned to science, research and innovation for urban development.

Vienna is characterised by a strong administration and high social responsibility. Many areas of human life are covered by enterprises and companies of the City of Vienna, e.g. housing (wiener_wohnen - Vienna

3 Boyd Cohen, Global Ranking 2011.
4 Mercer Survey, UN-Habitat, etc.
5 Evaluierung der Umsetzung des Klimaschutzprogramms (KliP II) der Stadt Wien, 2011.
6 Dritter Bericht des Beauftragten der Stadt Wien für Universitäten und Forschung.
7 Data: Statistics Austria 2011

Fund for Housing Construction and Urban Renewal), Vienna Water, the hospital sector and Vienna Public Utilities, which provides mobility and energy services through Wiener Linien, Wien Energie and Wiener Netze. It is definitely the political strategy of Vienna to keep the reins on the infrastructure required to deliver basic services.

The city regularly evaluates its quality of living,[8] a field in which the Austrian capital holds a special position at a European scale. This is not only expressed in annual international rankings but also reflects the opinion of Vienna's population. Surveys show that the manifold urban leisure and cultural activities, the environmental quality – which is high for a metropolis this size –, the ample social facilities and services for the population and the wide range of publicly subsidised housing options are particular assets of Vienna and contribute to its good image overall. A comparison with German and Swiss cities shows moreover that Vienna scores best in related issues concerning e.g. employment and housing opportunities, provision of social and nursing care services, public transport or environmental quality. Although Vienna is compact and Austria's smallest federal province by far, more than 50% of its territory is covered by green spaces. Vienna has launched environmental initiatives at a very early date and continues to pursue this approach. This is reflected in the field of ecology, which aims for close co-operation between science, public administration and business and presents tight collaborations between municipal administration and universities.

But Vienna wants to further fine-tune this balanced development and may rely on a very strong base towards this goal.

8 Studies "Leben in Wien" (1994 and 1995), "Leben und Lebensqualität in Wien" (2003), "Wiener Lebensqualitätsstudien" (2008 and 2013).

4

What challenges does Vienna face today?

More than ever before in human history, the 21st century is and will be the “century of cities«. The traditional role of big cities as drivers of scientific, technological, cultural and social innovation will continue to grow. Strong metropolises serve a decisive formative function regarding the main issues of tomorrow. Since time immemorial, cities have been the hubs of innovation, and this trend will increase further. Immigration and growth will continue in huge dimensions in the mega-cities of Asia, Africa and Latin America. As the example of Vienna shows, many European cities, too, present high dynamism that must not merely be mastered but rather should be viewed as an opportunity.

In this context, the Smart City framework strategy should thus be understood as follows: what will Vienna do to tap the opportunities of change and dynamic growth? How can we position ourselves as a venue of innovation and new solutions?

The resource issue is even more strongly contingent on cities, as energy consumption, CO_2 emissions, mobility patterns and hence quality of living of their citizens are at the centre of attention. The advantages related to short distances and spatial compactness are juxtaposed with challenges that result from rising consumption volumes as much as from the difficulty of contributing to changes in established technologies and lifestyles of various population groups.

Vienna boasts a long-standing tradition of resource conservation and protection. Instances of this are the impressively high share of public transport, the ramified district heating network with its cogeneration and waste incineration installations or numerous examples of resource-conserving production in Vienna’s industrial plants.

However, Vienna also must cope with specific challenges that stem on the one hand from urban growth and on the other hand from necessary processes of change. Examples in this context comprise the further restructuring of energy systems, the organisation and financing of building rehabilitation including thermal rehabilitation as well as changing demands made on the mobility system, which is marked by a steep increase in the shares for walking and cycling. Better fine-tuning of processes between city and surrounding region is another challenge for the future.

Although the excellent status quo of Vienna actually renders further improvements somewhat more difficult, the Austrian capital is firmly committed to the EU climate change objectives for 2030 and 2050 and wants to make the best possible contribution towards their attainment. However, it is equally clear that the goals formulated below cannot be fully achieved without corresponding frame conditions laid down by third parties (Federal Republic, EU). This includes the safeguarding of funds for specific climate protection measures in Vienna. Vis-à-vis the Federal

The special thing about Vienna's Smart City Framework Strategy lies in the fact that the aspects of **social inclusion** are considered essentialfor all three dimensions.

Fig. 2
The three dimensions of Smart City Wien

Republic of Austria and the European Union, Vienna will therefore advocate frame conditions that duly support the attainment of these goals.

The interaction of the three dimensions of Smart City Wien shown below – i.e. resources, quality of living and innovation – and the three interlaced superordinate goals allow for balancing the various needs and approaches and avoiding overly costly or biased and hence excessively risky strategies. Chapters 6 to 8 will specify how these objectives of **resource preservation, quality of living** and **innovation** are formulated in greater detail.

Quality of living

Three impulse generators formulate Vienna's specific approach to becoming a smart city. Vienna strives for optimum quality of living combined with the attainment of the necessary resource-related objectives. In this way, Vienna builds on existing strengths in the areas of social inclusion, healthcare and environment.

Resources

To be able to attain the ambitious goals of Smart City Wien as a resource-conserving forerunner city, politics and administration are committed to setting important steps in the core areas of energy, mobility, buildings and infrastructure. This comprises issues pertaining to energy systems, energy generation, pre-existing and new city quarters, future means of transport and the use of pioneering information and communication technology. The core areas form the main focus of the strategy, since they are primarily decisive for the question of resource preservation but also have a strong impact on quality of living and innovation.

Innovation

Three other impulse generators are decisive for the field of innovation, which supports and paves the road towards the Smart City Wien targets and is characterised by the intelligent and systematic use of cutting-edge technologies and social innovation. Education prepares the ground, and research, technology and innovation (RTI) produce novel, smart technological and social solutions. Finally, a strong economy implements these solutions and provides employment.

5 Overview of objectives and policy areas

Smart City Wien combines the three essential and interlinked basic elements of resources (resource preservation), quality of living and innovation. In this way, it builds on typical strengths of Vienna and includes externally imposed binding goals.

The definition of Smart City Wien:

Smart City Wien defines the development of a city that assigns priority to, and interlinks, the issues of energy, mobility, buildings and infrastructure. In this, the following premises apply:

- **radical resource preservation**
- **development and productive use of innovations/new technologies**
- **high and socially balanced quality of living**

This is to safeguard the city's ability to withstand future challenges in a comprehensive fashion. The elementary trait of Smart City Wien lies in the holistic approach pursued, which comprises novel mechanisms of action and co-ordination in politics and administration as well as a wider leeway of action assigned to citizens.

These objectives are long-term, allow for flexibility to do justice to continuous social change and should be understood as inextricably linked to the existing targets set by different specialised strategies of the City of Vienna (Fig. 3). The framework strategy does not substitute the targets of these specialised strategies but is to act like a magnet, i.e. as a superordinate and thematic framework that is in its turn encapsulated in existing plans, strategies, catalogues of targets and works.

The existing plans, strategies, etc. mostly follow a medium-term horizon, focus on one sector only and often comprise extensive and detailed packages of measures. The Smart City Wien framework strategy is more comprehensive (but not exhaustive), pursues a long-term horizon (2050) and does not offer detailed packages of measures. However, concrete sub-projects with a shorter timeframe will definitely be formulated and implemented – and also serve the purpose of clarifying the very nature of what a "smart city Vienna" might be like. In this way, the Smart City Wien framework strategy lays down an aid to orientation for the next generation of specialised strategies in such areas as climate protection, innovation, urban planning, mobility, etc.

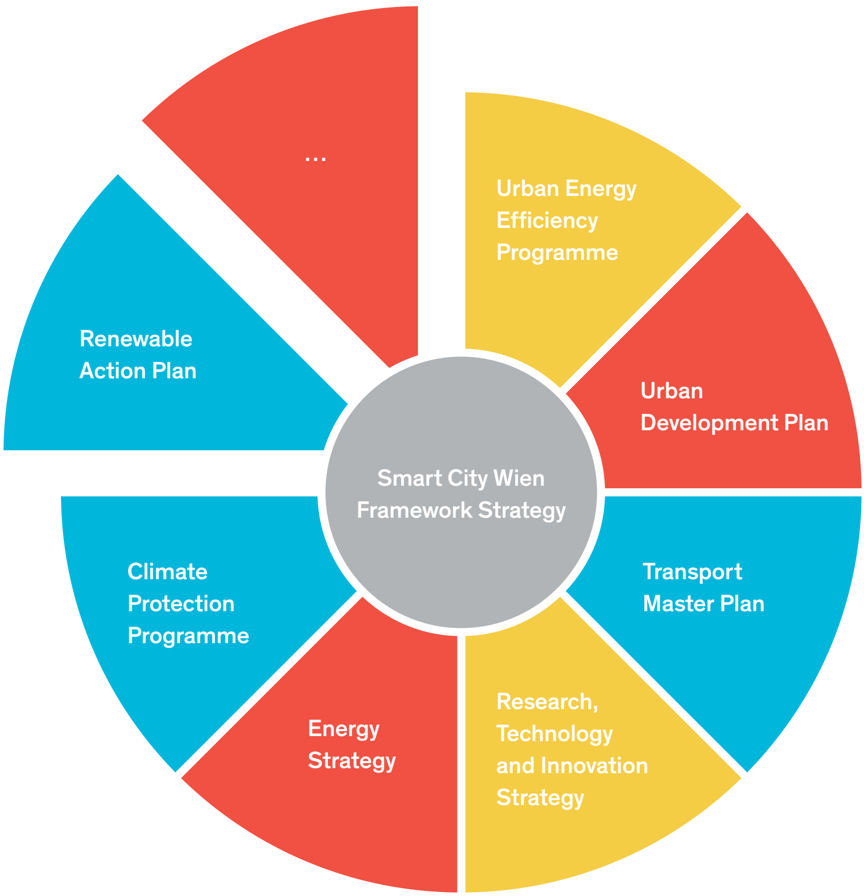


Fig. 3 Interaction of Smart City Wien framework strategy with existing and future strategies

Resources

Objective: Per-capita greenhouse gas emissions in Vienna drop by at least 35% by 2030 and by 80% by 2050 (compared to 1990).

Resources

Objectives Energy:

- Increase of energy efficiency and decrease of final energy consumption per capita in Vienna by 40% by 2050 (compared to 2005).
- At the same time, the per-capita primary energy input should drop from 3,000 watt to 2,000 watt.
- In 2030, over 20%, and in 2050, 50% of Vienna's gross energy consumption will originate from renewable sources.

Objectives Mobility:

- Strengthening of CO_2-free modes (walking and cycling), maintenance of high share of public transport and decrease of motorised individual traffic (MIT) in the city to 20% by 2025, to 15% by 2030, and to markedly less than 15% by 2050.
- By 2030, the largest possible share of MIT is to be shifted to public transport and non-motorised types of traffic or should make use of new propulsion technologies (e.g. electric-powered vehicles).
- By 2050, all motorised individual traffic within the municipal boundaries is to make do without conventional propulsion technologies.
- By 2030, commercial traffic originating and terminating within the municipal boundaries is to be largely CO_2-free.
- Reduction of energy consumption by passenger traffic across municipal boundaries by 10% in 2030.

Objectives Buildings:

- Cost-optimised zero-energy building standards for all new structures, additions and refurbishments from 2018/2020 and further development of heat supply systems towards even better climate protection levels.
- Comprehensive rehabilitation activities entail the reduction of energy consumption of existing buildings for space heating/cooling/water heating by one percent per capita and year.

Objectives Infrastructure:

- Maintenance of the high standards of Vienna's infrastructure facilities.
- In 2020, Vienna is the most progressive European city with respect to open government.
- The next 100 apps in three years.
- Pilot projects with ICT enterprises are to serve as showcases for the city and its economy.
- In three years, Vienna will have a comprehensive WLAN.

Innovation

Objectives Research, Technology and Innovation Strategy (RTI):

- In 2050, Vienna is one of the five biggest European research and innovation hubs.
- By 2030, Vienna attracts additional research units of international corporations.
- In 2030, Vienna is a magnet for international top researchers and students.
- By 2030, the innovation triangle Vienna-Brno-Bratislava is one of the most future-oriented cross-border innovation regions of Europe.

Objectives Economy:

- In 2050, Vienna remains one of the ten European regions with the highest purchasing power based on per-capita GDP.
- Vienna further strengthens its position as the preferred company headquarters city in Central/South-eastern Europe.
- 10,000 persons annually set up an enterprise in Vienna.
- The direct investment flows from and to Vienna have doubled as compared to 2013.
- The share of technology-intensive products in the export volume has increased to 80% by 2050 (as compared to 60% in 2012).

Objectives Education:

- Comprehensive provision of whole-day and integrated schools and further upgrading of high-quality childcare offerings.
- Even after 2020, it will be important for Vienna to make sure that a maximum number of young people will continue their education beyond compulsory schooling and thus attain a higher educational level.
- Even after 2020, it will be important for Vienna to safeguard positiv conditions for acquiring formal school-leaving qualifications through second-chance education and for the recognition of foreign educational attainments by adults.

Innovation

Objective: In 2050, Vienna is an innovation leader due to top-end research, a strong economy and education.

Quality of living

Objective: Vienna maintains its quality of living at the current superlative level and continues to focus on social inclusion in its policy design: as a result, Vienna in 2050 is the city with the highest quality of living and life satisfaction in Europe.

Quality of living

Objectives Social inclusion:

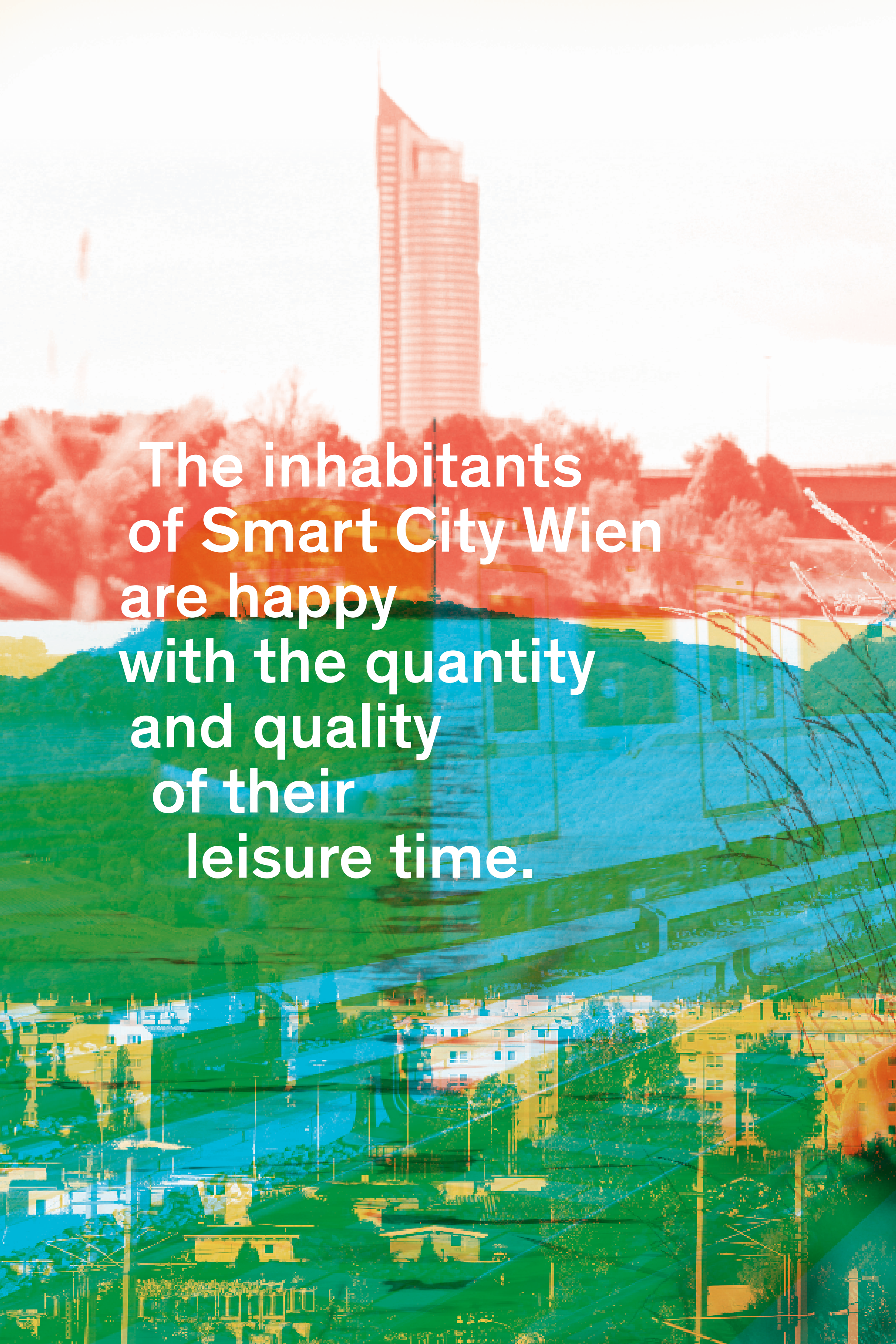
- All people in Vienna enjoy good neighbourly and safe life conditions irrespective of their background, physical and psychological condition, sexual orientation and gender identity. Vienna is a city of diversity that is expressed to the fullest in all areas of life.
- High-quality, affordable housing and an attractive housing environment are made accessible to the largest possible share of the population.
- Active participation at work as well as the performed work must be adequately remunerated and ensure the coverage of all basic needs in life.
- Women are involved in planning, decision-making and implementation processes in keeping with their share in the total population. All persons involved in these processes dispose of gender competence.

Objectives Healthcare:

- Strengthening of health-promoting conditions of life and health literacy of all population groups.
- Safeguarding of medical care at the highest level due to demand-oriented and efficient supply structures and processes (best point of service) for all citizens as well as reduction and shortening of hospital stays.
- To safeguard a strong and socially equitable public healthcare system, the Vienna Hospital Association and its facilities will remain a publicly-owned enterprise. Potentials for greater efficiency must be systematically reviewed and used in all areas.
- "Outpatient over inpatient" is the organisational principle of nursing services – letting persons stay in their own home for as long as possible while offering top-notch nursing quality.
- The inhabitants of "smart Vienna" are happy with the quantity and quality of their leisure time.

Objectives Environment:

- By 2030, the share of green spaces must be kept at over 50%. Especially in a growing city, additional recreational areas must be safeguarded to keep up with the rising population figures.
- In 2020, the savings achieved by municipal waste management have already attained approx. 270,000 tonnes of CO_2 equivalents as a result of further planned measures and improvements.


The inhabitants
of Smart City Wien
are happy
with the quantity
and quality
of their
leisure time.

SMART CITY

RESOURCES

6 Objective: highest possible resource preservation

The combustion of fossil fuels in cities and their surroundings causes approx. 70 to 75% of worldwide CO_2 emissions.[9] In developing and emerging countries as well as in highly developed industrial nations, migration towards the cities continues unabated. Cities thus play a key role for the energy turnaround and must undertake intensified efforts to attain ambitious goals. Therefore Vienna pursues the following big objective:

Objective: reducing per-capita greenhouse gas emissions in Vienna by 80% by 2050 (as compared to 1990) [10]

This objective is derived from the recommendations of the UN Intergovernmental Panel on Climate Change to limit anthropogenic global warming in the long term to two degrees centigrade. This objective is also endorsed by the EU and results in the requirement of reducing CO_2 emissions to a sustainable level of approx. one tonne per capita and year. Until 2050, the CO_2 target of 80 to 95% supported by the heads of state and government sets an underlying framework (see page 13). In addition, Vienna has committed itself to a continuous reduction of greenhouse gas emissions in the context of the Climate Alliance.

The 20-20-20 targets of the EU are the first important step towards this goal. As the next step, the CO_2 emission level within the EU is to be reduced by 40% until 2030. To attain this target, the sector covered by emissions trading (ETS) is to contribute a reduction by 43% while the »non-ETS sector« is to generate a reduction by 30% (referred to 2005 values).

It is an advantage of cities that condensed settlement types tend to trigger lower energy requirements for mobility and space heating. Vis-à-vis many other regions, Vienna already in 1990 held a pioneering role with 4.1 tonnes of CO_2 per capita[11] for the energy balance segment comparable

9 Cf. e.g. Bouton et al.: How to Make a City Great, McKinsey&Company; Burdett and Sudjic: Living in the Endless City.

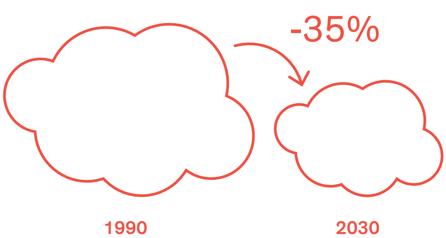
10 This objective, like the energy- and climate-related targets mentioned in the following, can only be attained if Vienna's activities are supported by corresponding conditions established by the Federal Republic of Austria and the EU, which includes the consideration of early actions.

11 AEA 2012: Evaluierung der Umsetzung des Klimaschutzprogramms (KliP II) der Stadt Wien – When calculating Vienna's CO_2 volume, the emissions subject to emissions trading are deducted, as is all fuel consumption that cannot be attributed to Vienna (fuel tourism; fuel purchases attributed to the federal province where the company's HQ is located).

to the »non-ETS sector«. When adopting and implementing the first Climate Protection Programme in 1999, Vienna stepped up its forerunner position and subsequently adopted KLiP II in 2009 as a follow-up programme until 2020. In 2011, Vienna was at 3.1 tonnes of CO_2 per capita, with a primary energy consumption of slightly under 3,000 watt of continuous output per person[12]. As a result, the sustainable, long-term levels for 2050, i.e. not more than one tonne of CO_2 per capita and 2,000 watt per capita[13], seem certainly attainable in Vienna. Vienna finds itself in a markedly better position than comparable ambitious cities in the European or even worldwide context. This can be explained above all by a systematic focus on the expansion of the public transport network, excellent building and rehabilitation standards, the cogeneration of electricity and district heating in modern plants and waste incineration facilities. In addition, the share of renewable energy was more than doubled since the mid-1990s.

This top position is an incentive for Vienna. The track record shows clearly that a lot of innovations and changes are possible in this city and in fact can be implemented in sometimes tough, gradual processes by many actors. Moreover, it should also be noted that, while long-term plans are necessary for resource-related objectives, the actual effects of change are due to concrete actions within foreseeable periods!

Intermediate objective: reduction of per-capita CO_2 emissions in Vienna by at least 35% until 2030 (compared to 1990)



With this reduction target, Vienna responds to the probable EU reduction goal for the non-ETS sector. In this respect, the EU is striving for a reduction of CO_2 emissions by 30% until 2030. As a result, the EU-wide per-capita value of 5.9 tonnes (in 2005) or 5.5 tonnes (in 2010) would decrease to 3.9 tonnes in 2030. Already today, Vienna attains a markedly better value (i.e. 3.1 tonnes per capita) than aimed-for by the EU for 2030 and wants to arrive at approx. 2.6 tonnes per capita in 2030.

Towards this goal, assistance must be provided by corresponding frame conditions to be set by the EU and/or the Federal Republic of Austria. Moreover, the measures contained in KliP 2 must be systematically implemented, followed by an ambitious KliP 3 that should contain Vienna's

12 When determining primary energy consumption per capita – and contrary to the calculation of CO_2 emissions –, none of the energy consumption volumes shown in the energy balance for Vienna are deducted (e.g. fuel tourism, emissions trading) or, conversely, added (e.g. Schwechat Airport, »grey energy« from production attributable to products consumed in Vienna).

13 Cf. 2,000-Watt Society of Zurich: to attain a sustainalbe and just society, the City of Zurich has adopted the 2,000-watt model. According to this model, the known primary energy volume suffices to cover a continuous output of 2,000 watt per person, which corresponds to an annual energy requirement of approx. 17,500 kWh per capita. On its way towards a 2,000-watt society, Zurich wants to attain approx. 2,500 watt of energy consumption and not more than one tonne of CO_2 per person by 2050.

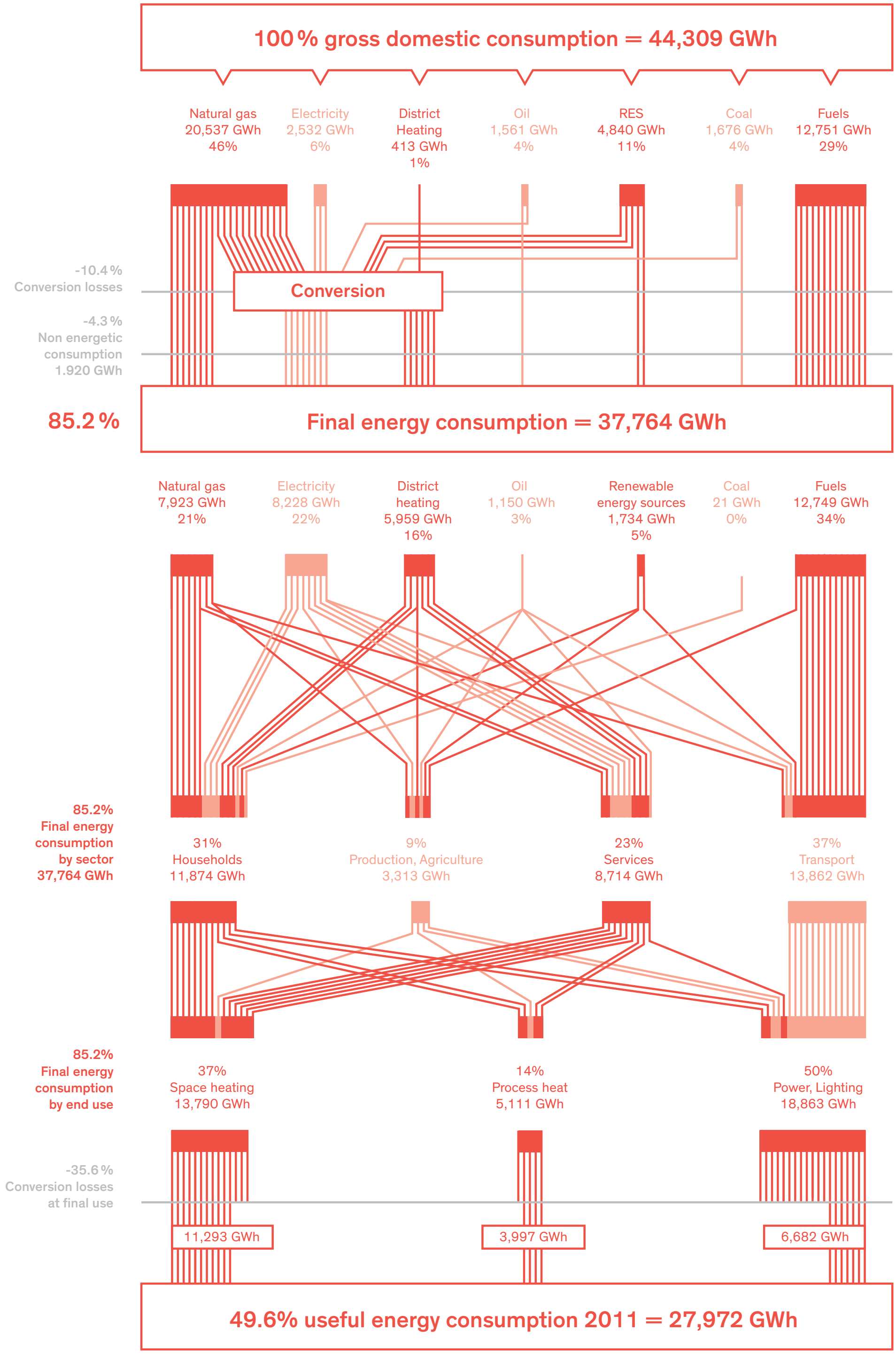


Illustration: Fig. on basis of an illustration by MA20/Typejockeys

Fig. 4 Energy flowchart for Vienna (status of 2012, data of 2011, source: Wien Energie, data by Statistics Austria) climate protection measures for the 2021-2030 period. It is assumed that the growth of renewables and power applications – in particular in transport – will be even more dynamic between 2030 and 2050. In 2050, the supply of urban regions with renewable energy originating in rural areas will be possible on a large scale. By the same token, even stricter energy-saving measures are assumed for the period from 2030 to 2050. These should be supported – possibly in the wake of price hikes for fossil energy sources – both by intensified market penetration of energy-efficient technologies and more energy-conscious behaviour on the part of consumers.

The objective of highest possible resource preservation embedded in the Smart City Wien framework strategy and the related sub-objectives are supported by specialised strategies and above all advanced by means of revisions or updates of these documents over the coming years. These e.g. include the Energy Strategy 2030 of the City of Vienna, the Climate Protection Programme KliP, the Security of Supply Plan including the Renewable Energy Action Plan, the Urban Energy Efficiency Programme, the new Urban Development Plan STEP2025 and the Mobility Concept derived from it.

These objectives are made possible by actions in the core areas of energy, buildings and mobility, since these areas are those with the highest energy consumption (see Fig. 4). Here, too, the importance of energy efficiency measures boosted by regulatory and financial instruments as well as due to changes in behaviour patterns becomes evident. Resource conservation is actively implemented by many people in Vienna as part of their lifestyle and assigned an overall positive image: it is a crucial task of the municipal administration to further anchor and foster these trends by means of strong incentives. The Climate Protection Programme of the City of Vienna has already achieved much in this respect: thus Vienna's greenhouse gas emissions decreased by 10% to 5.5 million tonnes of CO_2 between 1990 and 2010.[14] This was inter alia attained by means of improvements of building shells and energy-efficient technical services for buildings or intensified urban condensation, local shopping options and measures to step up cycling and walking.

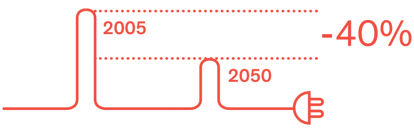
The objectives, the strategies taken to achieve them as well as a brief status quo overview of the four core areas will be presented in the following sections.

14 AEA 2012: Evaluierung der Umsetzung des Klimaschutzprogramms (KliP II) der Stadt Wien – When calculating Vienna's CO_2 volume, the emissions subject to emissions trading are deducted, as is all fuel consumption that cannot be attributed to Vienna (fuel tourism; fuel purchases attributed to the federal province where the company's HQ is located).

6.1 Efficient energy use and renewable energy sources

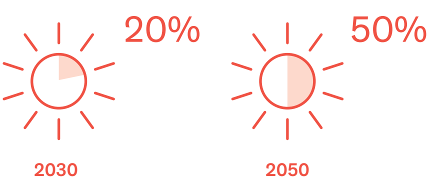
CO_2 reduction is achieved by increasing energy efficiency, stepping up the use of waste heat and renewable energy and cutting final energy consumption. The target values for energy consumption reflect the abovementioned long-term target of minus 80% of CO_2 per capita by 2050.

Increase of energy efficiency and decrease of final energy consumption per capita in Vienna by 40% by 2050 (compared to 2005). At the same time, the per-capita primary energy input should drop from 3,000 watt to 2,000 watt.



The remaining residual energy will be increasingly provided from renewable sources.

In 2030, over 20%, and in 2050, 50% of Vienna's gross energy consumption will originate from renewable sources[15].



The main emphasis regarding the further development of Vienna's energy system is on the intense use of local renewable sources and waste heat for both heat and electricity generation. The tapping and development of renewable energy sources (e.g. deep geothermics) for the district heating system and the considerable potentials of near-surface geothermics (depths of up to 300 metres) and solar energy are of great importance for the attainment of the above targets.

The importance of the production of renewable energy will continue to grow significantly in rural regions. As a result, a substantial portion of the renewable energy needed by cities will be generated in rural regions. These developments will be taken account of in the city's energy plans. By

15 These do not have to be located on municipal territory.

2050, electricity and gas imports from renewable sources will probably contribute decisively to overall energy supply.

EXAMPLE

Energy-autonomous wastewater purification

Thanks to an efficient purification plant, the waters of the Danube leave Vienna as clean as they were on arrival. But wastewater purification is very electricity-intensive. For this reason, the main wastewater treatment plant of ebswien already today opts for "SternE", i.e. electricity from renewable sources. A Kaplan turbine and a hydrodynamic screw produce green electricity, as do a wind turbine and a photovoltaics system. With the large-scale EOS project ("Energy Optimisation through Sludge Treatment"), the main wastewater treatment plant will also make optimum use of the energy contained in sewage sludge. As of 2020, ebswien will be able to generate all energy needed for wastewater purification independently from sewage gas, a renewable energy source.

EXAMPLE

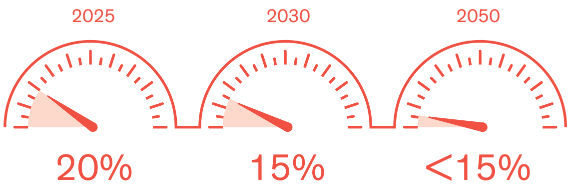
Zero Emission Liesing

The Liesing Mitte project is aimed at positioning an entire urban development zone as a central element of Smart City Wien by involving stakeholders and external experts and at taking concrete steps towards wholly renewable energy as well as massively reduced energy and resource input by at least a factor of 10. A parallel goal lies in enhancing the quality of living in the area, with special consideration of social aspects.

6.2 Resource-conserving mobility

Vienna is growing, and so is the number of trips taken within the city. In the field of mobility, attention is paid to ensure sufficiency as well as efficiency. Both the time required by citizens for everyday mobility and the number of trips taken are subject to only minimal variations. If the mode share of motorised individual traffic remained unaltered, the population growth would also lead to a rising number of car trips by 2025, resulting in its turn in increasing energy demand and pollutant emissions due to traffic as well as in the intensified use of already scarce urban space. This does not tally with the objective of high quality of living for all urban dwellers. Short distances can be easily covered by bike or on foot. Conversely, a shift in traffic that favours walking and cycling can in the long term strengthen urban structures with manifold service, shopping and leisure attractions in the immediate surroundings. Resource-conserving mobility means combining the claim of high quality of living with short distances. If the use of motorised vehicles is a necessity, these should run on CO_2-free, energy-saving propulsion types and renewable energy sources.

Strengthening of CO_2-free modes (walking and cycling), maintenance of the high share of public transport and decrease of motorised individual traffic (MIT) in the city to 20% by 2025, to 15% by 2030, and to markedly less than 15% by 2050.



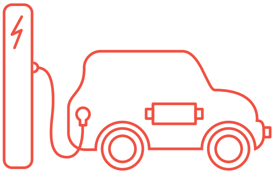
This calls for the timely use of effective measures in the mobility sector. These measures are developed and fine-tuned in terms of sector-bound strategies on the basis of the defined targets. A precondition for attaining the ambitious target for 2030 (MIT at 15%) lies in a marked increase of the share of cycling. Public transport (PT) will invariably remain essential. Public transport consumes only 5 to 6% of all energy used by the mobility sector[16] but delivers roughly the same transport performance in passenger kilometres as MIT. Therefore the energy consumption per passenger kilometre is only about one twentieth for PT as compared to MIT. If CO_2 emissions are compared, PT comes off even better since roughly three fourths of its energy consumption can be covered with hydropower or cogenerated electricity, both of which impact the CO_2 emission output only to a very limited degree. It is therefore planned to further reduce MIT and to substitute fossil fuels with electricity and renewables. The roadmap outlined in the White Paper of the European Commission[17] provides a path-breaking recommendation towards this goal.

By 2030, the largest possible share of MIT is to be shifted to public transport and non-motorised types of traffic or should make use of new propulsion technologies (e.g. electric-powered vehicles).

16 Municipal Department 20 (ed. 2013): Energiebericht der Stadt Wien; Energieverbrauch im öffentlichen Verkehr 1993-2011, acc. to Wiener Linien; p. 70.

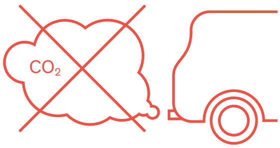
17 COM(2011) 144 final: "Halve the use of 'conventionally-fuelled' cars in urban transport by 2030; phase them out in cities by 2050"; achieve essentially CO_2-free city logistics in major urban centres by 2030; p. 10)

By 2050, all motorised individual traffic within the municipal boundaries is to make do without conventional propulsion technologies.



An important role for resource-conserving mobility is moreover to be attributed to urban logistics. Together with the logistics sector, the city will focus on the optimisation of goods and traffic flows by involving electromobility.

By 2030, commercial traffic originating and terminating within the municipal boundaries is to be largely CO_2-free.



Regarding traffic, it is also very important to adopt a regional view that extends across the municipal boundaries of Vienna, e.g. by means of integrated mobility and regional development with a special focus on commuter flows.

Reduction of energy consumption by passenger traffic across municipal boundaries by 10% in 2030.

The method of measuring and optimising regional traffic via the factor of energy consumption would constitute a novel approach. The intention is to embody this strategy at the regional level on the basis of city/environs mobility partnerships and transnational mobility management. The concept of multimodality and the establishment of mobility hubs can constitute a first step in this direction.

EXAMPLE

E-mobility on demand

The Vienna model region, with the “e-mobility on demand” research project at its centre, focuses on a gradual switch towards an integrated, comprehensive transport system. Public transport is thus effectively complemented by electromobility and e-car sharing. The new customer services will be simpler and easier to use. Electric-powered cars are to substitute fossil-powered trips in commercial traffic and ensure mobility in situations where walking, cycling and public transport use is not possible.

6.3 Buildings: built environment and new structures

In 2011, Vienna was composed of 165,000 buildings, 149,000 of which of a residential type, while the rest served service, industrial/commercial or other purposes.[18] Due to demographic change and the related population growth, it is estimated that approx. 120,000 new dwellings will have to be constructed by 2025. To meet the Smart City Wien goals, a new and ambitious framework for new buildings must be thus established. For this reason, energy standards, above all with a view to neighbourhoods and urban quarters in combination with new energy supply systems, must be redefined; likewise, the thematic complex covering buildings, energy and energy systems must be jointly discussed.

Even today, new buildings are planned and built according to very high standards of energy efficiency (low-energy standard and its successors). Energy and heating systems are always included in these considerations as well. According to the EU Energy Performance of Buildings Directive (EPBD) of 2010, the zero-energy standard will be mandatory for all new building types, with cost optimisation being taken account of in defining and updating all related requirements.[19]

Cost-optimised zero-energy building standards for all new structures, additions and refurbishments from 2018/2020 and further development of heat supply systems towards even better climate protection levels.

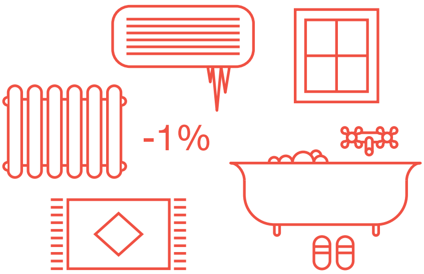
This demands excellent energy performance of buildings, which in combination with the systematic and optimised use of renewable energy sources and on-site waste heat or district heat entails minimal CO_2 emissions.

In addition to new buildings, great attention is also paid to building rehabilitation. Above all in municipal and co-operative housing construction, great advances were made in this field over the past 20 years. Now the task lies in also tackling challenging segments of the built environment, e.g. with regard to façade structures in need of protection, the situation in dwellings with tenant protection, the selection of energy sources or incentives for commercial or industrial premises, etc.

18 Statistics Austria

19 Cf. EU Directive 2010 (EPBD) and its implementation in Austria in 32 provincial laws.

Comprehensive rehabilitation activities entail the reduction of energy consumption of existing buildings for space heating/cooling/water heating by 1% per capita and year.[20]



This requires frame conditions that go beyond mere subsidisation, as the experience of recent years has shown that the latter tool does not generate a leverage effect strong enough to permit the attainment of ambitious rehabilitation goals. In the future, rehabilitation rates and qualities will have to be additionally boosted by means of other instruments, e.g. of a regulatory or fiscal policy type. Changes in housing legislation (e.g. tenancy law) or tax law would be useful in this context but are normally a federal competence. In keeping with the National Plan under the EPBD, Vienna will focus on further developing detailed quality requirements for the thermal and energy rehabilitation of buildings (e.g. concerning in-house utilities, heating systems, insulation, etc.). In addition, it is essential for the city to lead by example regarding its own buildings, with special priority assigned to the use of green systems (district heating and renewables) for space and water heating. Moreover, (roof)space potentials are to be systematically drawn upon for tapping solar energy.

EXAMPLE

Cities as the world's biggest mines

Cities as treasure troves: urban mining projects, too, are being initiated in Vienna, with Wiener Linien as a case in point. The intention lies in obtaining a better understanding of the building stock of the city and documenting it in order to be able to reuse valuable resources and raw materials ("the city can recycle itself").

6.4 Infrastructure and information and communication technology

Vienna is a city that functions excellently with regard to basic infrastructure facilities such as water supply, wastewater disposal, i.e. sewerage and rainwater management, as well as waste recycling and management.

20 This target value assumes that the Federal Republic of Austria and the EU will provide corresponding, supportive frame conditions.

Maintenance of the high standards of Vienna's infrastructure facilities.

Due to the Spring Water Mains originating in the spring protection zones of Rax, Schneeberg and Hochschwab, Vienna boasts drinking water of outstanding quality and abundance to meet the needs of even a growing population. Since 1873, 95% of Vienna's water supply has been safeguarded by gravity conduits, i.e. without artificial pressure boosts by pumping stations. Vienna's drinking water also contributes significantly to the production of green electricity: at the moment, 15 drinking water power stations generate approx. 65 million kilowatt hours of electricity annually, which corresponds to the power demand of around 20,000 households. Moreover, per-capita water consumption is also slightly on the decrease due to novel, water-conserving technologies.

Issues such as wastewater disposal, waste recycling and waste disposal are dealt with by specialised strategies of the City of Vienna. There e.g. exist targeted efforts for the use of rainwater management or strategies forming part of the Waste Management Plan and the Waste Avoidance Programme for Vienna.[21] An international comparison shows that Vienna's performance standard in this field is extremely high; it was moreover possible to create a positive image for waste avoidance and waste separation in the minds of citizens.

ICT as the "nervous system" of Vienna as a smart city

Information and communication technology is a central driver of innovation and a special asset of Vienna. The city assigns high priority to this sector – from science to business and public services – under the Smart City Wien framework strategy. This pertains to both the generic and infrastructure character of ICT and the role of ICT in shaping many services in an innovative fashion. Here, the most important task lies in the fact that the city views itself as an advanced client, provider and enabler of digital services. In this context, Vienna is committed to the open government principle and the related concepts of participation and transparency, but also to data security. The further development of high-quality e-government services of the City of Vienna is on the way. This concerns important issues such as the Virtual Office or the open government data catalogue, which is currently meeting with great interest on an international scale as well. In this way, innovative applications can be created for the benefit of citizens in such areas as energy, health, culture, environment, transport or housing, thus enhancing the intensity of use by both inhabitants and business.

21 Planning period: 2013-2018

Services offered by the City of Vienna are to be made more easily available, in particular with the aid of mobile end devices; this also calls for improved WLAN provision. In this context, care is taken to ensure that this offer will be balanced and attractive for different target groups to safeguard equal opportunities.

In 2020, Vienna is the most progressive European city with respect to open government.

The systematic expansion of digital public services taps economic possibilities. This is true of apps developed by individuals as well as of business opportunities for small and big companies, e.g. in the context of innovative pilot projects and applications. In this, ICT should be understood quite broadly as ranging from communication ventures to applications in areas such as health, energy supply or education. Pilot projects are to change processes in exemplary fashion and at the same time help to access efficiency potentials in combination with staff skills. Concurrently, new services are emerging, as are new forms of presenting the city in texts and visuals.

The next 100 apps in three years.

Pilot projects with ICT enterprises are to serve as showcases for the city and its economy.

Furthermore, ICT is a strong component of infrastructure. Communication infrastructure facilities should be viewed as the "neural pathways" of Vienna as a smart city. The new challenges such as big data initiatives – as well as users – need strong on-site infrastructure. Investments in latest-generation glass fibre and radio networks support the economy, users and the public sector.

In three years, Vienna will have a comprehensive WLAN.

Finally, the city will intensify its co-operation with universities, research institutions and universities of applied sciences to further strengthen the ICT competence of Vienna as a business location.

SMART CITY

INNOVATION

7 Objective: **innovation leader through cutting-edge research, a strong economy and education**

Smart City Wien means making use of innovations early and in a intelligent fashion, developing competencies and potentials and enabling the city to take a dynamic road into the future. The motto of the Smart City Wien campaign – "Vienna has 1.7 million brains. Let's use them!" – is nowhere more applicable than here. It is all about intelligence, creativity and critical analysis. The more diverse a city, the higher its potential for dynamic development. These potentials of society must be made use of, and the inclusion of broad strata of the population is a prerequisite of any successful, innovative city. This calls for opportunities for all to develop according to their possibilities as well as for good education and training options or skill building across the entire population. In its turn, this presupposes a major task of the city, i.e. to provide frame conditions, institutions and supporting services from infant pedagogy at the kindergarten to universities and an innovative economy. Using 1.7 million brains therefore means that top-end innovation must rest on a strong and broad foundation.

Objective: in 2050, Vienna is an innovation leader due to top-end research, a strong economy and education.

Innovation is key for linking resource preservation to high quality of living for all Viennese. This harbours the great opportunity for Viennese enterprises and research institutions to develop new solutions and competencies and in this way to garner respect and revenue across the world. Vienna holds the potential to be a major centre of excellence in the field of municipal services or urban technologies, but also in terms of balancing interests and supporting social innovations. However, many of these ambitious goals cannot be attained by simply perpetuating time-tried formulas: new forms of service delivery must be developed for Vienna. Conversely, the economy, society and administration must also be ready and able to absorb innovations from all over the world in a timely and intelligent manner. This calls for openness to science, research and innovation at all levels and in a way that is dynamic, reflective and, if possible, participatory. Education, research and, last but not least, a dynamic economy are crucial aspects for the future evolution of Vienna into a truly smart city.

7.1 Research and use of new technologies

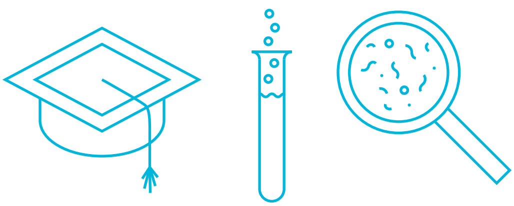
Research and new technologies generate added value, ensure high-quality workplaces and help to protect quality of living. At the same time, in order to realise its smart city ambitions, Vienna needs both the ability to absorb research findings from all over the world in a productive manner and the capacity to be a high-ranking research and university location in its own right. With currently in excess of 185,000 students, a strong industrial research and innovation sector and about half of all basic research activities in Austria, Vienna disposes of excellent institutional preconditions. Fields such as life sciences or information and communication technology present a high density of scientific achievements, industrial competencies and numerous new companies that are emerging day after day. Measured by the number of publications, citations, top-end researchers or patents, Vienna is an important research and innovation hub already today, but must still undertake efforts to catch up with the global top batch.

In 2050, Vienna is one of the five biggest European research and innovation hubs.

The next steps along this road will be defined in 2014/2015 with the development of Vienna's research and innovation strategy "Innovative Vienna 2020". Smart City Wien will play a key role in this strategy process. This involves the following relevant issues: Vienna disposes of top-end research equipment infrastructure that is equally used by science, industry and small or medium-sized companies. Research and innovation efforts in the areas of energy, mobility, climate and sustainability – all of which are particularly relevant for the smart city angle – are fostered, and great attention is paid to social innovation. Vienna places special emphasis on top-end research and the further strengthening of already successful areas; this also helps to enhance its attractiveness for research institutions of international corporations and top-class researchers. The city generates strong incentives to promote the continuous increase in the number of R&D workers and organisations engaged in research. In the context of the task sharing between the Federal Republic of Austria and the federal provinces, Vienna will earmark corresponding resources and pay special attention to know-how transfer directed at the economy and society. The contribution of women to key future-oriented sectors is to be specially promoted.

By 2030, Vienna attracts additional research units of international corporations.

In 2030, Vienna is a magnet for international top researchers and students.



Over the next decades, the »innovation triangle« Vienna-Brno-Bratislava will become an innovation system characterised by strong dynamism and networking, with Vienna acting as the innovation driver of the region.

By 2030, the innovation triangle Vienna-Brno-Bratislava is one of the most future-oriented cross-border innovation regions of Europe.

Vienna systematically pursues the ambition of positioning itself as an innovation leader in public service delivery. The Austrian capital favours living labs in the deployment of innovations so as to further develop and network its public services. The municipal administration and its spinoff enterprises promote technical and social innovations and the use of the latest technologies in their various areas of work so as to both improve public service quality and render administrative processes more efficient.

The city acts as a key customer of innovative products and contributes its know-how actively to regional research and innovation projects.

EXAMPLE

Delivery of innovative solutions

The city embodies innovation-oriented and resource-conserving public procurement methods clearly in its structures (pre-commercial procurement) and intensifies the "WienWin« programme, which likewise serves as a setting for pilot projects. Equal opportunity aspects are moreover increasingly used as a criterion as well.

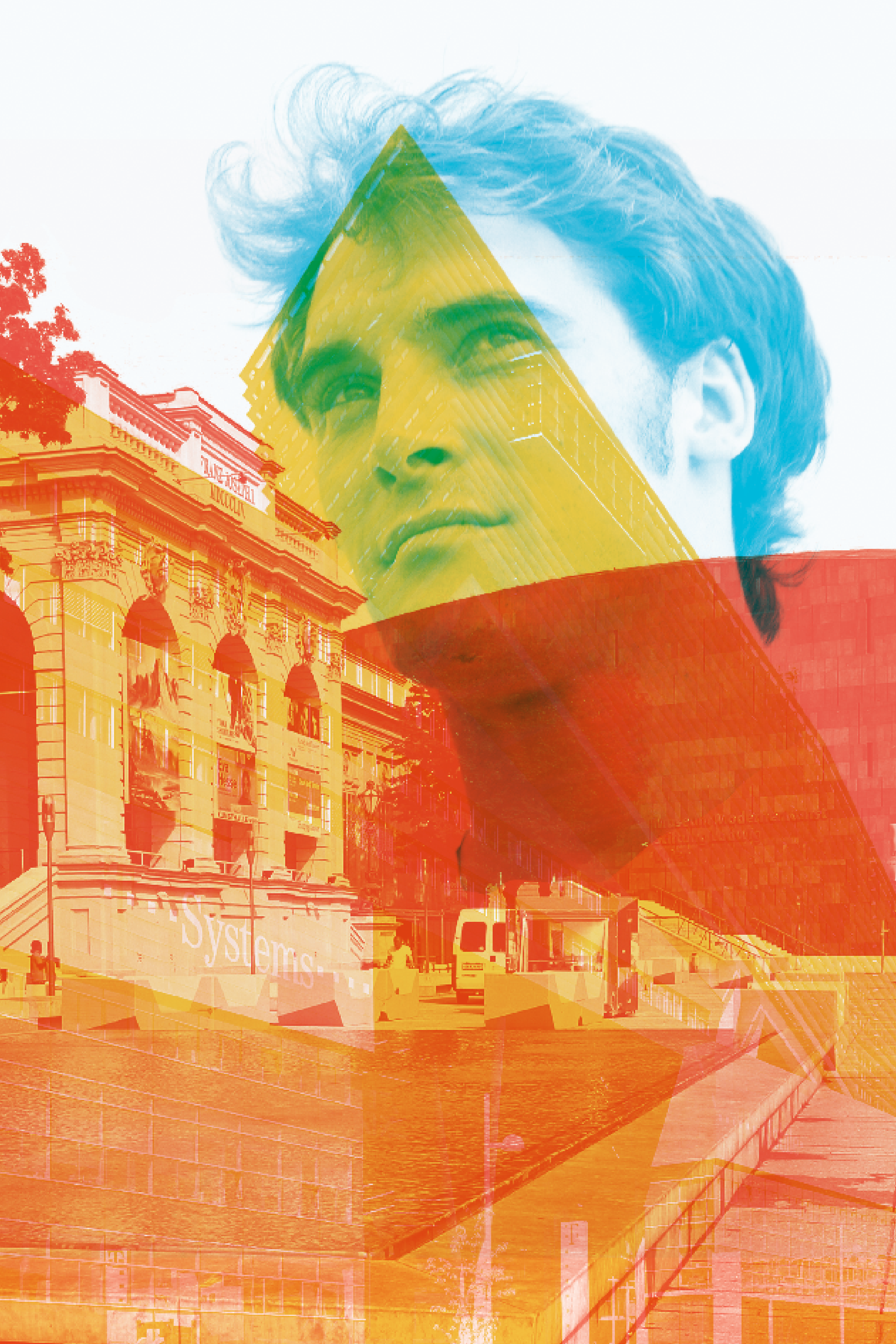
7.2 The 21st-century economy originates in the city

Due to their diversity, density and innovative clout, cities are ideal breeding grounds for a strong economy. Only recently, Vienna was mentioned as a prototypical prospering city by the UN[22]. Any prospering city also needs a diversified and varied economic structure, in which different industries and company sizes can survive and grow with success and co-operate with manifold organisations. This diversity entails resilience vis-à-vis critical developments and is to be safeguarded for the future. While the service sector predominates, the city does dispose of an excellent industrial basis with highly resource-conserving production methods, good productivity and strong export orientation. Regarding the consideration of environmental aspects ("green GDP"), too, Vienna may be called exemplary.

In 2050, Vienna remains one of the ten European regions with the highest purchasing power based on per-capita GDP.

In 2050, Vienna continues to enjoy the highest gross regional product per capita of all Austrian cities and towns, i.e. the urban economy presents a very high productivity level. Prosperity goes beyond mere material security; for this reason, Vienna aims for a top position in existing and future alternative methods of measuring prosperity. Vienna offers and enables meaningful work for all. In this, workplaces in Vienna correspond to the criteria of "good work", i.e. employment is for an indeterminate period and fulltime (if requested by the worker); payment corresponds to a

22 UN-HABITAT study: "State of the World's Cities 2012/2013"


Eva
Hesse
Systems

“living wage” standard; collective bargaining regulations are complied with. Access to the labour market is low-threshold and equitable, in particular also for less advantaged parts of Vienna’s population. Persons with special needs are integrated into the primary labour market. Women and men contribute equally to generating this prosperity; there are no wage or salary gaps. The possibility of productive participation of all persons in the many embodiments of the labour market makes for a decisive factor of quality of living in this city and an equally decisive contribution of the economy to Vienna as a smart city. The attractiveness of Vienna as a workplace is strong enough to motivate people from all over the world to come to Austria in order to work here. Smart City Wien is a cosmopolitan city all around.

Universities, universities of applied sciences and vocational training make sure that the skills profile of labour supply will largely coincide with labour demand. In this way, the Vienna 2020 Qualification Plan and its revisions play a decisive role in shaping Vienna’s labour market.

Vienna is a diversified and competitive economic hub characterised by a high level of co-operation within the city and the surrounding larger region. This region presents a diversified company structure in technology-intensive industries that include both the service and the manufacturing sectors. Enterprises are able to position their know-how and products successfully in the global market.

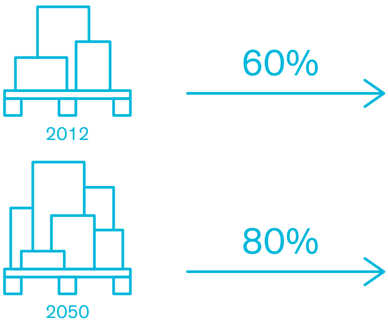
Particularly in technology-intensive industries, both the share of workers and the contribution to value creation overall has increased. Enterprises new to Vienna can draw on a wide range of services and counselling options with state-of-the-art ICT technologies.

Vienna further strengthens its position as the preferred company HQ city in Central/South-eastern Europe.

10,000 persons annually set up an enterprise in Vienna.

The direct investment flows from and to Vienna have doubled as compared to 2013.

The share of technology-intensive products in the export volume has increased to 80% by 2050 (as compared to 60% in 2012).



With a highly differentiated subsidy policy, the city ensures that future-oriented enterprises can find their niche in the market and use Vienna as a launching pad to conquer global markets. This also includes comprehensive counselling and assistance during the start-up phase. With counselling services and the provision and development of suitable real estate, the city offers a broad, integrated service range.

7.3 Education and qualification as a basis for Vienna as a smart city

A high level of education signifies equal opportunities and possibilities of participation. Education touches the very core of any smart city: the possibility of changing things either as an individual or as a group and to be able to realise one's professional and private plans underpin each and every structure and system. The instruments to pursue this approach are provided by ample opportunities for education. Qualification and education are likewise central starting-points for prosperity, quality of living and innovation. Therefore nobody must be excluded from educational processes; overcoming the educational disadvantages besetting certain social groups is a declared objective of the city.

In the future, only a strategy of lifelong learning will make it possible to meet the high standards of the modern work environment. The city will take measures to ensure a correspondingly high level of education and training of workers. The Vienna 2020 Qualification Plan serves as a central document towards this purpose.

As a smart city, Vienna is particularly interested in safeguarding that as many children and young people as possible will complete a highly qualified educational and training curriculum. Priority is given to creating a level playing field for all youngsters. For this purpose, the range of crèches and kindergartens will be further enlarged; many (mainly elementary) schools will be transformed into whole-day schools, and the objective of an integrated school for all children aged 10 to 14 years will be systematically pursued.

Comprehensive provision of whole-day and integrated schools and further upgrading of high-quality childcare offerings.

These objectives in particular serve to boost the educational potentials of children and young people in their school careers and support the improved compatibility of work and family. Another aim is to markedly decrease the number of early school-leavers, i.e. those youngsters aged 18 to 24 years who have not completed upper secondary schooling[23]. The Vienna 2020 Qualification Plan with its clearcut objectives and orientations constitutes a particularly important strategic document, e.g. with the targets of reducing the number of early school-leavers to 8.1% by 2020 (from 11.8% in 2009) and increasing the number of non-apprenticeship-based occupation permits by 2015 as well.

Even after 2020, it will be important for Vienna to make sure that a maximum number of young people will continue their education beyond compulsory schooling and thus attain a higher educational level.

The Vienna 2020 Qualification Plan focuses above all on education, occupational skills and labour market entrance possibilities and offers a strategy to reduce the share of poorly qualified persons. More young people are to obtain formal educational qualifications above the compulsory level in the primary educational system; more adults are to acquire formal school-leaving qualifications through second-chance education. One instrument of the Qualification Plan is the qualification pass (underway in 2014), which serves to document partial qualifications and skills of individuals in order to allow for their further development. Special attention is paid to gender-equitable pedagogy.

23 AHS (higher general secondary schools), BHS (higher vocational schools), BMS (intermediate vocational schools), apprenticeships and vocational schools

Even after 2020, it will be important for Vienna to safeguard positive frame conditions for acquiring formal school-leaving qualifications through second-chance education and for the recognition of educational attainments obtained abroad by adults.

EXAMPLE **Vienna Campus Plus Model**

At the moment, the **Vienna Campus Plus model** (school campus model) is being continued and further developed in nine sites under the "Construction Plan for New Educational Facilities 2012-2023". According to this model, kindergartens, schools and leisure education are concentrated in one location. This co-operation is to ensure the optimised use of all resources. These educational centres offer integrated teaching and ancillary spaces to include children with special needs, enabling them to fully participate in the educational process on an equal footing. The implementation of the Vienna campus plus model entails the construction of energy-efficient buildings whose spatial configuration also allows for a variety of multiple uses and efficient logistics. Gender aspects are specifically taken aboard during the planning and design of the buildings.

SMART CITY

QUALITY OF LIFE

8 Objective: ensure top-level quality of living

International rankings and studies emphasise that Vienna presents particularly high quality of living and an equally high degree of social participation. This is a key location factor and plays a central role for urban development. Quality of living is a multidimensional phenomenon and the sum total of numerous influencing variables. In addition to socioeconomic parameters and material living conditions, further key dimensions serve to define it: thus the individual's life satisfaction interacts with his or her career and educational attainments, with the quality of the environment and nature, with safety and security, with social participation, leisure quality and culture as well as with openness and diversity of gender roles; finally, physical, psychological and social health, too, is important. In this context, it is essential to continue the Viennese approach of social inclusion, which ensures the participation of all.

Objective: Vienna maintains its quality of living at the current superlative level and continues to focus on social inclusion in its policy design: as a result, Vienna in 2050 is the city with the highest quality of living and life satisfaction in Europe.

In addition to objective factors, the subjective level is of great significance here. As a result, it is imperative to look at individual statements and assessments to obtain a clear picture. As a smart city, Vienna takes account of the different urban living environments and realities of women and men. Comprehensive equality of a political, social and economic kind is a key element to ensure quality of living.

8.1 Social inclusion

Social inclusion stands for an open society and solidarity, good neighbourly relations, mutual respect and acceptance. Social pluralism and diversity are viewed as an opportunity. This calls for a corresponding design of the social safety net – it must be as tightly knit as necessary and as individual as possible.

In this, all residents of the city are united by a common language. Social and political participation of all population groups[24] – in particular of migrants – is promoted, as are learning German and plurilingualism. Recognition and nostrification mechanisms for qualifications obtained abroad by newcomers to Vienna should be optimised. The high level of cultural events and access to these for all citizens constitute another, equally important aspect. Special attention is paid to the potentials of youngsters from families with a migration background.

All people in Vienna enjoy good neighbourly and safe life conditions irrespective of their background, physical and psychological condition, sexual orientation and gender identity. Vienna is a city of diversity that is expressed to the fullest in all areas of life.

With regard to affordability, special emphasis is placed on housing and housing costs.[25] Particular attention should be paid above all to persons at risk of poverty or persons unable to pay energy bills due to straitened circumstances. Moreover, the housing environment, e.g. attractive public spaces, shopping and service provision "around the corner", access to cultural and educational facilities and easy accessibility, should be given priority here.

High-quality, affordable housing and an attractive housing environment are made accessible to the largest possible share of the population.

24 At all levels: education, work and career, culture, social affairs and health, safety and security, housing and spatial context, barrier-free design, identification

25 Housing costs = rent, service charges, heating, energy and maintenance (minus any housing or rent allowances) + payment of interest on loans for the creation or rehabilitation of housing space. Housing costs are deemed unreasonable if the amount paid for housing minus allowances (if any) exceeds one fourth of the disposable annual income.

In addition to possibilities for individual development and the fostering of social contacts and social skills, active participation in the world of employment must be adequately remunerated to safeguard economic participation and a "living wage", i.e. coverage of all basic needs in life.

Active participation at work as well as the performed work must be adequately remunerated and ensure the coverage of all basic needs in life.

To ensure that all residents of the city will be able to realise their life plans, the equitable participation of women and men in social and political decision-making processes is a declared key goal of the Smart City Wien framework strategy. The know-how and experience of both sexes are needed to safeguard a truly humane city.

Women are involved in planning, decision-making and implementation processes in keeping with their share in the total population. All persons involved in these processes dispose of gender competence.

Female and male city-dwellers have a different (subjective) sense of security. Thus public spaces should be rendered attractive for women – and hence voluntarily used by them – as well.

EXAMPLE

Vienna belongs equally to both women and men

To attain gender-specific equality of opportunities, the City of Vienna decided to conduct a gender equality monitoring along 15 selected thematic areas and based on approx. 120 selected indicators. Systematic and continuous data monitoring is to draw attention to relevant developments in the gender equality process and demands necessary corrections where required. In due course, this will foster the further development of specific actions for women so as to gradually eliminate the existing gender gaps in various fields.

8.2 Health as a prerequisite

Physical but above all psychosocial health is an essential factor for the individual's wellbeing and life satisfaction. All Viennese residents should enjoy maximum quality of living and life satisfaction on a daily basis, irrespective of their sex, background and age.

Health-promoting conditions of life must be further strengthened, e.g. sustainable safeguarding of high-quality natural resources (air, water and soil) or healthy nutrition with high-quality foodstuffs. Preventive measures must be rendered as effective as possible. This calls for the fostering of the health literacy of all population groups to prevent diseases and disorders triggered by health-impairing lifestyles. A health-promoting environment should be instituted, safeguarded and nurtured for all age groups and life circumstances, from infancy to adulthood, from kindergartens and schools to health promotion at the workplace.[26] To ensure the healthy development of children and adolescents, awareness of healthy nutrition modes and physical exercise must be stimulated already in kindergartens and schools. Physical exercise in daily life, e.g. walking or cycling, must be rendered attractive by corresponding design in an everyday context and requires promotion to stimulate interest.

Strengthening of health-promoting conditions of life and health literacy of all population groups.

Security of service provision requires equal opportunities regarding access to medical services for all citizens of Vienna and correspondingly presupposes the sustainable safeguarding of solidarity-based funding for the public healthcare sector.

Efficiency means demand-oriented treatment and care process management across all care levels[27] with a focus on the "best point of service".[28] Future structures and care processes in the healthcare system deliver better quality of medical care for patients and the proper medical service, provided at any time, in the right place and by suitably qualified medical professionals.

26 See Viennese measures relating to the concept of the Provincial Health Promotion Fund

27 See Health Target Control of Federal Target Control Contract; primary care, outpatient specialised care and inpatient care

28 This safeguards that the correct service is delivered at the right moment and in the right place by providing optimum medical and nursing quality in the macro-economically most cost-efficient manner.

Safeguarding of medical care at the highest level due to demand-oriented and efficient supply structures and processes (best point of service) for all citizens as well as reduction and shortening of hospital stays.

Efficiency and strategic focuses in medical care are necessary contributions to enable long-term security of service delivery according to the principles of solidarity. The Vienna 2030 Hospitals Concept pursues the goal of concentrating medical services to improve their quality and ensure optimum use of existing resources in order to prepare Vienna's hospital system for future financial and quality-related challenges.

To safeguard a strong and socially equitable public healthcare system, the Vienna Hospital Association and its facilities will remain a publicly-owned enterprise.

Potentials for greater efficiency must be systematically reviewed and used in all areas.

Moreover, demographic change and an ageing society pose rising demands regarding nursing care for the elderly in geriatric centres and nursing homes.

"Outpatient over inpatient" is the organisational principle of nursing services – letting persons stay in their own home for as long as possible while offering top-notch nursing quality. "Work/life balance" is a concept that assigns greater importance to time and leisure activities when assessing quality of living. As a smart city, Vienna contributes to optimising the everyday life of its residents and fosters innovative solutions to promote flexibility of site and control over time, i.e. quick and efficient processes on the one hand and deceleration on the other hand.[29]

The inhabitants of Smart City Wien are happy with the quantity and quality of their leisure time.

29 This refers to accessibility, tightly knit structures of service provision, removal of barriers, avoidance of motorised individual traffic, multimodality, e-government, etc.

BEISPIEL **Investments in health**

The Vienna 2030 Hospitals Concept aims to concentrate services at seven central hospital organisations in order to enhance the quality of service delivery. Other goals are modernisation, efficiency and process improvement – for this purpose, extensive investments for the benefit of hospital structures will be implemented until 2030. At the same time, the City of Vienna is responding to the increasing demands regarding nursing care for the elderly in geriatric centres and nursing homes with the objective of favouring structures near clients' homes and promoting outpatient care and assistance in order to help these individuals to remain in their own homes for as long as possible while enjoying maximum quality of living.

8.3 Vienna as an environmental model city

The share of green makes up half of Vienna's municipal territory. As a "green lung", these areas contribute significantly to Vienna's high quality of living. The city depends on the functioning of the ecosystems that surround and permeate it. Spacious and attractive green and open spaces within the municipal territory must be safeguarded even in a growing city; they must be easily reachable by eco-friendly means and should be barrier-free and publicly accessible. This is a key contribution to quality of living and life satisfaction.

By 2030, the share of green spaces must be kept at over 50%. Especially in a growing city, additional recreational areas must be safeguarded to keep up with the rising population figures.

This calls on the one hand for the preservation and networking of large-scale protection zones, such as the extension of the Vienna Woods in the north-eastern part of the city. On the other hand, it also calls for a dense network of local green and open spaces and smaller vertical and horizontal gardens. Further measures to stabilise inner-city temperatures include the planting of trees in the city, façade and roof gardens, new buildings with greened flat roofs and neighbourhood gardens. Current guidance values for green space provision are being updated by adding such factors as accessibility, efficiency of supply and other qualities. A high share of green spaces is another key element in the city's efforts to safeguard biodiversity.

Air pollutant and noise emissions negatively impact human health and quality of living and cause huge economic costs, which must be minimised by forward-thinking planning as well as by early avoidance and protection measures. The City of Vienna realised its responsibility already at an early date and thus has developed comprehensive measures that are reflected in the measured values, which are largely gratifying. For example, the three packages of measures implemented by the City of Vienna to combat particulate matter made it possible in the past two years to comply with the European limit values. This approach is systematically pursued. Close co-operation with other relevant decision-makers, in particular the Federal Republic of Austria and the EU, is a key pillar towards this goal. On the basis of the EU Environmental Noise Directive, the City of Vienna has joined forces with the Federal Republic[30] to develop action plans, which contain numerous already successfully implemented measures to curb noise for the benefit of Vienna's population.

Traffic in Vienna is the main causal agent of noise and air pollutants. Smart urban development should create the spatial-planning and structural preconditions to motivate the population to switch voluntarily to eco- and climate-friendly mobility types.

Another objective must be to protect the soil because of its manifold ecological functions. Soil is a habitat, pollutant filter, infiltration body and CO_2 storage medium of the planet's biogenesis and contributes positively to the microclimate. Resource-conserving and efficient urban

30 Federal Ministries, ASFINAG, Austrian Federal Railways (ÖBB) expansion, i.e. minimal land consumption due to compact designs, the favouring of internal condensation and brownfield development are to curb soil sealing.

A notable part of Vienna's green spaces is used for agriculture: horticulture, viticulture and field crops. This is unique for a metropolis of this size. The municipal administration and the city government have for many years been promoting environmentally friendly, ecologically oriented methods of organic and genetically unmodified production.

Vienna's waste management system contributes not only essentially to making the Austrian capital a very clean city; it is also strongly oriented towards the Smart City Wien objectives and due to cutting-edge technology saves 550,000 tonnes of CO_2 equivalents annually, which is more than it actually causes through waste management (i.e. 420,000 tonnes of CO_2 equivalents[31]). In this way, it actually over-accomplishes the objective of a zero-emission city. This is achieved by producing district heat from residual waste incineration, the fermentation of kitchen waste in Vienna's biogas plant, waste separation and waste recycling as well as the use of compost for organic farming.

In 2020, municipal waste management attains approx. 270,000 tonnes of CO_2 equivalents by savings as a result of further measures and improvements.

EXAMPLE

A garden at the doorstep

More and more often, Vienna's citizens, too, like to take up a spade to create gems of shared, self-determined cultivation – it's called "urban farming". Over recent decades, new forms of urban farming have sprung up across the world, e.g. the community garden movement. The great benefits of this gardening culture lie in fostering social encounters and a sense of community, thereby creating the substrate for more intense commitment on behalf of the neighbourhood (called "Grätzl" in the Viennese dialect).

EXAMPLE

Ecological production and ecological consumption

Exemplary programmes, such as the Vienna EcoBusinessPlan and "ÖkoKaufWien", lead to significant reductions of the negative environmental impact of production and consumption. The former supports Viennese enterprises in the development of environmental and sustainability measures. More than 1,000 enterprises have taken part in this programme so far and saved over Euro 120 million in operating costs by means of approx. 15,000 measures for the benefit of the environment. The effect on the environment is

31 Study "Klimarelevanz der kommunalen Wiener Abfallwirtschaft" by denkstatt, commissioned by Municipal Departments 22 and 48 and Wien Energie Fernwärme. For the waste management system, this results in a savings surplus of 130,000 tonnes of CO_2 equivalents: this quantity corresponds to the emissions of electricity generation for 130,000 households or the emissions of 60,000 cars travelling 15,000 km each.

equally impressive: inter alia, waste production was curbed by 123,570 tonnes, which equals 1.53 times the capacity of Vienna's Ernst Happel Stadium; moreover, 1.04 TWh of energy was saved, corresponding to the annual consumption of 208,000 Viennese households; the emission of carbon dioxide was curbed by 305,000 tonnes, which can be equated with the volume of 51,300 hot-air balloons; 93.4 million transportation kilometres were likewise rendered unnecessary, corresponding to 2,330 trips around the equator. Furthermore, drinking water consumption was reduced by 2.573,500 cubic metres, equal to the volume of 830 Olympic-sized swimming pools.

"ÖkoKaufWien" is a programme for sustainable public procurement and covers everything from detergents and office supplies to services and construction works.

SMART CITY

WIEN

9 Links between the individual objectives

Being a smart city also means nurturing constant evolution and creating space for new developments. These new developments – be they services, forms of social encounter, business concepts, mobility types or expressions of culture – cannot always be contained in pre-appropriated spaces and certainly do not respect cut-and-dried competencies and rules of procedure. Some novelties and changes produce their added value quickly and at several levels, e.g. by conserving resources and improving the quality of living. Others may initially strain existing structures. This routinely puts the management of a smoothly functioning city to the test and requires it to prove its adaptability. First of all, this calls for openness and the willingness to question time-tried procedures from all actors involved. Past successes can only be projected into the future if this willingness is a given.

The special effect of the framework strategy should and will find its expression in the development of stronger links between the individual thematic areas and objectives:

- First of all, the multiple benefits of activities are rendered more clearly visible: thus high quality of living is often a result of resource-conserving changes, and innovations – ideally "invented here" – frequently reduce resource input.
- Moreover, potentials relating to all three objectives – resource preservation, quality of living and innovation – can only be tapped if tasks are viewed in a cross-cutting manner and limits of responsibility are overcome. If properly supported, local action is often able to break down such boundaries. A lack of co-operation entails high costs in terms of inconsistencies, duplications of effort or gaps.

- In addition, exemplary and awareness-creating measures in the field of the City of Vienna's activities as an entrepreneur under private law contribute significantly to stimulating awareness among the population, which is a key actor of the Smart City Wien Initiative. This includes the use of renewable energy sources in buildings, car pools or innovative pilot projects to demonstrate the positive attitude of the city towards new developments and innovations.

But not all interactions are solely positive. As a rule, Vienna is attuned to identifying such developments: both administration and citizenry keep a wary eye on potential innovations for potentially harmful side effects. As a smart city, however, Vienna invests as much energy in spotting positive effects and mutual benefits.

Already today, there exist many examples of activities that are "net contributors" to the smart city concept. Some are listed here.

Strengthening of sub-centres and neighbourhoods
The strengthening or establishment of sub-centres that offer a wealth of shops and services, short distances, lively open spaces and multifunctionality is to proceed in urban areas that so far lack such centres. Thus even a growing city should be able to safeguard short distances between housing and schools. The average share of 90% of primary school places close to pupils' homes is to be maintained. Other social infrastructure facilities, too, should be in close proximity and offer easy access for all users. Multiple forms of use of public spaces including work, housing and social activities must be fostered even more than today for reasons of efficiency and capacity use and also to promote a good mix.

Sustainability at the local level: from pilot district to urban quarter
In co-operation with the Municipal District Office of the 22nd district Donaustadt, sustainability was systematically implemented by monitoring the development of this specially selected district. This successful pilot project will be gradually extended to the entire city. The planned project identifies/initiates urban activities and places considered suitable for rendering sustainability "palpable". The corresponding places/activities will be concentrated – similar to the "Sustainable Donaustadt" map – and complemented by other activities to create even greater visibility.

Energy consumption in hospitals
With regard to their energy consumption, hospitals are considered "cities within the city". Big hospitals need as much energy as small towns – for medical equipment, lighting, heating, cooling or ventilation. Their contribution to the resource preservation objective e.g. lies in the intensified use of energy contracting, the production of their own energy, the use of renewables or electromobility.

Networked spatial, mobility and energy planning
More comprehensive solutions are necessary to meet future challenges. In order to network all aspects of spatial, mobility and energy planning already in early phases, integration at urban quarter and neighbourhood level in a joint process to identify optimal infrastructure solutions offers the possibility to consolidate Vienna's standing as a smart city. Networking and novel structures and processes must hence be conceived of as spatial (at the urban quarter and neighbourhood level) and in time (consider/involve all stakeholders already in early phases).

Fig. 5 Examples of activities that contribute in multiple ways to the Smart City Wien initiative

Temporary use of sites for cultural purposes

Vacant houses and shop premises offer the possibility of putting them to temporary use for artistic and cultural projects, which entails benefits for urban development. Over the coming years, the concrete objective lies in attracting a greater number of cultural institutions to urban expansion areas. The co-ordination project “einfach-mehrfach” of the City of Vienna initiates, supports and promotes projects for multiple and temporary use in all municipal districts and functions as an instrument of internal and external structural support. This implementation model also comprises aspects of authentic requisition (statement of needs) and participation, promotes self-organisation and co-operates with universities on a permanent basis.

aspern – Vienna’s Urban Lakeside

In the 22nd municipal district of Vienna, north-east of the city centre, aspern – Vienna’s Urban Lakeside is emerging in several stages until 2028 as one of Europe’s biggest urban development projects. Numerous exemplary initiatives are being implemented and often interlinked, e.g. in the fields of mobility, urban planning, innovation or energy, with the objective of creating a new, multifunctional and attractive part of Vienna with housing, offices and a commercial, science, research and education quarter grounded in the latest findings in energy efficiency, building standards and forms of use. One example is Smart City Research GmbH & Co KG (ASCR), a company established by the City of Vienna together with Siemens AG with the remit of researching and implementing energy efficiency solutions on the basis of real-life buildings in Vienna’s urban development zones over the coming years.

New building for Wien Museum Karlsplatz

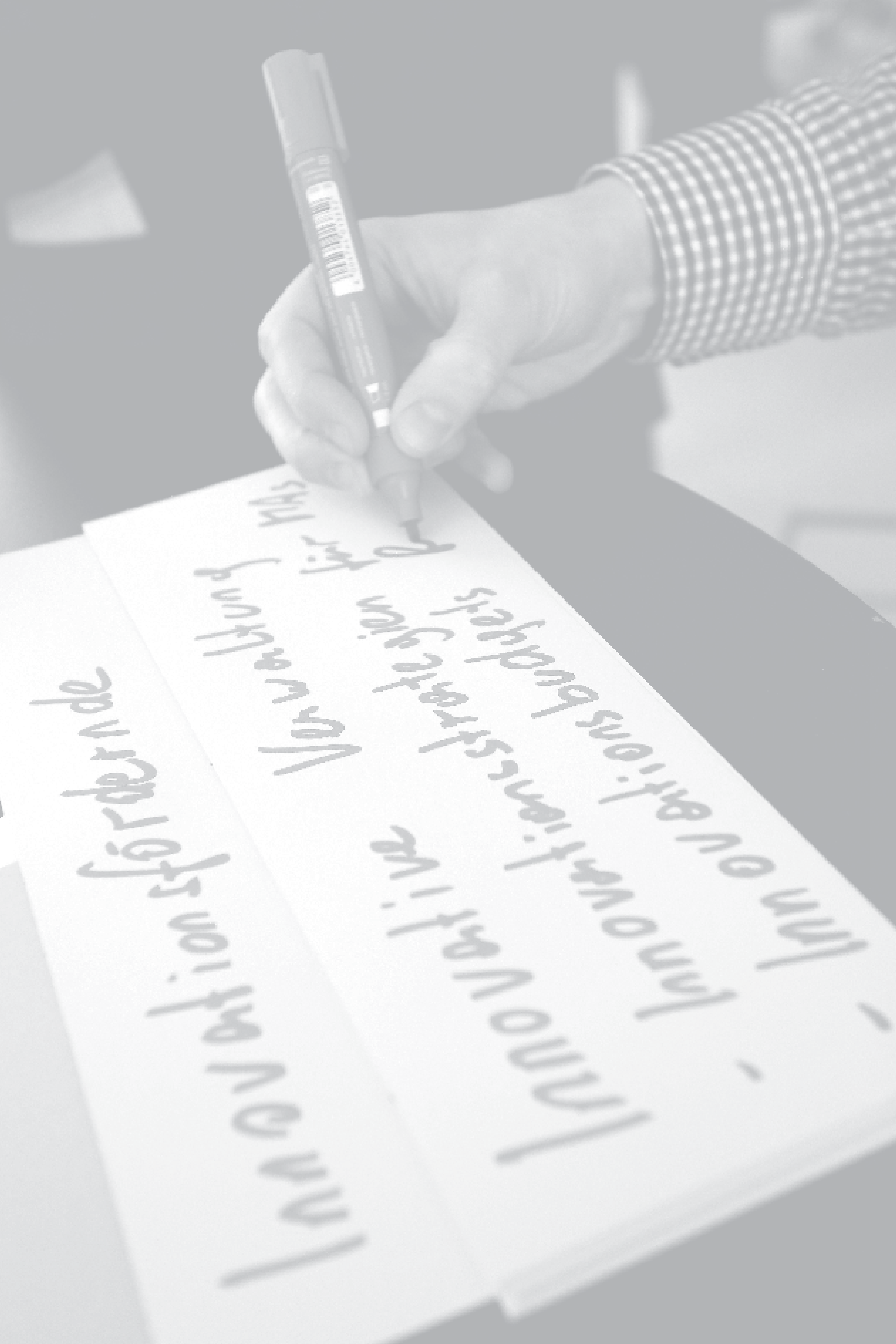
The City of Vienna will construct a future-oriented new building in Karlsplatz square (the current site of the museum) to house the Wien Museum. The new project will make use of the already existing structure. With this decision, Vienna is setting a clear signal of urban renewal, revitalisation, conversion, addition and condensation. The new building will correspond to all criteria of an ecologically sustainable and energy-efficient, 21st-century building typology.

Culture mediation

In the next few years, the numerous existing mediation programmes in the cultural field – e.g. the “Wiener Kulturpass” (culture pass for persons with low incomes) or the “Go for Culture” initiative – are to be complemented with others. In this way, innovative ideas include large strata of the population and whet their interest in cultural events.

Energy balance of educational establishments

On the basis of cost/benefit analyses, the energy balance of educational establishments and municipal office buildings should be improved by means of energetic refurbishment of the existing building stock and energy-efficient construction and operation methods for new educational infrastructure buildings to promote the Smart City Wien objectives.



10 Governance

Vienna's objective to be a smart city and the implementation of the framework strategy confront the Austrian capital with special challenges. Many objectives cannot be tackled through individual activities or competencies but require superordinate thematic management. The strong innovation orientation that is a hallmark of the smart city concept not only affects Vienna as a hub of research, education and business but also generates new instruments and approaches that govern the way in which the municipal administration and its enterprises design processes and, above all, render services. For the city and its residents, this means that service quality remains very high, and the ways of service delivery will evolve constantly, with special account taken of the different needs of users (gender and diversity).

The smart city approach has two primary levels of implementation: on the one hand, it concerns the political level and hence the privilege of setting political priorities and defining policies in view of increasing complexity coupled with tight resources. On the other hand, the smart city concept poses challenges for the operative level, also because many tasks can only be handled by cutting across individual organisational units. For the staff members and organisational units of the City of Vienna, this calls for even tighter co-operation within and outside the municipal administration. This "outside" harbours a particularly important aspect, i.e. the necessity of further intensifying the consultation processes with the Federal Provinces of Lower Austria and Burgenland, for example regarding mobility and regional development issues, on the basis of existing structures like PGO and SUM.

Perhaps the most essential task of the Smart City Wien framework strategy lies in the additional assistance it can provide for the numerous specialised strategies underway in Vienna. In this way, these planning documents, which usually reflect a time horizon of seven to ten years and deal with future-oriented questions of energy supply, climate protection, urban planning, occupational qualifications or research/innovation pick up greater momentum to define and pursue ambitious objectives and measures.

Towards this purpose, the City of Vienna will in the course of the coming years take the following steps in the areas specified on the next pages.

10.1 Co-ordination and co-operation, establishment of lighthouse projects

Smart City Wien means change and the tackling of larger thematic areas going beyond narrow departmental confines. Already existing examples include the SMILE mobility card of Vienna Public Utilities, ÖBB (Austrian Federal Railways) and municipal actors; ongoing Smart City calls of the Vienna Business Agency or the URBEM-DK co-operation between the Vienna University of Technology and Vienna Public Utilities in the context of a doctorate course on energy and mobility modelling.

To allow for interdepartmental strategic and "smart" control that also includes municipal enterprises, a suitable organisational model needs to be developed.

This might e.g. contain the following elements and implementation steps:

- Regular Smart City Wien steering rounds chaired by the Chief Executive Director of the City of Vienna with the assistance of a scientific advisory board.
- A Smart City Wien Agency as the central co-ordination point for all internal and external stakeholders. It should cover the areas of co-ordination, stakeholder management, inquiry management and communication and would record, evaluate and initiate projects on behalf of all relevant partners within and outside the City of Vienna. The objective lies in the interdisciplinary promotion of networking between municipal administration, research, business and industry.
- Setting-up of larger innovation projects while taking account of a broad base of different departments, municipal enterprises and third parties to cope with major challenges.
  - → For ventures of this kind, it is suggested to appoint project area managers who at the same time serve as "faces" and testimonials of Smart City Wien.
  - → These ventures would facilitate access to corresponding European innovation platforms and their funding in the Grand Challenges context of the Horizon 2020 framework programme.

    - → Examples of such thematic project areas might include mobility management or ambient assisted living (the latter of the Vienna Social Fund).
    - → Individual larger lighthouse projects with an innovative character will contribute to the attainment of key Smart City objectives.
    - → These projects will follow gender and diversity aspects.
- Strengthening Vienna's co-operation on smart city issues with universities and research institutions: setting up long-term collaborations, support in the recruitment of additional key personnel at Viennese universities and research institutions, "urban issues" as study content, research topics and testing grounds for social innovations.

10.2 Strengthening the participation possibilities of citizens and experts

Smart City Wien means creating a wider leeway for action for all Viennese. Codetermination and modern management go together, both in direct interpersonal contact and via the Internet.

Exemplary implementation steps:

- Large-scale rollout of open government as a principle and driver of innovation.
- Regular Smart City Wien stakeholder forums.
- Development of formats that transport Smart City Wien issues to kindergartens, schools and other educational establishments: a major initiative makes topics like energy efficiency, low-impact mobility, virtual worlds or coexistence in a city without poverty part of the syllabus and enables children and young people to build their own smart Vienna: "100,000 kids design their very own smart city".

10.3 Human resource development, training and recruitment

Smart City Wien projects offer a possibility for employees of the Vienna City Administration and its enterprises to learn about new things and test novel forms of co-operation. For this reason, questions relating to human resource development, training, recruitment and knowledge management are at the centre of the Smart City Wien Initiative.

Exemplary implementation steps:

- Further implementation of the knowledge management strategy of the City of Vienna. The objective lies in making optimum use of networking and further developing the store of knowledge accumulated by staff members. A tool for self-analysis ("Self Check") enables municipal departments to meet their most urgent needs with suitable methods and to align their work with the overall strategy of the Vienna City Administration ("Strategy House"). At the same time, a strategic unit is set up and a community of practice is initiated to safeguard the coverage of cross-cutting thematic knowledge.
- Further development of the diversity-oriented human resource management methods of the City of Vienna as well as of equality between women and men in human resource management.

10.4 Information and brand management for Smart City Wien

The Smart City Wien Initiative is designed to focus important steps towards change for the coming decades. A strong and broad-based communication strategy is to give vibrant life to this concept. This will only be possible for the Vienna City Administration through constant exchange and dialogue with the population as well as with numerous other partners. In this way, Vienna can be positioned as a strong brand in the international competition between cities.

Exemplary implementation steps:

- Continuation of the Smart City Wien campaign launched in 2013 with concrete projects and testimonials.
- Establishment of a Smart City Wien award with three categories – resource preservation, innovation and quality of living – for exemplary projects; in co-operation with external partner organisations.
- Stronger involvement of enterprises that support the City of Vienna in its strategy with their Smart City (Wien) projects. Applies both to external effects and co-funding.

10.5 Alliances, lobbying and consultation processes

With regard to innovation, energy and climate issues, cities are more and more at the centre of interest and policy design. Through co-operation, cities can give more weight to their concerns, e.g. the safeguarding of the principles of public services and services of general interest or the eligibility for subsidies in important areas.

Exemplary implementation steps:

- Intensification of city alliances in Austria but above all with other European metropolises to formulate demands useful to attain Smart City objectives.
  - → Active conducting of a debate with other cities and regions on how infrastructure investments could again be accorded a special role in the calculation of government debt ratios.[32]
- Development and implementation of three joint projects with the Federal Ministry for Transport, Innovation and Technology (BMVIT) based on the "Memorandum of Understanding between the City and Vienna and BMVIT" over the next three years.
- Systematic defence of the interests of Vienna and other big cities (lobbying, services of general interest, subsidies) on the European level.
- Canvassing of 20 patrons from the corporate sector, associations and civil society for important Smart City Wien projects over the next three years in combination with joint PR work and – where possible – financial contributions of these patrons. In general, relevant stakeholders outside the municipal administration are integrated into the processes on a long-term and binding basis.

32 The terms "investment-related government debt concept" or "golden rule" refer to methods of calculating the government debt ratio that do not count investments as augmenting the government debt level.

- Setting-up of a joint strategy development process for issues crossing the municipal boundaries by means of the Smart City Region Platform in the context of PGO. This is to result in concrete key projects in such areas as mobility and regional development.

11 Monitoring

For the implementation of the Smart City Wien objectives, a coherent monitoring and reporting process with a limited number of core indicators is to be established. The set of core indicators comprises status, target and policy indicators. For this purpose the established objectives and intermediate objectives with defined time axes are drawn upon. The degree to which the objectives of the Smart City framework strategy are met is measured by means of the core indicators assigned to each key objective. Detailed indicators are limited to the individual specialised strategies in the context of a fine-tuned process.

This includes the development of an ongoing process with

- an analysis of indicators
- the development of packages of measures and a definition of responsibilities
- decisions regarding the implementation of these packages
- continuous reporting and adaptation of the strategy, i.e. a definition of who may take necessary improvement measures, and what sort of measures these might be
- the implementation of these measures with suitable implementation monitoring

At regular, brief intervals, the status of implementation is determined. This can be done by means of mandatory, scheduled and shorter data assessments at certain moments in the form of a status report and by means of an analysis and interpretation of outcomes in strategic reports compiled at longer intervals. This makes it possible to monitor progress; it also allows for the fine-tuning and adjustment of objectives.

The monitoring process involves all departments in a cross-cutting fashion. To take account of target groups, it also evaluates process data in a gender- and diversity-specific style.

12 International level and models to follow

Cities are faced with many challenges of a social, technological and ecological kind. In the international context, the smart city concept is above all couched in terms of resource conservation and CO_2 reduction. At the moment in Europe alone, cities account for 70 % of total energy consumption. Moreover, the global urbanisation trend is continuing. While Smart City Wien also includes quality of living and social aspects, the international discourse on CO_2-related questions is still necessary.

If Europe wants to develop an innovative low-carbon economy and society, metropolises will therefore play the central role. The European Union is aware of this coherence and has already taken the first essential steps towards finding joint solutions. Thus the European Commission launched the “European Smart Cities and Communities” initiative, whose main goal is to network European cities and to promote future co-operation projects in order to jointly identify solutions for urban challenges and increase the energy efficiency of European cities. Moreover, the European Union is continuously adapting its established funding schemes and implementing new possibilities for improved co-operation to achieve the transformation into smart cities.

In keeping with the ambition to take an international pioneering role in the context of a future-oriented innovative smart city, the Smart City Wien Initiative has participated in European networks and joint funding projects with international partners from the get-go. Ongoing exchange and proactive involvement in European networks – e.g. the EU Smart Cities and Communities stakeholder platform, the Eurocities network, the Covenant of Mayors and many others – safeguard constant information flow and exchange with other European cities. Thus Vienna already today takes a solid role in the field of smart cities and urban technologies, which is reflected in a great number of international inquiries. Established networks and contacts entail co-operation in joint European research projects.

The Smart City Wien framework strategy and its preparatory process are embedded in European projects and parallel activities. Examples of these are TRANSFORM[33] and EU-GUGLE as part of the Seventh EU Framework Programme for Research, Transform+[34] and the INTERREG IV C project CLUE[35]. These projects allow for the generation of comprehensive and broad-based new knowledge. Through all these activities, Vienna maintains continuous contacts with other forerunner cities such as Amsterdam, Copenhagen, Hamburg and Stockholm. In the future, too, a great deal of attention will be paid to this international level and to corresponding co-operation projects.

33 Transformation Agenda for Low Carbon Cities
34 National project of the Climate and Energy Fund of FFG (Austrian Research Promotion Agency) to complement TRANSFORM
35 Climate Neutral Urban Districts in Europe

13

The strategy process

Under the aegis of Mayor Michael Häupl, the City of Vienna launched the Smart City Wien Initiative in 2011. Smart City Wien builds on existing approaches to environmental and climate policy, concentrates the available resources and makes sure that the collaboration between all actors will facilitate a joint focus on superordinate goals.

A crucial starting-point was provided by a broad-based stakeholder process initiated in 2011 with the project "smart city wien" [36] funded by the programme "Smart Energy Demo – FIT for SET" of the Climate and Energy Fund and continuously updated ever since.

The regularly organised forums – conceived as platforms for the exchange of ideas and opinions regarding successes, current developments and future challenges for actors, decision-makers and experts from municipal administration, research, business and industry – are an important element of the initiative and generate impulses for further project developments and participations.

In spring 2013, the Smart City Wien steering group chaired by Chief Executive Director Erich Hechtner decided to have a Smart City Wien framework strategy developed to support the further transformation of Vienna into a smart city and to identify the fundamental objectives necessary towards this goal. Municipal Department 18 (MA 18) – Urban Development and Planning was charged with heading the development of this Smart City Wien framework strategy. In a participatory process involving numerous group discussions, thematic workshops and interviews with more than 100 experts, the main topics of the framework strategy were then rendered concrete, leading to the formulation of objectives.

The Smart City Wien framework strategy should be viewed as a long-term umbrella strategy that spans the entire period until 2050 and encompasses all areas of municipal administration and urban policy in Vienna. Combined with other existing and future documents, plans and pro-

36 Project "smart city wien" with vision for 2050, Roadmap for 2020 and Beyond, Action Plan 2012-15 grammes, it defines a productive and structuring thematic framework. In addition to a vision that must be always kept in mind as well, phased objectives, concrete strategies and exemplary activities, instruments and projects are to lay the ground for co-ordinated political action, with a special focus on thematic areas of vital importance for Vienna.

On 24 July 2013, Mayor Michael Häupl and Federal Minister Doris Bures signed a memorandum of understanding (MOU) between Vienna and the Federal Republic of Austria to advance the smart city cause. The objective lies in initiating projects via a joint steering group and to obtain funding at a European level in order to support the implementation of the Smart City Wien framework strategy and the further transformation of Vienna into a smart city.

After a political consultation process, the Smart City Wien framework strategy was adopted by the Vienna City Council on 25 June 2014.

I. List of abbreviations

CO_2	Carbon dioxide
EnEffG	Federal Energy Efficiency Act (“Bundes-Energieeffizienzgesetz “)
EnStrat	Vienna Energy Strategy (“Energiestrategie Wien”)
EP	European Parliament
EPBD	EU Energy Performance of Buildings Directive
ICT	Information and communication technology
KliP	Climate Protection Programme (“Klimaschutzprogramm”)
MIT	Motorised individual traffic
MPV	Transport Master Plan (“Masterplan Verkehr”)
PT	Public transport
PGO	Eastern Austrian Planning Association (“Planungsgemeinschaft Ost”)
RAP_VIE	Renewable Action Plan Vienna
RTI	Research, Technology and Innovation
SCW FS	Smart City Wien Framework Strategy
SEP	Urban Energy Efficiency Programme (“Städtisches Energieeffizienz Programm”)
SMEs	Small and medium-sized enterprises
STEP	Urban Development Plan (“Stadtentwicklungsplan”)
SUM	City-Environs Management (“Stadt-Umland-Management”)
t	Tonnes

II. List of illustrations

III. List of references and bibliography

Beauftragter der Stadt Wien für Universitäten und Forschung (Commissioner of the City of Vienna for Universities and Research), Wissen schafft Wert – Hochschulen und Forschung als wichtiger Standortfaktor, Dritter Bericht des Beauftragten der Stadt Wien für Universitäten und Forschung (Third Report of the Commissioner of the City of Vienna for Universities and Research), Vienna, 2013, http://www.universitaetsbeauftragter-wien.at/wp-content/uploads/2013/12/VdB_Report2013_final_.pdf

Bundesministerium für Gesundheit (Federal Ministry of Health), Rahmen-Gesundheitsziele – Richtungsweisende Vorschläge für ein gesünderes Österreich (Framework Health Targets), Vienna 2012, http://www.bmg.gv.at/cms/home/attachments/5/5/3/CH1435/CMS1394200255665/bericht_rgz3_redaktionsteam.pdf

Bundesministerium für Wirtschaft, Familie und Jugend (Federal Ministry of Economy, Family and Youth), Bundesgesetz über die Steigerung der Energieeffizienz bei Unternehmen und dem Bund (Federal Energy Efficiency Act – EnEffG), Vienna 2013, http://www.ris.bka.gv.at/Dokumente/RegV REGV_COO_2026_100_2_855453/REGV_COO_2026_100_2_855453.pdf

Bürgerschaft der freien Hansestadt Hamburg (Parliament of the Free and Hanseatic City of Hamburg), Hamburger Masterplan Klimaschutz (Hamburg Master Plan for Climate Protection), 2013, http://www.coolbricks.eu/fileadmin/Redaktion/Dokumente/Publications/masterplan-klimaschutz.pdf

Cohen B., The Top 10 Smartest European Cities, 2012, http://www.fastcoexist.com/1680856/the-top-10-smartest-european-cities

Mercer LLC, Worldwide Quality of Living Survey, 2012, http://www.mercer.com/articles/quality-of-living-survey-report-2011

UN-HABITAT, State of the World's Cities 2012/13, Prosperity of Cities, 2012, http://www.unhabitat.org/pmss/listItemDetails.aspx?publicationID=3387

UN Intergovernmental Panel on Climate Change (IPCC), Stern Review on the Economics of Climate Change 2006, 2013

European Council, Conclusions of the Presidency, 7224/1/07, REV 1, CONCL 1, Brussels 2007, http://register.consilium.europa.eu/doc/srv?l=EN&f=ST 7224 2007 REV 1

Germanwatch, Klima und Klimaschutz in Hamburg: Informationen, Projekte und Adressen rund um die Klimaexpedition, 2007, http://germanwatch.org/klima/hh07.pdf

Siemens AG, Sustainable Urban Infrastructure, Ausgabe München – Wege in eine CO2-freie Zukunft, 2009, https://www.cee.siemens.com/web/at/de/corporate/portal/Nachhaltigkeit/Documents/munich_de.pdf

Institute for Ecological Economy Research, Energiekonzept 2020 – Energie für Berlin, 2011, https://www.berlin.de/imperia/md/content/sen-wirtschaft/energie/energiekonzept.pdf?start&ts=1302593601&file=energiekonzept.pdf

European Commission, Horizon 2020, The EU Framework Programme for Research and Innovation, http://ec.europa.eu/programmes/horizon2020/

European Commission, Energy Roadmap 2050, Brussels 2011, http://ec.europa.eu/energy/energy2020/roadmap/doc/roadmap2050_ia_20120430_en.pdf
Vision 2050, Roadmap for 2020 and Beyond, Action Plan for 2012-15, Smart Energy Demo – FIT 4 SET, Climate and Energy Fund, Vienna, Wien 2010-2012

Shannon Bouton et al., How to Make a City Great, McKinsey & Company

Ricky Burdett, Deyan Sudjic, Living in the Endless City. The Urban Age Project, New York, 2011, p.17

European Commission, Roadmap for moving to a competitive low carbon economy in 2050, 2011, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0112:FIN:EN:PDF

European Commission, Roadmap to a Single European Transport Area - Towards a competitive and resource efficient transport system, 2011, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0144:FIN:EN:PDF

European Parliament, Directive on the energy performance of buildings (recast), 2010, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:153:0013:0035:EN:PDF

Federal Ministry of Agriculture, Forestry, Environment and Water Management, Energiestrategie Österreich (Austrian Energy Strategy), 2010, http://www.energiestrategie.at/images/stories/pdf/longversion/energiestrategie_oesterreich.pdf

City of Vienna, Municipal Department 18 – Urban Development and Planning, Transport Master Plan Vienna 2003, Vienna 2003

City of Vienna, Municipal Department 18 – Urban Development and Planning, Urban Development Plan 2005, Vienna 2005; update: Urban Development Plan 2025, in preparation

City of Vienna, Municipal Department 20 – Energy Planning, Renewable Action Plan Vienna, Vienna 2011

City of Vienna, Municipal Department 20 (2013): Energiebericht der Stadt Wien (Energy Report of the City of Vienna); Energieverbrauch im öffentlichen Verkehr 1993-2011

City of Vienna, Municipal Department 20, Energie! voraus, Energiebericht der Stadt Wien (Energy Report of the City of Vienna); Vienna 2013 (2011 data), http://www.wien.gv.at/stadtentwicklung/energieplanung/pdf/energiebericht2011.pdf

City of Vienna, Municipal Department 48 – Waste Management, Street Cleaning and Vehicle Fleet, Wiener Abfallvermeidungsprogramm und Wiener Abfallwirtschaftsplan (Waste Avoidance Programme and Waste Management Plan, 2013-2018 planning period), 2012, http://www.wien.gv.at/umwelt/ma48/service/pdf/awp-avp-2013-2018.pdf

City of Vienna, Energiestrategie 2030 (Energy Strategy 2030), in preparation

City of Vienna, Municipal Department 22, Vienna Climate Protection Programme (KliP I), Vienna 1999, http://www.wien.gv.at/umwelt/klimaschutz/pdf/klip.pdf

City of Vienna, Vienna Climate Protection Programme, 2010-2020 update (KliP II), 2009, http://www.wien.gv.at/umwelt/klimaschutz/pdf/klip2-lang.pdf

City of Vienna, Urban Energy Efficiency Programme (SEP), Vienna 2006, http://www.wien.gv.at/stadtentwicklung/energieplanung/sep/

City of Vienna, Security of Supply Plan, in preparation

Wien Energie, Energy Flowchart for Vienna, Vienna, 2012, http://www.wien.gv.at/stadtentwicklung/energieplanung/pdf/energiebericht2010.pdf

Statistics Austria, Stock of Buildings and Dwellings, 2011 (also in English), http://www.statistik.at/web_de/statistiken/wohnen_und_gebaeude/bestand_an_gebaeuden_und_wohnungen/index.html

City of Vienna et al., Vienna 2020 Qualification Plan, 2013, http://www.waff.at/html/index.aspx?page_url=Qualifikationsplan&mid=358

Statistics Austria, Early School Leavers, 2012 (also in English), www.statistik.at/web_de/statistiken/bildung_und_kultur/formales_bildungswesen/fruehe_schulabgaenger/index.html

Statistics Austria, How's Austria? (also in English), 2012, www.statistik.at/web_de/statistiken/wie_gehts_oesterreich/

Statistics Austria, Time Use (also in English), 2008/09, www.statistik.at/web_de/statistiken/soziales/zeitverwendung/zeitwohlstand/index.html

United Nations Human Settlements Programme (UN-Habitat), State of the World's Cities 2012/2013 – Prosperity of Cities, 2013, http://www.unhabitat.org/pmss/listItemDetails.aspx?publicationID=3387

City of Vienna, Municipal Department 22 – Environmental Protection, Klimarelevanz der kommunalen Wiener Abfallwirtschaft (Climate Relevance of Vienna's Municipal Waste Management), Vienna 2012, http://www.wien.gv.at/umweltschutz/pool/pdf/klimarelevanz-2012.pdf

City of Vienna, Municipal Department 24 – Health Care and Social Welfare Planning, Vienna Social Welfare Report 2010, Vienna 2010, http://www.wien.gv.at/gesundheit/einrichtungen/planung/pdf/sozialbericht-2010.pdf

OECD Publishing, OECD Economic Surveys, 2013, http://www.keepeek.com/Digital-Asset-Management/oecd/economics/oecd-economic-surveys-austria-2013/austria-s-well-being-goes-beyond-gdp_eco_surveys-aut-2013-4-en#page6

City of Vienna, Municipal Department 18 – Urban Development and Planning, Wiener Lebensqualitätsstudien - Sozialwissenschaftliche Grundlagenforschung für Wien 2008 (Vienna Quality of Life Studies, Sociological Basic Research for Vienna 2008 – Summary Report), Vienna 2009, http://www.wien.gv.at/stadtentwicklung/studien/pdf/b008123.pdf

The Economist Intelligence Unit Limited, Global Liveability Report, 2013, www.eiu.com/site_info.asp?info_name=The_Global_Liveability_Report

Wiener Stadtwerke Holding AG (Vienna Public Utilities), Herausforderung Energiearmut und der Beitrag der Wiener Stadtwerke – Ursachen und Auswirkungen von Energiearmut und die Arbeitsweise der Wien Energie Ombudsstelle für soziale Härtefälle (The Challenge of Energy Poverty and the Contribution of Vienna Public Utilities), Vienna 2013

Further reading

CLUE, Climate Neutral Urban Districts in Europe, 2013, http://www.clue-project.eu/

EU-GUGLE, Sustainable Renovation Models for Smarter Cities, 2013, http://eu-gugle.eu/

Fachstelle 2000-Watt-Gesellschaft, 2000-Watt-Gesellschaft, 2013 (2,000-Watt Society), http://www.2000watt.ch/

Klimabündnis Österreich GmbH, Klimabündnis Österreich (Climate Alliance Austria), www.klimabuendnis.at/

The New Economics Foundation (NEF), Happy Planet Index HPI, 2013, www.happyplanetindex.org/

TRANSFORM, Transformation Agenda for Low Carbon Cities, 2013, http://urbantransform.eu/

TRANSFORM+, 2013, http://www.transform-plus.at/

IV. Process actors

Chief Executive Office of the City of Vienna – Steering Group for Smart City Wien:
Erich Hechtner (Chief Executive Office); chair
Dieter Groschopf (wiener_wohnen – Vienna Fund for Housing Construction and Urban Renewal)
Marc Hall (Vienna Public Utilities)
Veronika Haunold (TINA Vienna)
Gerhard Hirczi (VBA – Vienna Business Agency)
Ina Homeier (Municipal Department 18)
Brigitte Jilka (Chief Executive Office – Executive Group for Construction and Technology)
Stefan Leeb (Chief Executive Office – Communication)
Thomas Madreiter (Chief Executive Office – Executive Group for Construction and Technology)
Wolfgang Müller (Chief Executive Office – Executive Group for Organisation and Security)
Michael Stampfer (WWTF – Vienna Science and Technology Fund)
Ilse Stockinger (Vienna Public Utilities)
Oliver Stribl (Municipal Department 53)
Andreas Trisko (Municipal Department 18)
Angelika Winkler (Municipal Department 18)

Chief Executive Office of the City of Vienna – Smart City expert advisory board:
Brigitte Bach (AIT – Austrian Institute of Technology)
Robert Horbaty (ENCO, Zurich)
Andreas Knie (InnoZ, Berlin)
Verena Madner (Vienna University of Economics and Business)
Rudolf Scheuvens (Vienna University of Technology)

Project leader:
Ina Homeier (Municipal Department 18)

Process monitoring:
Donia Lasinger (WWTF – Vienna Science and Technology Fund)
Michael Stampfer (WWTF – Vienna Science and Technology Fund)

Thematic process monitoring (working group):
Elvira Franta (Municipal Department 53)
Stefan Gara (ETA)
Stephan Hartmann (Municipal Department 18)
Pia Hlava (Municipal Department 18)
Thomas Kreitmayer (Municipal Department 20)
Anna Mayerthaler (Vienna Public Utilities)
Kurt Mittringer (Municipal Department 18)
Manfred Mühlberger (ETA)
Volkmar Pamer (Municipal Department 21)
Eva Persy (Municipal Department 22)
Elke Plate (Senate of Berlin, exchange in the context of Leonardo Programme)
Volker Schaffler (TINA Vienna)
Christof Schremmer (ÖIR – Austrian Institute for Spatial Planning)
Katharina Söpper (Municipal Department 18)
Ilse Stockinger (Vienna Public Utilities)
Monika Sturm (ASCR)
Lara Tiede (Municipal Department 18)
Dominic Weiss (TINA Vienna)

Interview partners and process contributors:
Ursula Adam (WAFF – Vienna Employment Promotion Fund)
Thomas Bednar (Vienna University of Technology)
Agnes Berlakovich (Municipal Department 24)
Eva Buchinger (AIT – Austrian Institute of Technology)
Karin Büchl-Krammerstätter (Municipal Department 22)
Office of the Executive Policy Group for Integration, Women's Issues, Consumer Protection and Personnel
Michael Cerveny (TINA Vienna)
Gabriele Domschitz (Vienna Public Utilities)
Dieter Drexel (IV – Federation of Austrian Industries)
Andreas Eigenbauer (Chief Executive Office – Executive Office for Strategic Energy Policy)
Andrea Faast (Vienna Economic Chamber)
Susanne Fabian (Municipal Department 21)
Roland Fischer (IFF University of Klagenfurt)
Christine Fohler-Norek (Chief Executive Office – Executive Office for the Co-ordination of Climate Protection Measures)
Wolfgang Förster (Municipal Department 50)
Harald Frey (Vienna University of Technology)
Regine Gerike (University of Natural Resources and Life Sciences Vienna)
Michael Gervautz (Imagination Computer Services GesmbH)
Rudolf Giffinger (Vienna University of Technology)
Birgit Ginzler (TINA Vienna)
Dietmar Griebler (Municipal Department 5)
Simon Guy (University of Manchester)
Christian Habernig (Municipal Department 14)
Dieter Häusler (Municipal Department 18)
Rainer Hauswirth (Executive Policy Group for Education, Youth, Information and Sports)
Klemens Himpele (Municipal Department 23)
Roland Hierzinger (AEA – Austrian Energy Agency)
Josef Hochgerner (ZSI – Centre for Social Innovation)
Claus Hofer (Technology Agency of Vienna)
Johannes Höhrhan (IV – Federation of Austrian Industries)
Daniela Kain (Climate and Energy Fund)
Gerti Kappel (Vienna University of Technology)
Thomas Keller (Chief Executive Office – Executive Group for Construction and Technology)
Norbert Kettner (Vienna Tourist Board)
Gunda Kirchner (AEA – Austrian Energy Agency)
Roland König (Executive Policy Group for Public Health and Social Affairs)
Isabella Kossina (Vienna Public Utilities)
Leopold Liechtenstein (IV – Federation of Austrian Industries)
Daniel Löcker (Executive Policy Group for Cultural Affairs and Science)
Jutta Löffler (Executive Policy Group of Finance, Economic Affairs and Vienna Public Utilities)
Wolfgang Loibl (AIT – Austrian Institute of Technology)
Wolfgang Lutz (Vienna University of Economics and Business, IIASA)
Markus Marterbauer (Chamber of Labour)
Isabelle Meran-Waldstein (IV – Federation of Austrian Industries)
Johann Mittheisz (Chief Executive Office – Executive Group for Organisation and Security)
Gerhard Mörtl (Municipal Department 5)
Gerald Murauer (Siemens)
Hans-Martin Neumann (AIT – Austrian Institute of Technology)
Franz Oberndorfer (Head of the Executive Office of the Mayor)
Sigrid Oblak (Wien Holding)
Andrea Paukovits (Executive Policy Group for the Environment)
Ferdinand Pay (Executive Policy Group of Finance, Economic Affairs and Vienna Public Utilities)
Gabriele Payr (Vienna Public Utilities)
Margit Peter (Municipal Department 22)
Christian Pöhn (Municipal Department 39)
Wolfgang Polt (Joanneum Research)
Regina Prehofer (Vienna University of Economics and Business)
Clemens Rainer (denkstatt)
Michael Rauhofer (FTW – Telecommunications Research Center Vienna)
Beatrix Rauscher (Municipal Department 18)
Helmut Rechberger (Vienna University of Technology)
Herbert Ritter (Municipal Department 20)
Edith Rudy (Executive Policy Group for Education, Youth, Info mation and Sports)
Martin Russ (AustriaTech)
Ingolf Schädler (BMVIT – Federal Ministry for Transport, Innovation and Technology)

Gerhard Schmid (TINA Vienna)
Waltraud Schmid (TINA Vienna)
Martina Schmied (Head of the Executive Group for Personnel and Internal Auditing of the City of Vienna)
Andrea Schnattinger (Ombudsperson for Environmental Protection)
Josef Michael Schopf (Vienna University of Technology)
Monika Schuh (IV – Federation of Austrian Industries)
Hans-Günther Schwarz (BMVIT – Federal Ministry for Transport, Innovation and Technology)
Sigrid Semlitsch (Executive Policy Group of Finance, Economic Affairs and Vienna Public Utilities)
Mahshid Sotoudeh (ÖAW – Institute for Technology Assessment)
Christiane Spiel (Vienna University)
Christine Spieß (Project leader for aspern - Vienna's Urban Lakeside)
Sonja Starnberger (EIW – Energy Institute for Business)
Bernhard Steger (Executive Policy Group for Urban Planning, Traffic & Transport, Climate Protection, Energy and Public Participation)
Lukas Stockinger (TINA Vienna)
Gregor Stratil-Sauer (Municipal Department 18)
Monika Stumpf-Fekete (Vienna University)
Gregory Telepak (Municipal Department 18)
Gregor Thenius (AEA – Austrian Energy Agency)
Andrea Trattnig (Executive Policy Group for Education, Youth, Information and Sports)
Michaela Trojan (wiener_wohnen – Vienna Fund for Housing Construction and Urban Renewal)
Gabriele Trummer (Executive Policy Group for Public Health and Social Affairs)
Bernd Vogl (Municipal Department 20)
Andrea Wagner (Executive Policy Group for Housing, Housing Construction and Urban Renewal)
Wolfgang Wais (Vienna Public Utilities)
Christine Wanzenböck (Municipal Department 23)
Matthias Watzak-Helmer (TINA Vienna)
Norbert Weidinger (Chief Executive Office – Executive Group for Organisation and Security)
Peter Wieser (Municipal Department 23)
Gregor Wiltschko (raum & kommunikation GmbH)
Beate Wimmer-Puchinger (Vienna Commissioner for Women's Health)
Christian Wurm (Municipal Department 23)
Thomas Zemen (FTW – Telecommunications Research Center Vienna)
Axel Zuschmann (Ecker & Partner)

Forums:
Smart City Wien stakeholder forums (29 April 2013, 26 November 2013)
Press and Information Services (Municipal Department 53), workshop "Smart City Wien" for department heads (14 June 2013)
KliP workshop in Litschau (2-4 October 2013)
Intake workshop in the context of the EU-supported TRANSFORM project (17-18 September 2013)
IT company round (25 September 2013)
Mobility workshop (24 July 2013, 6 August 2013, 30 October 2013, 10 February 2014)
Thematic focus groups (26 August 2013, 28 August 2013, 30 September 2013, 21 October 2013)

V. Glossary

20-20-20 targets of the European Council for 2020
The European Union has set itself ambitious energy- and climate-policy objectives. Until 2020, these 20-20-20 targets obligate EU Member States to reduce greenhouse gas emissions by at least 20% vis-à-vis 1990, to improve their energy efficiency by 20% and to raise the share of EU energy consumption produced from renewable resources to 20%.

2,000-Watt Society of Zurich
With the initiative of the "2,000-Watt Society", the City of Zurich is leading the way to an energy-efficient, sustainable city of the future. Concretely, this means that Zurich:
· is committed to sustainable development,
· will reduce its energy consumption to 2,000 watt per person,
· will reduce its CO_2 production to one tonne per person and year by 2050,
· promotes renewable energy sources and energy efficiency,
· and will not renew its participation in nuclear power plants.
In this way, Zurich wants to contribute to climate protection. As a 2,000-watt society, Zurich is better prepared for times of scarce and costly energy resources.

Air pollutant emission
Air pollutant emission is the primarily anthropogenic release of certain substances into the atmosphere and stratosphere. This is mainly caused by traffic and industry. The increasing volume of air pollutant emissions directly impacts humans and the environment.

Ambient Assisted Living (AAL)
Age-appropriate assistance systems for a self-determined life comprise technical systems to support persons in need of assistance in their everyday lives. The objective lies in preserving and promoting independent lifestyles of persons including the very aged and in improving the quality of assistance, support and domestic services. The techniques and technologies are user-centred – i.e. focus on the client – and are integrated into this person's immediate living environment. Thus technology adapts to the client's needs and not vice versa.

Best point of service
To ensure a highly efficient healthcare system, this term describes demand-oriented treatment and care process management across all care levels. This safeguards that the correct service is delivered at the right moment and in the right place by providing optimum medical and nursing quality in the macroeconomically most cost-efficient manner.

Campus sites/Vienna Campus Model
The Vienna Campus Model concentrates kindergartens, schools and leisure education in one location. This co-operation is to ensure optimised use of all resources: all rooms of the building(s) are available for the teaching of all children. Joint projects make it possible to learn together as well as from each other. Smoother transitions between the different age groups and the combination of study and leisure activities are clear advantages of this model.

Climate neutrality
This term defines the reduction and compensation of greenhouse gas emissions as well as processes or certain conditions that do not lastingly change or impair the global climate. Often various products, too, are called or advertised as "climate-neutral". The most consistent form of climate-neutral energy use is the exploitation of greenhouse gas-free energy sources such as sunlight, wind and hydroelectricity.

Climate Protection Programme (KliP)
The current Climate Protection Programme (KliP II) of the City of Vienna, which was adopted by the City Council in 2009, defines climate protection targets to be attained by 2020 and stipulates a corresponding package of measures. The update of KliP I (1990) comprises a total of 385 individual measures in various fields of action: energy supply, energy use, mobility and urban structure, procurement, waste management, agriculture and forestry, nature conservation and public relations.

CLUE (Climate Neutral Urban Districts in Europe)
This INTERREG IV C project scheduled from January 2012 to December 2014 tackles the challenges sustainable cities and urban quarters aiming for an improved CO_2 balance are facing. A key objective lies in increasing local and regional capacities to develop political concepts for the implementation and assessment of new solutions and technologies. The consortium co-ordinated by the City of Stockholm comprises nine local and regional authorities as well as three universities from a total of nine European countries.

CO_2
Carbon dioxide, the most important greenhouse gas, mainly results from the incineration of the fossil fuels carbon, petroleum and natural gas. Throughout the Smart City Wien framework strategy, the term "CO_2" is used synonymously with "CO_2 equivalents" for reasons of simplicity.

CO_2 equivalent
The greenhouse gas effect of different greenhouse gases, e.g. methane, nitrous oxide (laughing gas) or fluorinated (F) gases, varies. The CO_2 equivalent value describes the greenhouse gas potential of a gas across an observation period of usually 100 years as compared to CO_2. The quantity in tonnes of the gas in question thus emitted is multiplied by the factor by which the greenhouse gas effect of the gas exceeds that of CO_2. The outcome is the emission volume in tonnes of CO_2 equivalents. Due to the uniform presentation of the environmental effects thus caused, the emissions of different greenhouse gases can be added up. In Vienna, 94% of all greenhouse gas emissions (expressed in CO_2 equivalents) in 2009 were due to CO_2; methane, laughing gas and F gases accounted for only 6%.

Cogeneration
Cogeneration (or combined heat and power – CHP) is the simultaneous production of power and heat, with power normally used directly to generate electric energy. As a rule, the heat generated is used for heating purposes by either feeding it into a heating grid or directly on-site. The combined use of power and heat results in high overall efficiency, which entails primary energy savings. In Austria, a great part of thermal power production is done in CHP plants.

Doctorate course URBEM-DK
Wiener Stadtwerke Holding AG (Vienna Public Utilities) and Vienna University of Technology have jointly instituted a doctorate course entitled "Urban Energy and Mobility Systems" (URBEM-DK). The goal is the research and development of scenarios for the path to a "sustainable, supply-secure, affordable and liveable city", using the example of Vienna with an integrated and interdisciplinary approach. Starting with the 2013/2014 winter semester, 10 university graduates will receive a three-year position at Vienna University of Technology to enable them to complete their doctoral theses.

Domestic traffic/transport
This term defines traffic and transport within the borders of a state or confederation of states.

Drinking water power plant
The principle of drinking water power plants is based on the exploitation of the difference in elevation between spring and reservoir and the water quickly flowing through the nozzle to generate electricity. Thus the drinking water turbine serves the function of a pressure-reducing valve and produces additional electric energy from the desired pressure reduction. Before the

construction of drinking water power plants, the high gravitational pressure in Vienna had to be reduced mechanically (hydraulic surge control valves = Clayton valves) in order to obtain optimum pressure in the pipe network. This method has no effect whatsoever on drinking water quality.

Early school leavers
According to the definition of the EU, early school leavers are people aged 18-24 years who are no longer in education or training and have no higher secondary education (no attainments above ISCED level 3c). Adapted to the Austrian situation, young people do not count as early school leavers if they, before leaving the educational system, have at least completed an apprenticeship or a multiyear vocational school, while the completion of a lower secondary school, a polytechnic secondary school or a one-year vocational middle school (e.g. one-year domestic-science schools) is not sufficient.

e-car sharing
On an average, private motorised vehicles are used only one hour per day and are parked, mostly in public space, for the remaining 23. To promote a more efficient use of cars and parking space, several initiatives have sprung up in Austria over the past few years to offer either commercial or privately organised car sharing services. At the moment, car sharing services are mainly located in cities; a comprehensive range of different forms of car sharing and appropriate vehicles will certainly impact the future of mobility. e-car sharing describes car sharing services based on electrically powered vehicles for individual traffic.

EcoBusinessPlan Vienna
EcoBusinessPlan Vienna is the environmental service package of the City of Vienna for enterprises in the Austrian capital. Established in 1998 by Municipal Department 22 – Environmental Protection, EcoBusinessPlan Vienna supports companies in their implementation of eco-relevant measures in day-to-day work and contributes to the reduction of operating costs.

Eco-friendly means of transport
These are all means of transport whose pollutant and noise emissions and land consumption do not damage the environment: walking, cycling (including public bike rental such as the "Citybike" system), public transport (suburban trains, Underground, trams, buses) and, in the wider sense, taxis, car sharing or car pools; combinations of these means of transport are also included.

e-government
"e-government" (or electronic government) is synonymous with modern, efficient administrative methods. Concretely, it stands for the use of information and communication technology (ICT) by public administrations in combination with organisational changes and new skills in order to improve public services and democratic processes and facilitate the design and delivery of public policies.

Electromobility
This term defines the use of electric-powered vehicles to meet different individual mobility needs and is hence a cross-cutting issue that extends across traffic/transport, infrastructure, technology, energy and environment.

Energy Efficiency Act of 2013
The Energy Efficiency Act of the Federal Republic of Austria was adopted by the Council of Ministers but failed to obtain the required two-thirds majority in Parliament. It comprised the following key aspects:

- increase of the share of renewable energy sources in total energy consumption, power generation and the traffic/transport sector
- saving energy by stepping up cogeneration, increasing energy intensity, energy checks for households and thermal rehabilitation
- active climate protection by means of energy, environmental, location and transport policies

Energy Roadmap 2050 of the European Commission
To attain the target of reducing emissions by over 80% by 2050, the European Commission presented the Energy Roadmap 2050 in December 2011. The Roadmap explains how this target can be achieved without impairing security of energy supply and competitiveness. Starting from the analysis of several scenarios, the effects of a CO_2-free energy system and the political framework necessary to attain this goal are described in this paper. Member States should take the necessary energy policy decisions on this basis and be able to create a stable business environment for private investments.

EU-GUGLE
It is the objective of the EU-GUGLE project, conducted in six pilot cities, to demonstrate the feasibility of nearly-zero energy building renovation models in view of triggering large-scale, Europe-wide replication in smart cities and communities by 2020. Taking on the challenge of sustainable renovation in urban areas, the cities of Vienna (AT), Aachen (DE), Milan (IT), Sestao (ES), Tampere (FI) and Bratislava (SK) are committed to renovating a total of 226,000 sq m of living space during the five years of the project, with the objective of achieving 40 to 80% of primary energy savings per pilot district while increasing the share of renewable energy sources by 25% by 2018. Gothenburg (SE) and Gaziantep (TR) will take part in the five-year project as associated cities and will be expected to start smart renovation activities during the project's lifetime.

European Union Emissions Trading System (EU ETS)
EU ETS, the EU-wide greenhouse gas emissions trading system, has been in place since 2005. A certificate must be obtained for every tonne of CO_2 emitted. Since the number of certificates is fixed across the EU and reduced year by year, this system will in the long term lead to the reduction of emissions without stipulating a specific emission target for individual market players. The set reduction in the number of available certificates has engendered a functioning market in certificates. The CO_2 price created by supply and demand reveals which measures for CO_2 reduction are more efficient than paying for certificates.

EU White Paper on Transport
White Papers published by the European Commission contain proposals for common action in a given area. In part, they tie in with Green Papers, which initiate a consultation process at the European level. A White Paper can only evolve into a concrete action programme if it meets with a positive response by the Council. The White Paper "Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system" (White Paper on Transport) was submitted in 2011. Against the background of the EU climate and energy efficiency targets, one of its main objectives is the reduction of greenhouse gas emissions due to transport by approx. 60% by 2050 (compared to 1990 values).

Forerunner cities
This term describes a group of cities that are considered European or worldwide pioneers with regard to a (more or less specific) area of application.

Gender budgeting
The basic concept behind gender budgeting is to analyse the effects of administrative actions and budget policy in particular with regard to the distribution and earmarking of public funds for women and men and, if necessary, to take corrective measures.

Gender gaps
in sociology and political economics, gender gaps define differences between women and men.

Gender mainstreaming
Gender mainstreaming is a concept to denote a gender-equitable society with equal social structures, starting points and framework conditions. Gender equality is no longer a peripheral issue but a matter-of-fact element of all processes and measures. Women and men are thus not viewed as a uniform, homogeneous group; rather, their respective social, ethnic or age-related differences are taken account of.

Gentle urban renewal
"Gentle urban renewal" is a successful Viennese model. The primary aspects of gentle urban renewal, which was introduced in Vienna nearly 40 years ago, have always been affordable housing based on financial incentives, social sustainability and a comprehensive strategy of renewal of the "core city". The positive trends of the architectural, social and economic development of Vienna's urban renewal areas – above all within the perimeter of the suburban boulevard, the Gürtel – are evident. Targeted rehabilitation and new construction measures upgrade and rejuvenate neighbourhoods and quarters. The experts of Gebietsbetreuung Stadterneuerung (Area Management Office, GB*) serve as the hub for all renewal processes, but also for all questions relating to housing, the housing environment and good neighbourly relations within neighbourhoods. As of today, 320,000 dwellings have been rehabilitated through gentle urban renewal.

Governance
In our time, traditional forms of government are no longer sufficient to ensure effective use of the volume and variety of information as well as of the creativity of society. For this reason, administrations open up and complement hierarchic types of control with co-operative action together with other social groups, e.g. citizens, business circles and other local or regional authorities.

Gross final energy
This term defines the energy remaining after the loss-entailing conversion of raw energy. This energy may take the form of e.g. electricity, district heat, process gases or fossil fuels. Gross final energy is defined as final energy plus network distribution losses plus the energy demand of power plants.

Gross regional product
The gross regional product is the regional equivalent of the gross domestic product (GDP). It is usually calculated nominally (in market prices of the year in question) and serves on the one hand to analyse regional economic development and on the other hand to establish comparisons with other federal provinces. To calculate the gross regional product, the national subsidies and taxes on products are allocated to the individual federal provinces in keeping with their shares in regional gross value creation, resulting in the GRP.

Grätzl
"Grätzl" is a Viennese dialect word for an urban neighbourhood that usually comprises a few blocks and is considered the smallest urban "unit". Such neighbourhoods are defined by their differences as compared to adjoining zones or a special, unique atmosphere; there are no official delimitations or territorial boundaries.

Gründerzeit
This term encompasses the period from post-1850 to 1914, when the Austro-Hungarian monarchy was economically modernised over several phases, leading to the industrialisation of numerous regions and the construction of railway lines throughout the Empire. Many buildings for educational and cultural purposes were erected in this era; cityscapes and city dimensions changed quickly over a short period.

Horizon 2020
Is an EU subsidy programme for research and innovation that is organised by the European Commission. Close to Euro 80 billion are earmarked by the EU from 2014 to 2020 for research and innovation purposes. The funding and subsidy instruments target everything from basic research to innovative product development. Individual researchers, enterprises and collaborations between science and industry are crucial target groups of Horizon 2020. Excellent research, highly competitive research and industry locations, more innovative products and services as well as simplified regulations are the essential objectives pursued by this new programme for research and innovation.

Housing costs
This stands for the average monthly costs of a dwelling (rent or loan repayments/annuities for owner-occupied flats plus service charges as well as all heating, garage/parking lot costs, if any; always including value-added tax).

ICT Strategy
The ICT Strategy 2007 is aligned alongside the Business Strategy of the City of Vienna and wants to optimise the economic value added of ICT services and products for the implementation of the overall strategy. The ICT Strategy describes the key objectives, measures and processes to effectively and efficiently support the attainment of business goals on the part of the City of Vienna by means of ICT. It is the strategic guideline for ICT use and diversification by all municipal departments.

Information and communication technology (ICT)
Information and communication technology may be used for three types of applications: the transmission of information through space (from point A to point B, i.e. "communication"), the transmission of information in time (from moment 1 to moment 2, i.e. "storage"), and the structured transformation of information in space and time, based on an algorithm (i.e. "computer calculations").

Objectives of Vienna's ICT strategy
In the context of administrative reform, e-government enables citizens and business circles to use services of the City of Vienna simply and easily by drawing on state-of-the-art information and communication technology (ICT).
Towards this purpose, "ICT" embodies the following three principles:
"I" stands for information: the Virtual Office of the City of Vienna provides support through customised information.
"C" stands for communication: citizens and business circles can communicate with the City of Vienna via contact and feedback forms. Contact forms can be found at the bottom of every wien.at web page. Feedback forms are available on every information page of the Virtual Office of the City of Vienna.
"T" stands for transaction: the Virtual Office also offers the possibility of submitting online applications.

Innovation
Innovations are planned and controlled changes or new developments of a social system as a result of new ideas or techniques. Moreover, the term encompasses the implementation of novel, progressive solutions to specific problems, the introduction of new products or the application of novel procedures. The majority of innovations results from novel combinations of previously already known procedures, products or systems.

Lifelong learning
The document "LLL:2020 – Strategie zum lebensbegleitenden Lernen in Österreich" (Strategy for Lifelong Learning in Austria) was adopted by the Austrian federal government in July 2011. Ten action lines detail visions, the status quo and individual projects and measures. The action lines emphasise that lifelong learning comprises a great variety of forms of learning throughout all life phases. In addition to aspects of elementary and youth education, informal learning processes and learning in old age are likewise covered.

Life sciences
Life sciences comprise those fields of science that involve the scientific study of living organisms, such as bioscience, medicine and many others.

Lighthouse project
An exemplary project that in addition to its underlying purpose also entails a signal effect for numerous follow-up ventures. This is contingent on the success and public profile of the project.

Living lab
Living labs embody a research concept – often collocated in a user-centred, territorial context (e.g. a city or region) – that examines the integration of research and innovative processes within the scope of a public-private partnership.

Memorandum of understanding (concluded between the City of Vienna and the Federal Ministry for Transport, Innovation and Technology – BMVIT)
Mayor Michael Häupl and Federal Minister Doris Bures signed a Smart City memorandum of understanding (MOU) concluded between Vienna and the Federal Republic of Austria represented by BMVIT. The objective lies in initiating projects via a joint steering group and to obtain funding at a European level.

Metropolitan region
A metropolitan region is an area comprising at least one big city of international importance, serving a variety of high-level centralised functions, and the region surrounding it. As a rule, such a region encompasses, in addition to the urban agglomeration, further core settlements as well as rural zones that are closely linked to the regional centre, e.g. by commuter traffic flows. As central nuclei of co-operation and competition, metropolitan regions are hubs of international networking. Metropolitan regions may have up to 20 million inhabitants, may be transboundary, are politically defined in a variety of ways and usually lack institutions of their own.

Multimodality
In the traffic sector, multimodality applies if more than one means of transport is or can be used to cover a distance or to transport goods; for travelling from and to work, a person may e.g. use either a bicycle or a car, or merchandise may be brought to its destination by a combination of navigation and railway.

Multiple use
Purpose of the strategic project "einfach-mehrfach" (simple-multiple) of the City of Vienna, which was commissioned in 1998. The concept of multiple use mainly refers to municipal infrastructure facilities like schools, above all with regard to the opening of green/open spaces (or sports grounds) for the neighbourhood.

Noise emissions
Noise emissions are sounds that impair the physical, psychological and social wellbeing of humans. The perception of noise varies greatly from person to person, and noise exposure can have manifold physical and psychological effects. Noise effects are grouped into effects on hearing and effects on the organism as a whole.

ÖkoKauf Wien"
In 1998, the City of Vienna initiated the "ÖkoKauf Wien" programme to promote climate protection. The objective lies in aligning the procurement of goods, products and services in all fields of municipal administration more strongly with ecological principles. According to a decree issued by the Chief Executive Director of the City of Vienna, all outcomes of "ÖkoKauf Wien" (catalogues of criteria, position papers, studies, sample folders) must be bindingly applied.

Open data
Offers possibilities of data use at many different levels. Enterprises and citizens can develop new applications and services based on the available data material. The participation of citizens in political decision-making processes is likewise fostered by open data. Science and research benefit equally from simplified data exchange. https://open.wien.at/site/open-data/

Open government
The term "open government data" describes the idea that public data collected by administrative bodies should be rendered freely accessible. These data should be provided to the population in machine-readable format to permit their automated dissemination. Open standards for interfaces and software ensure greater transparency, participation and collaboration. In addition to technical interfaces, the administration must also create a legal framework. Examples of such public data include geodata, traffic data, environmental data, budget data or statistical data. Personal data are not rendered public.

Participation
In this context, the term defines the participation of a person or group in decision-making processes or courses of action taking place within superordinate structures or organisations.

PGO (Planungsgemeinschaft Ost)
This is a joint organisation of the Federal Provinces of Vienna, Lower Austria and Burgenland to consult and agree on relevant spatial planning issues. Its scope of work comprises the development of joint spatial planning objectives, the thematic co-ordination and scheduling of plans impacting regional space, the representation of joint interests and the implementation of research projects of importance for regional planning in the three federal provinces. In addition to a political decision-making body and a body co-ordinating the three provincial administrations, PGO also operates a joint office to implement its work programme.

Pre-commercial procurement
Pre-commercial procurement (PCP) is an attractive (because unbureaucratic and flexible) instrument for enterprises to competitively develop new ideas and solutions for public procurement purposes. Public bodies benefit from this, as their problems are solved; moreover, PCP serves as a viable instrument for demand-oriented innovation policy.

Primary energy
This term defines the raw energy that has not yet been converted into useful energy. Final energy results from the loss-entailing conversion of raw energy. Raw energy may be available in many different forms, e.g. sunlight, wind, biomass, fossil materials or nuclear power.

Protection zones
It is not only that one third of the municipal territory of Vienna is covered by protection zones – rather, the city also boasts part of a national park, nature reserves, protected landscapes and landscape elements, ecological development zones, protected biotopes, the protected zone of the Vienna Green Belt and parkland. Moreover, parts of Vienna's westernmost districts were declared the Vienna Woods Biosphere Park. The landscape of the Vienna Woods has been under protection since 1905 on the basis of the Vienna Green Belt Decision as well as due to a protection zone category assigned to it under the Building Code for Vienna (Sww category), which practically equals a total construction ban, and the Nature Conservation Act. STEP05 shows these restrictions in its mission statement for green spaces in Vienna.

Quality of living
Quality of living is the basically subjective perception of an individual regarding his or her position in life in relation to the culture and value systems of his or her life environment and focuses on that person's goals, expectations, standards and concerns. However, key traits of quality of living also involve political, social, economic and environmental aspects. This is compounded by factors such as personal security, health, education and transport options as well as other public services.

Rainwater management
Rainwater management offers the possibility of retaining water precipitated on built-up and sealed areas within the natural water cycle and thus relieving sewers. This can have a positive effect on temperature, air quality and climate.

Renewable Action Plan (RAP_VIE)
The Plan aims to strengthen the use of regenerative energy sources in and for Vienna. This is important for both existing and new buildings, but also plays a key role in mobility. The major potentials in urban areas lie in the generation of power and heat from solar energy, in the use of ambient and waste heat and, in the long term, in the tapping of deep geothermics. In the field of traffic and transport, the switch to alternative powertrain types, above all electromobility from renewable sources, is most crucial alongside the promotion of eco-friendly means of transport. Commuter and goods traffic, too, play a significant role in this context. Wind and hydropower supply Vienna with energy mainly harnessed in the region or in other federal provinces. The Action Plan defines measures to cover these areas.

Renewable energy sources
This term defines energy sources that are constantly renewed or replenished and hence are permanently available. Renewable energy sources e.g. include sunlight, wind, hydroelectricity, biomass, geothermal heat and waste. Renewable energy sources are CO_2-neutral; thus their use does not negatively impact the climate. Even with renewable energy sources, sustainable use is only safeguarded if the consumption rate does not exceed the renewal rate.

Research and development spending
R&D (= research and (experimental) development) stands for creative activities conducted systematically by using scientific methods in order to improve the state of knowledge and develop new applications of this knowledge. Statistics Austria regularly conducts surveys on R&D in Austria by interviewing enterprises and university institutions as well as public institutions engaging in R&D (Federal Republic, federal provinces, municipalities, chambers, …) about their research and development activities and spending on staff, ongoing administrative expenses and investments directly linked to these activities.

Research, Technology and Innovation Strategy (or Research and Innovation Strategy "Innovative Vienna 2020")
With this strategy, the City of Vienna has set itself the task to join the ranks of the leading European metropolises of research, technology and innovation (RTI). Vienna's RTI strategy adopted in 2007 comprises five action areas that supply answers to five central challenges for Vienna as an RTI location.
Human resources: smart brains for Vienna
Thematic focuses: visible and relevant
Research and the city: communication, learning and publicity
A breeding ground of research and innovation: enabling new developments
A research and innovation location in Europe: Vienna as an international network hub

Resilience
This word describes the ability of a system to deal with change. Resilience strategies may be inter alia preventive (providently accumulated resistance to withstand negative external influences), adaptive (ability to return quickly to the original situation) or innovative (active use of advantages resulting from changing environmental conditions).

Resource preservation
The use of natural resources and the competition for scarce resources such as freshwater, land and raw materials (see Resources below) are intensifying all over the world. Thus a key challenge lies in ensuring the sustainable and conservation-oriented use of limited resources.

Resources
Resources are natural stores or sources of something needed for a specific purpose, in particular for feeding humans and economic production.

Security of supply
Security of supply is the steady and long-term safeguarding of basic human needs. These include water supply, clean air, provision for old age, basic income, healthcare, etc. In energy policy, this term describes environmentally compatible, efficient energy supply. The umbrella term "security of supply" also incorporates the quality of supply, which in general is structured into reliability of supply, voltage quality, operative security of supply and commercial quality (services).

Services of general interest
This term describes the judicious and reliable provision of public goods and services that entail a special responsibility for the community at large. These comprise e.g. energy, water, waste and wastewater disposal, education, culture, medical services or public transport. They are characterised by a guarantee of equal, universal access, security and continuity of provision and – where the State is the provider – by democratic scrutiny and public accountability. In Austria, the definition and type of provision of services of general interest typically entail a high degree of municipal autonomy.

Showcase
This defines a project or programme that serves to convey innovative aspects in a clearcut, exemplary and instructive manner.

Smart City Wien stakeholder forum
The core element of the Smart City Wien initiative is a long-term stakeholder process, in whose context all interest groups within and outside the municipal administration are organised in both general purpose and thematically specialised advisory groups.

SMILE mobility card
SMILE (Smart Mobility Info and Ticketing System Leading the Way for Effective E-Mobility Services) is the prototype of an Austria-wide multimodal mobility platform that is to provide all public and individual mobility services for customers. It is the objective of this joint research project of Vienna Public Utilities, Wiener Linien and Austrian Federal Railways (ÖBB) to develop a platform that shows the various possibilities of travelling from A to B. With open and uniform interfaces, the mobility platform is designed to enable other mobility service providers (e-car sharing, e-bike rental, car parks, charging stations, etc.) and other, likeminded projects to easily hook up with this system. The cooperation of the two biggest mobility service providers ÖBB and Wiener Linien lays the ground for a future, nationwide smart mobility platform.

Social housing
This term defines state-subsidised housing construction in particular for social groups that are unable to cover their housing needs at market conditions. With the Viennese model of housing subsidisation, the City of Vienna ensures the creation and safeguarding of high-quality, demand-oriented and environmentally compatible dwellings. Manifold measures moreover guarantee that housing will remain affordable for persons with medium or lower incomes.

Social inclusion
The idea of inclusion aims for a society steeped in human rights, appreciation and respect, which grants the same, full rights of individual development and participation to all its members, irrespective of abilities, gender, sexual orientation, social or ethnic background, etc. Inclusion strives to eliminate all processes of exclusion and fights marginalisation risks such as poverty and discrimination.

Stakeholders
These are all persons or groups that have a justified, vested interest in the course or outcome of a process of project.

Sustainability
Sustainable use of the environment equals the conservation of the natural capital (like groundwater, habitats and rare species). It stipulates that the use of renewable resources and energy carriers should not exceed their replenishment. Non-renewable resources should not be consumed faster than they can be substituted by enduring, renewable ones. Sustainable use of the environment also means that pollutant emission must not exceed the ability of air, water and soil to bind and break down these pollutants. The needs of our time must be met so that later generations will not have to suffer the consequences.

Testimonial
This word drawn from the advertising industry describes concrete praise and/or satisfaction for/with a product, service, idea, project or institution on the part of a person known and respected as competent in the field in question. It can also mean "experience report".

TRANSFORM
"Transformation Agenda for Low Carbon Cities" is a project under the Seventh Framework Programme for Research of the European Commission with a length of 30 months and a financial volume of approx. Euro 7.5 million. By means of implementation-oriented strategy development, the TRANSFORM project is to support cities in their transformation into smart cities. For this purpose, comprehensive strategies and transformation agendas for the entire city are developed; on the basis of selected urban quarters of the partner cities – called smart urban labs – these are embedded in a context corresponding to the special requirements of these neighbourhoods to derive individually fine-tuned, concrete implementation plans.

Transform+
The "Transform+" project is subsidised by the Climate and Energy Fund of the Austrian Research Promotion Agency and is scheduled to run from March 2013 to February 2016. It is the objective of the project to prepare and support the Austrian contributions for the EU project TRANSFORM thematically and operatively. Moreover, the outcomes are edited for an exchange of experience with the partner cities. This includes the organisation of a citywide stakeholder process, the development and adaptation of the database for decision-supporting models and the work on implementation plans and pilot projects for the model urban quarters "Liesing-Gross Erlaa" and "aspern - Vienna's Urban Lakeside" (smart urban labs). In this context, strategy development proceeds in a targeted fashion for planning processes that concern urban planning, housing construction, transport and energy system development. International exchange and feedbacks within the TRANSFORM project generate process impulses such as the goal-focused networking of relevant actors, the concentration of know-how and experience, the joint formulation of sustainable strategies and the tapping of interdisciplinarity and synergy effects.

Transport Master Plan (MPV)
The Transport Master Plan Vienna 2003 pays great attention to Vienna's role in a new Europe: Vienna as a TEN node, Vienna as a potential main beneficiary of EU enlargement, Vienna as a metropolis of technology and an economic hub. The quality of living in the Austrian capital is to be kept as high as it is, and this can be strongly impacted by smart and sustainable mobility. For this reason, eco-friendly means of transport (public transport, cycling, walking) are further promoted, resulting in a modern and future-oriented transport concept for the next two decades.

Umbrella strategy
An umbrella strategy is a superordinate strategy based on different individual strategies and/or combining them under one heading (the "umbrella").

Urban Development Plan 2025 (STEP2025)
The Urban Development Plan serves as a guideline for all matters of the city which have spatial effects and therefore need to be co-ordinated. It is prepared every ten years. In 2014 the Vienna City Council enacted the current version: STEP2025. The predicted population growth as well as new ways of co-operation and participation bring about novel urban development tasks: the focus lies on the mobility system, green infrastructure, and building compact and attractive urban quarters that combine space for living, work, and leisure. STEP2025 will be supported and substantiated by subsequent strategies for implementation.

Urban Energy Efficiency Programme (SEP)
SEP I was adopted in 2006 by the Vienna City Council and comprises guidelines for consumption-side energy policy until 2015. Municipal Department 27 was charged with drafting the programme, which describes two scenarios (business-as-usual and energy-saving scenarios) and links them to an analysis of the energy situation and background data. The energy-saving scenario assumes an energy consumption growth of 7% between 2003 and 2015. Another part of the programme includes numerous measures in the fields of buildings, rehabilitation, electric appliances and plants, heating systems and air conditioners, illumination and transport – all this concerning private individuals, businesses, industry and the public sector. An update of the programme as SEP II is currently under discussion.

Urban mining
Urban mining refers to the identification of anthropogenic "deposits", the quantification of the secondary raw materials contained in them, feasibility considerations against the background of available technical recovery options and current and predicted revenue as well as, last but not least, the commercial treatment and recycling of the reusable materials identified and the integral management of anthropogenic deposits. In this context, humans are not only viewed as consumers but also as producers of valuable resources.

Urban technologies
This term covers products and services in such thematic areas as telecommunications, transport and mobility, energy and environment, supply and disposal, construction and housing as well as conservation. Urban technologies are a focal point of Vienna's urban research activities and of the Technology Agency of the City of Vienna.

Vienna Charter
The Vienna Charter was launched in March 2012 and constituted a hitherto unique citizen participation project in Europe. Pursuing the objective of jointly shaping good neighbourly relations in Vienna, a total of 8,500 Viennese citizens conducted discussions in 651 groups to voice their ideas. The Charter process was initiated by the City of Vienna and implemented together with over 325 partner organisations. The preamble to the Charter reads, "Vienna is home: for women and men, for young and old, for those who were born here and for those who moved to the city, for people with different worldviews, beliefs and needs. To get along well, we need to respect each other. Respect means accepting other people the way they are – as we ourselves wish to be accepted and respected. Human rights are our common basis."

Vienna Spring Water Mains
Vienna's mountain spring water is tapped in the Styrian and Lower Austrian Alps and transported to Vienna by means of the First and Second Spring Water Mains without the use of pumping stations, based solely on gravity. The drinking water of the Second Vienna Spring Water Main originates in the Styrian Hochschwab massif and covers a distance of approx. 180 kilometres before reaching the Austrian capital.

WienWin
The “WienWin” initiative is a database for innovative products and services of Viennese entrepreneurs. This clearly structured overview of Vienna’s potential for innovation offers municipal project managers and enterprises of the City of Vienna an optimum understanding of the situation as well as detailed information about innovations from the Austrian capital. This innovation pool should be seen as a network of future-oriented projects of the City of Vienna.

Work/life balance
Working life and private life should present an even balance without interfering with each other; ideally, they should be mutually supportive.

Zero-energy building standard
This standard refers to the construction of exclusively (nearly) zero-energy buildings, which was imposed by the EU for the construction of new public buildings starting from 2018 and for all buildings starting from 2020. Zero-energy buildings are characterised by very low energy requirements, with a considerable share of these requirements to be produced by means of renewable energy sources on-site (or close by). This standard is to be met either through stricter stipulations for the building shell (very low heating demand) or by means of the intensified use of renewable energy sources.

Imprint

Owner and publisher: Vienna City Administration

Project leader: Ina Homeier, Municipal Department 18,
City Development and Planning

Text and editing: Municipal Department 18, WWTF,
TINA Vienna and ÖIR
Content based on expert co-operation (see pages 100 to 101)

Technical co-ordination: TINA Vienna, Municipal Department 18

Layout and graphics: Beton | typischbeton.com
Text set in Theinhardt and Fishmonger

Photos people: Julian Mullan
Photos city: Archive of the City of Vienna, Municipal Department 18

Production: Lindenau Productions

Translation: Sigrid Szabó

For further information on the Smart City Wien Initiative:
www.smartcitywien.at
www.smartcity.wien

Printed on paper of the EcoBuy Vienna portfolio.

The translation is based on the German Smart City Wien
Framework Strategy, which was adopted by the
Vienna City Council on 25 June 2014.

ISBN 978-3-902576-91-0

Published in final edited form as:
Subst Use Misuse. 2003 December ; 38(14): 2049–2063.

Explaining the Geographical Variation of HIV Among Injection Drug Users in the United States

D. Ciccarone, M.D., M.P.H.[1,2,*] and **P. Bourgois, Ph.D.**[2]

[1]*Urban Health Study, Department of Family and Community Medicine, and* [2]*Department of Anthropology, History and Social Medicine, University of California, San Francisco, California, USA*

Abstract

Distinct physical and chemical types of street heroin exist worldwide, but their impact on behavior and disease acquisition is not well understood or documented. This article presents a hypothesis to explain the unequal diffusion of HIV among injection drug users in the United States by examining the distribution and use of one type of heroin—"Mexican black tar." Drawing on ethnographic, clinical, epidemiological, and laboratory data, we suggest that the chemical properties of black tar heroin promote the following safer injection practices: (1) the rinsing of syringes with water to prevent clogging; (2) the heating of cookers to promote dissolution; and (3) a rapid transition from venous injection to subcutaneous or intramuscular injections.

Keywords

Black tar heroin; Geography variation; HIV; Risk factors; Heroin type

---

Most of the literature regarding illicit heroin use regards the substance as uniform with few variations in either chemistry or in patterns of administration by injection drug users (IDUs). Phenotypical and chemical variation in street heroin types may, however, have had a substantial role in shaping the epidemic among IDUs in the United States. Using multiple primary and secondary sources of data we have developed a hypothesis, that use of Mexican-derived heroin, commonly referred to as "black tar" heroin (BTH), may have retarded the spread of HIV in those U.S. states where it predominates.

Geographical Risk of HIV Infection

In the United States, multi-site epidemiological studies of IDUs consistently identify location in a high-seroprevalence city (i.e., large metropolitan cities in the northeastern United States) as an independent risk factor for HIV (Friedman et al., 1995; Kral et al., 1998; Montoya and Atkinson, 1996). Geographical location, consequently, has emerged as a proxy variable for behavioral, environmental, historical, and/or structural factors that still need to be explained. The importance of geography as a risk for HIV among IDUs is more compelling given the relatively even geographical pattern of diffusion of HIV among men who have sex with men (MSM) (Holmberg, 1996). In the early 1990s in New York City, for example, the estimated HIV prevalence among MSM was 29.2% while that for IDUs was 41% (Holmberg, 1996). In contrast, in a comparable metropolitan area, Los Angeles, the MSM HIV prevalence was 22.6% while the IDU HIV prevalence was only 3.8%. In the center of the country in Denver the HIV

---

*Correspondence: D.Ciccarone, M.D., M.P.H., Urban Health Study, Department of Family and Community Medicine and Department of Anthropology, History, and Social Medicine, University of California, 3180 18th St., Suite 302, San Francisco, CA 94110, USA; Fax: (415) 476-3406; E-mail: ciccaron@fcm.ucsf.edu..

prevalence among MSM remained high at 25%, while the IDU population was only 3.8% HIV positive.

Initially, the HIV epidemic in the United States was expected to radiate out of its geographical epicenters, the largest being New York City (Des Jarlais et al., 1989). Public health researchers were concerned that the MSM-driven epidemic would crossover into the IDU population in the South, West, and Midwest as had already happened in the Northeast by the mid 1980s (Hahn et al., 1989; Lange et al., 1988). By the early 1990s, this crossover of HIV had taken place in the South, but it never occurred in epidemic proportions west of the Mississippi (Holmberg, 1996). Researchers have offered several hypotheses to explain the failure of HIV to spread in epidemic proportions to IDUs in the western United States. These have included: (1) differential interfaces between the MSM and IDU communities across cities (Bourgois, 1998b; Chaisson et al., 1987); (2) the absence of the phenomenon of "shooting galleries" on the West Coast early in the epidemic (Watters, 1989); (3) differential public health responses and law enforcement practices affecting overall availability of sterile injection equipment (Bourgois, 1998b; Des Jarlais et al., 1995; Friedman et al., 2001; Koester, 1994); and (4) different drugs of choice among street users, most notably cocaine (Bourgois and Bruneau, 2000; Chaisson et al., 1989). None of these possible explanations, however, claims to be definitive nor has satisfactory empirical evidence for them been published.

Heroin Type and HIV Prevalence

We obtained unpublished U.S. Drug Enforcement Agency (DEA) data on the types of heroin predominating in the 20 cities monitored by the Domestic Monitoring Program (DMP) from 1990–1993 (Drug Enforcement Agency, 1991–1993) and compared it with published HIV seroprevalence estimates for both IDU and MSM populations (Holmberg, 1996) (Table 1).[a] The majority of street heroin available in most cities west of the Mississippi River is "black tar heroin" (BTH): a dark, tacky resinous substance of Mexican origin. In contrast, heroin in East Coast cities, imported predominately from South Asia and South America, consists primarily of a white or light brown powder (Drug Enforcement Agency, 1991–1993).[b] Few cities have both BTH and powder heroin simultaneously presumably because of the logistics of smuggling and distribution networks (see Fig. 1).

Overall, across North America, IDU populations in cities with BTH had lower HIV prevalence than IDUs in cities with powder heroin. Two proximate cities, Vancouver and Seattle offer a dramatic contrast. Seattle, where only BTH was available had an HIV prevalence of 2.4%, whereas Vancouver, 225 kilometers away, where only powder heroin was available (reportedly from Southeast Asia) had an HIV prevalence of 23.2%.[c] Cities where the heroin was of mixed origin, such as New Orleans and St. Louis, had low to mid-range HIV prevalence. San Francisco's IDU HIV prevalence of 14.3%, the highest of any U.S. city with BTH, may be due to the fact that it was the West Coast epicenter for HIV in the MSM population and there may be an overlap in the sampling of IDU and MSM populations (Holmberg, 1996; see also Watters,

[a]In addition to the HIV seroprevalence estimates presented by Holmberg (1996) we also examined CDC seroprevalence estimates for IDUs (Centers for Disease Control and Prevention, 1994). These estimates provide an even stronger ecological association between heroin type and HIV (see also Hahn et al., 1989). We present the Holmberg estimates, however, because they are not limited to intreatment IDUs. For the DEA data on drug type we selected the years 1990 through 1993 to coincide with the years of information reviewed by Holmberg and because DEA data prior to 1990 is limited.

[b]The DEA Special Testing Research Laboratory analyses heroin obtained through undercover, retail-level purchases and broad-scale enforcement seizures. It identifies a unique chemical "signature" determining country of origin (Hast, 2001). Eighty percent of the heroin identified as of Mexican origin is black tar heroin (Drug Enforcement Agency, 2000). Powder heroin in the eastern U.S. is predominately from South America with diminishing amounts imported from South Asia.

[c]This HIV differential may also have been promoted by higher rates of cocaine injection in Vancouver than in Seattle (Bourgois and Bruneau, 2000).

1989). In addition, a significant cocaine injection epidemic was documented in San Francisco in the late 1980s (Chaisson et al., 1989).

Significantly, in the mid-to-late 1980s, San Francisco already had the highest MSM HIV seroprevalence in the US (40.7%), yet the prevalence among IDUs never approached those of comparable East Coast cities in later years; it has been stable around 10%–14% through 1999 (Kral et al., 2001). Longitudinal data (1993–1997) from the Centers for Disease Control show overall declining HIV prevalence for IDUs entering drug treatment across the United States. Still, however, the pattern continues with Western cities having markedly lower HIV prevalence than Eastern cities (Centers for Disease Control and Prevention, 2001).

We posit that the lower HIV prevalence among Midwest and Western IDUs can be largely explained by the fact that the chemical properties of BTH oblige IDUs to modify their drug-using behavior. A triangulation of evidence from epidemiological as well as clinical and ethnographical observations suggests that at least three mechanisms in concert potentially reduce HIV survival and transmission. Firstly, and probably most importantly, BTH obliges IDUs to thoroughly rinse their syringes following each injection in order to prevent the syringe mechanism from becoming obstructed. This has the unintended consequence of reducing residual blood volume and its potential HIV load. Secondly, heating is necessary to enhance drug solubility. This reduces the probability of transmitting HIV indirectly through paraphernalia (e.g., cookers) sharing (Clatts et al., 1999). Epidemiological self-report statistics in association with serotesting suggest that indirect sharing is not a primary means of transmitting HIV in the United States, although it may transmit hepatitis C (HCV) (Bourgois, 2002; Hagan et al., 2001). Thirdly, BTH promotes rapid venous sclerosis among injectors, leading them to seek alternative routes of injection (subcutaneous and intramuscular), which may transmit less blood-borne virus (Rich et al., 1998).

Ethnographic Data

Over the past decade we have engaged in participant-observation research among street-based heroin injectors in San Francisco and New York, supplemented by fieldwork visits to Montreal and Vancouver (Bourgois, 1998a; Bourgois, 1998b). Direct observation of street-based IDUs in their natural environments (i.e., shooting encampments, apartments, public restrooms, and vacant buildings, etc.) has allowed us to develop an understanding of the dynamics of risky practices with greater precision and fuller context than is possible through a self-report epidemiological survey.

We have observed that San Francisco heroin injectors predominately inject BTH. They frequently complain that their syringes become obstructed even after a single use. The BTH leaves a residue inside needles and syringe barrels, consequently, BTH injectors in San Francisco vigorously flush water through their equipment to keep the syringe mechanism and needle from clogging. We have observed that they usually flush their syringe at least once before injection to verify function and also after injection to preserve function. Rinsing, or flushing, thus has a dual effect: a direct one, to keep syringe mechanisms working and an indirect one, the reduction of residual blood volume in the syringe. By dramatically reducing blood volume, rinsing should reduce viral load in an HIV-contaminated syringe even if the rinse water used is not sterile, since there is lower viral load in residual rinse water than residual whole blood. This, consequently, diminishes the likelihood of transmitting HIV if an HIV-contaminated syringe is reused. Population-wide we might also expect increased syringe turnover in BTH cities, as unrinsed syringes become obstructed.

The solubility characteristics of BTH also require heating to place it into solution. Our observations support reported findings that BTH users consistently heat heroin solutions while

IDUs with other types of heroin do not (Clatts et al., 1999). Heroin solution that is cooled is often viscous and can be diffcult to inject though fine gauged needles. Hence, BTH-using IDUs heat their solutions more thoroughly.

In contrast to BTH, powder heroin dissolves easily in cool-to-warm water, and it does not leave a significant syringe residue. New York IDUs, consequently, do not complain of obstructed syringes nor do they rinse as regularly, or as thoroughly, as BTH injectors. They will sometimes flush their syringes with water to evacuate any blood that might have coagulated in the needle or syringe, but they do not engage in this rinsing practice as consistently or as iteratively as BTH injectors in San Francisco.

Early self-report surveys do not confirm the benefit of cleaning syringes with water (Chaisson et al., 1987; Marmor et al., 1987). Our ethnographic data suggests, however, that rinsing practices may be too unconsciously habitualized to be susceptible to accurate self-report. In addition, no studies have examined the intensity of rinsing practices. Rinsing by BTH users, according to our ethnographic data, is notably more universal and more vigorous than that by powdered heroin users. Furthermore, no multi-site epidemiological studies have examined the type of heroin injected and related it to the details of cleaning practices, or to the risk of HIV transmission across cities or countries.

Laboratory Data

Laboratory findings lend the weight of biological plausibility to our ethnographic observations and ecological analysis. Rinsing HIV-contaminated syringes three times with water in vitro reduces the number of syringes with recoverable virus 99% (Abdala et al., 2001). Rinsing blood-contaminated 1 mL syringes with water reported a 74% to 92% reduction in residual blood volume following a single rinse (Gaughwin et al., 1991). A laboratory testing HIV disinfectants reported, as an accidental finding, that “rinsing the syringe twice (with culture medium), probably even with water, reduces the number of culturable HIV-1 to a level below the sensitivity of our assay (Flynn et al., 1994). Rinsing with any fluid, consequently, should reduce the volume of residual blood in syringes and thereby lower the risk of HIV transmission.

Moreover, the higher temperatures and longer heating times required to put BTH into solution (compared to powder heroin) reduced levels of HIV in vitro (Clatts et al., 1999). This comparative heating data lends further plausibility to our ethnographic interpretations that BTH obstructs syringes due to its lower solubility and greater viscosity, especially when cooled. Thus, while rinsing may protect IDUs from *direct* HIV transmission (syringe sharing), heating may protect IDUs from *indirect* HIV transmission via ancillary paraphernalia (i.e., HIV-contaminated syringe A to cooker to syringe B).

Clinical Observations

Use of BTH has been associated with bacterial infections including tetanus, botulism, and gangrene as well as soft tissue infections (Ciccarone et al., 2001). Soft-tissue infections due to intramuscular and subcutaneous injection of BTH are very common in San Francisco (Ciccarone et al., 2001). Injectors who have used powder heroin prior to BTH report increased inflammation and venous scarring after switching to BTH. Our clinical and ethnographic data document that many IDUs who inject BTH lose their venous access early in their injection careers—sometimes within 6 months of initiation. Consequently, many BTH users are forced to switch to intramuscular injection. This early transition away from intravenous injection does not occur among those who use white powder heroin. In vitro evidence supports decreased HIV transmissibility with subcutaneous/intramuscular compared to venous injection (Rich et al., 1998).

Public Health Implications

Our findings unfold an intriguing hypothesis: Has heroin type helped shaped the spread of HIV among IDUs in the United States? If so, how strong a determinant is it and how does it interact with the complex multidimensionality of IDU risk taking? Would a simple public health campaign promoting vigorous water-rinsing of used syringes be a cost-effective and realistically implementable intervention in resource-poor settings?

The HIV epidemic among IDUs is currently stable or in a pattern of decline in many areas of the United States. This has been attributed to a dying out of the infected population, a decrease in risky practices, and increased access to sterile injection equipment (Des Jarlais, 1998). We emphasize that this decline cannot be taken for granted. If BTH is as protective as our analysis suggests, then a simple change in drug distribution patterns, e.g., intrusion of powdered heroin into traditional BTH cities, may undo what was thought to be effective public health intervention.

Global conflicts, economic restructuring, and migration patterns lead to changes in the accessibility and distribution of heroin (The Economist, 2001; Moore, 2002; Singer, In Press). Mixtures of heroin types and qualities exist in most European cities and correspondingly generate multiple use patterns (Strang et al., 2001). Concerns over the future spread of HIV and other infectious diseases related to injection drug use should account for how risky injection practices are affected by the regional and global distribution patterns of specific heroin types.

RESUMEN

Aún no se ha estudiado ni documentado adecuadamente la relación entre las distintas composiciones físicas y químicas con que la heroína llega al consumidor mundial y sus métodos de inyección y el contagio de enfermedades. En este trabajo ofrecemos una hipótesis que explicaría la irregular difusión del VIH entre consumidores de drogas inyectables en los Estados Unidos basada en el análisis de la distribución de la heroína tipo “chapopote mexicano.” proponemos, utilizando datos etnográficos, epidemiológicos y de laboratorio, que sus propiedades químicas fomentan métodos de inyección más seguros. Entre ellos se destacan: (1) enjuague de la jeringa para evitar que se obstruya, (2) calentamiento del disolvente para facilitar la disolución, (3) rápida transición de inyección intravenosa a subcutánea o intramuscular.

RÉSUMÉ

Ll existe différents types physiques et chimiques d’héoïnes a travers le monde mais la façon don’t ces différences influencent les comportements de risques et la contamination n’est ni bien comprise ni bien documentée. Cet article présente une hypothése qui explique la distribution inégale du VIH chez les injecteurs d’heroine aux-Etats-Unis en examinent la distribution et l’utillisation d’un type d’héroïne: la “Mexican black tar” (‘goudron mexicain’). D’après les données ethnographiques, cliniques, épidémiologiques et de laboratoire, nous suggérons que les propriétés chimiques de cette héroïne millitent pour un usage plus sanitaire de l’injection par: (1) rinçage des seringues pour éviter qu’ils se bouchent; (2) chauffage de l’ustensile pour une meilleure dissolution, et (3) transition rapide de l’injection intraveineuse à l’intramusculaire ou la transcutanée.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

THE AUTHORS

Daniel H. Ciccarone, M.D., M.P.H., is a clinician, anthropologist and epidemiologist who has been doing research among—as well as providing primary clinical care to—injection drug users and high-risk youth in San Francisco for years. He is Assistant Professor in the Department of Family and Community Medicine and the Department of Anthropology, History, and Social Medicine at UCSF. Dr. Ciccarone is also a staff research physician at the Urban Health Study where he directs a number of research projects utilizing both quantitative and qualitative methodologies, which aim to deepen our understanding of risk taking among socially marginalized groups.

Philippe Bourgois, Ph.D., is a cultural anthropologist who is Professor and Chair of the Department of Anthropology, History, and Social Medicine at the University of California, San Francisco. He is best known for his fieldwork among drug dealers and addicts in the U.S. inner city. His most recent book, *In Search of Respect: Selling Crack in El Barrio* (1995, with an updated second edition in 2003) won the C. Wright Mills and the Margaret Mead prizes, among others. He is currently conducting fieldwork among homeless heroin injectors and crack smokers in San Francisco with the photographer Jeff Schonberg to prepare a book for the University of California Press, *Righteous Dopefiend: Homeless Heroin Addicts in Black and White*.

Acknowledgements

This study was supported by a grant from the National Institute of Drug Abuse (R01DA10164; PI: Philippe Bourgois, Ph.D.). We thank Jeff Schonberg for his ethnographical work with us and Ann Magruder for her assistance with qualitative analysis and editing. Ben McMahan generously donated his time to create the maps from original drafts by Juan Ochoa. We also thank Brian Edlin and Alex Kral for commenting on earlier versions of this manuscript.

References

2001. War and drugs: another powder trail. *The Economist* pp. 19–20.

Abdala N, Gleghorn AA, Carney JM, Heimer R. Can HIV-1-contaminated syringes be disinfected?: implications for transmission among injection drug users . J Acquir Immune Defic Syndr 2001;28(5): 487–94. [PubMed: 11744839]

Bourgois P. Just another night in the shooting gallery. Theory, Culture, and Society 1998a;15(2):37–66.

Bourgois P. The moral economies of homeless heroin addicts: confronting ethnography, HIV risk, and everyday violence in san francisco shooting encampments [see comments]. Substance Use and Misuse 1998b;33(11):2323–2351. [PubMed: 9758016]

Bourgois P. Anthropology and epidemiology on drugs: the challenges of cross-methodological and theoretical dialogue. International Journal of Drug Policy 2002;13(4):259–269.

Bourgois P, Bruneau J. Needle exchange, HIV infection, and the politics of science: confronting Canada's cocaine injection epidemic with participant observation. Medical Anthropology 2000;18(4):325–350.

Centers for Disease Control and Prevention, (1994). *National HIV Serosurveillance Summary—Results Through* 1992. Atlanta: U.S Department of Health and Human Services.

Centers for Disease Control and Prevention, (2001). *HIV Prevalence Trends in Selected Populations in the United States. Results from National Serosurveillance*, 1993–1997. Atlanta: Centers for Disease Control and Prevention, pp. 1–51.

Chaisson RE, Bacchetti P, Osmond D, Brodie B, Sande MA, Moss AR. Cocaine use and HIV infection in intravenous drug users in san francisco [see comments]. JAMA 1989;261(4):561–565. [PubMed: 2909798]

Chaisson RE, Moss AR, Onishi R, Osmond D, Carlson JR. Human immunodeficiency virus infection in heterosexual intravenous drug users in San Francisco. Amer J Public Health 1987;77(2):169–172. [PubMed: 3467596]

Ciccarone D, Bamberger J, Kral A, Edlin B, Hobart C, Moon A, Murphy E, Bourgois P, Harris H, Young D. Soft tissue infections among injection drug users–San Francisco, California, 1996–2000. MMWR Morb Mortal Wkly Rep 2001;50(19):381–384. [PubMed: 11465906]

Clatts MC, Heimer R, Abdala N, Goldsamt LA, Sotheran JL, Anderson KT, Gallo TM, Hoffer LD, Luciano PA, Kyriakides T. HIV-1 transmission in injection paraphernalia: heating drug solutions may inactivate hiv-1. J Acquir Immune Defic Syndr 1999;22(2):194–199. [PubMed: 10843535]

Des Jarlais DC, Friedman SR, Novick DM, et al. HIV-1 infection among intravenous drug users in Manhattan, New York City, from 1977 through 1987. JAMA 1989;261:1008–1012. [PubMed: 2915408]

Des Jarlais DC, Hagan H, Friedman SR, Friedman P, Goldberg D, Frischer M, Green S, Tunving K, Ljungberg B, Wodak A, Ross M, Purchase D, Millson ME, Myers T. Maintaining low HIV seroprevalence in populations of injecting drug users. JAMA, The Journal of the American Medical Association 1995;274(15):1226. [PubMed: 7563513]

Des Jarlais DC, Perlis T, Friedman SR, Deren S, Chapman T, Sotheran JL, Tortu S, Beardsley M, Paone D, Torian LV, Beatrice ST, Debernardo E, Monterroso E, Marmor M. Declining seroprevalence in a very large HIV epidemic: injecting drug users in New York City, 1991 to 1996. Amer J Public Health 1998;88(12):1801. [PubMed: 9842377]

Drug Enforcement Agency, (1991–1993). Domestic monitoring reports jan–dec for 1991–1993. Washington, DC: U.S. Department of Justice.

Drug Enforcement Agency, (2000). The mexican heroin trade. Washington, DC: U.S. Department of Justice, pp. 13.

Flynn N, Jain S, Keddie EM, Carlson JR, Jennings MB, Haverkos HW, Nassar N, Anderson R, Conen S, Goldberg D. In vitro activity of readily available household materials against HIV-1- is bleach enough. J Acquir Immune Defic Syndr 1994;7(N7):747–753. [PubMed: 8207658]

Friedman SR, Jose B, Deren S, Des Jarlais DC, Neaigus A. Risk factors for human immunodeficiency virus seroconversion among out-of-treatment drug injectors in high and low seroprevalence cities. The national aids research consortium. Amer J Epidemiol 1995;142(8):864–874. [PubMed: 7572963]

Friedman SR, Perlis T, Des Jarlais DC. Laws prohibiting over-the-counter syringe sales to injection drug users: relations to population density, HIV prevalence, and HIV incidence. Amer J Public Health 2001;91(5):791–793. [PubMed: 11344889]

Gaughwin MD, Gowans E, Ali R, Burrell C. Bloody needles—the volumes of blood transferred in simulations of needlestick injuries and shared use of syringes for injection of intravenous drugs. Aids 1991;5(N8):1025–1027. [PubMed: 1777162]

Hagan H, Thiede H, Weiss NS, Hopkins SG, Duchin JS, Alexander ER. Sharing of drug preparation equipment as a risk factor for hepatitis c. Amer J Public Health 2001;91(1):42–46. [PubMed: 11189822]

Hahn RA, Onorato IM, Jones TS, Dougherty J. Prevalence of HIV infection among intravenous drug users in the united states. Jama 1989;261(18):2677–2684. [PubMed: 2651732]

Hast, R. (2001). *Review of the Drug Enforcement Administration's Heroin Signature and Domestic Monitor programs* Washington, DC: United States General Accounting Office, pp. 8.

Holmberg SD. The estimated prevalence and incidence of HIV in 96 large US metropolitan areas [see comments]. Amer J Public Health 1996;86(5):642–654. [PubMed: 8629714]

Koester SK. Copping, running, and paraphernalia laws—contextual variables and needle risk behavior among injection drug users in Denver. Hum Organ 1994;53(3):286–295.

Kral AH, Bluthenthal RN, Booth RE, Watters JK. HIV seroprevalence among street-recruited injection drug and crack cocaine users in 16 US municipalities. Amer J Public Health 1998;88(1):108–113. [PubMed: 9584014]

Kral AH, Bluthenthal RN, Lorvick J, Gee L, Bacchetti P, Edlin BR. Sexual transmission of HIV-1 among injection drug users in San Francisco, USA: risk-factor analysis. Lancet 2001;357(9266):1397–1401. [PubMed: 11356437]

Lange WR, Snyder FR, Lozovsky D, Kaistha V, Kaczaniuk MA, Jaffe JH. Geographic distribution of human immunodeficiency virus markers in parenteral drug abusers. Amer J Public Health 1988;78 (4):443–446. [PubMed: 3348473]

Marmor M, Des Jarlais DC, Cohen H, Friedman SR, Beatrice ST, Dubin N, El-Sadr W, Mildvan D, Yancovitz S, Mathur U, Et Al. Risk factors for infection with human immunodeficiency virus among intravenous drug abusers in New York City. Aids 1987;1(1):39–44. [PubMed: 3122788]

Montoya ID, Atkinson JS. Determinants of HIV seroprevalence rates among sites participating in a community-based study of drug users. J Acq Immun Defic Synd Hum R 1996;13(N2):169–176.

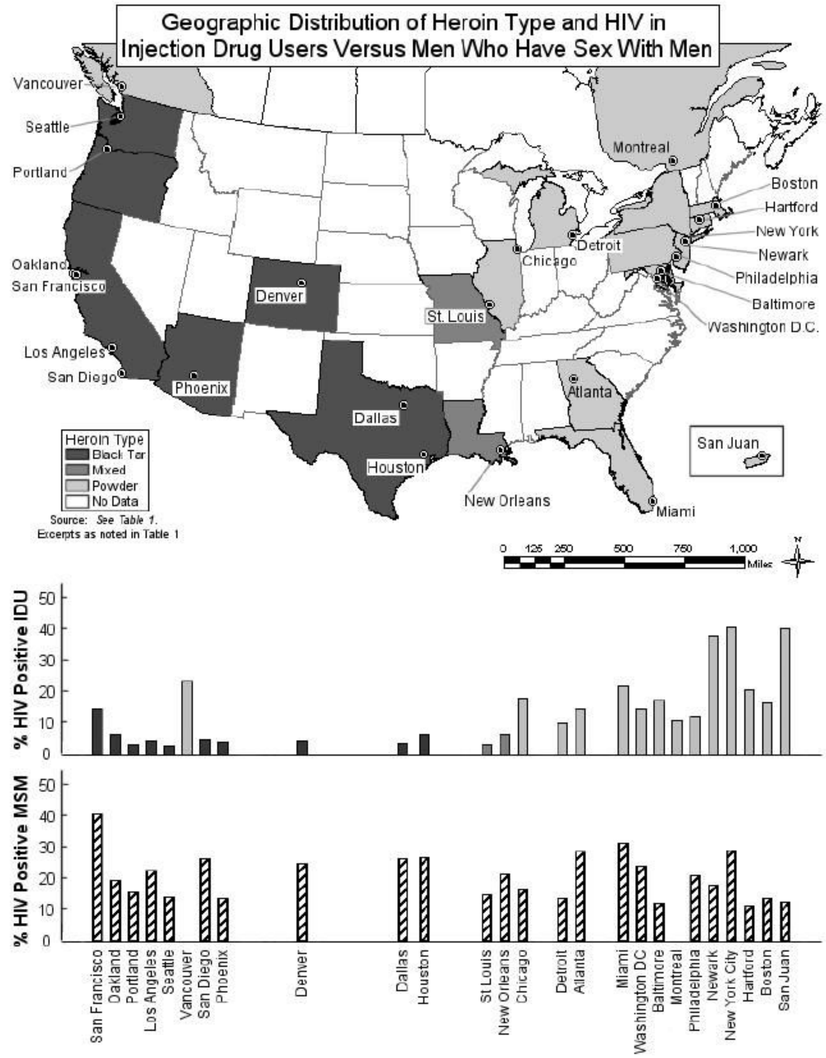
Moore LD. Surveillance, addiction, and policy: the examples of South Africa and Afghanistan. Soz-Praventivmed 2002;47:001–002.

Rich JD, Dickinson BP, Carney JM, Fisher A, Heimer R. Detection of HIV-1 nucleic acid and HIV-1 antibodies in needles and syringes used for non-intravenous injection. Aids 1998;12(17):2345–2350. [PubMed: 9863878]

Singer, M. (In Press). Fighting drugs, fighting change: rethinking the war on drugs from a public health perspective. *Medical Anthropology Quarterly*

Strang J, Keaney F, Butterworth G, Noble A, Best D. Different forms of heroin and their relationship to cook-up techniques: data on, and explanation of, use of lemon juice and other acids. Subst Use Misuse 2001;36(5):573–588. [PubMed: 11419488]

Watters JK. Observations on the importance of social context in HIV transmission among intravenous drug users. In J Drug Issues 1989;19:9.



Geographic Distribution of Heroin Type and HIV in Injection Drug Users Versus Men Who Have Sex With Men
Vancouver
Seattle
Portland
Oakland
San Francisco
Los Angeles
San Diego
Phoenix
Denver
Dallas
Houston
St. Louis
New Orleans
Chicago
Detroit
Atlanta
Miami
Montreal
Boston
Hartford
New York
Newark
Philadelphia
Baltimore
Washington D.C.
San Juan
Heroin Type
Black Tar
Mixed
Powder
No Data
Source: See Table 1.
Excerpts as noted in Table 1
0 125 250 500 750 1,000 Miles
% HIV Positive IDU
% HIV Positive MSM
0 10 20 30 40 50
San Francisco
Oakland
Portland
Los Angeles
Seattle
Vancouver
San Diego
Phoenix
Denver
Dallas
Houston
St Louis
New Orleans
Chicago
Detroit
Atlanta
Miami
Washington DC
Baltimore
Montreal
Philadelphia
Newark
New York City
Hartford
Boston
San Juan

Figure 1.

Table 1

HIV seroprevalence in IDU and MSM populations and type of heroin by city.

City	IDU Population (#)(1)	HIV seroprevalence[a] HIV+ IDU (%)	HIV+ MSM (%)	Heroin type (1991–1993)[b] % Tar heroin (91–93)
NYC	168,300	41.0	29.2	0
San Juan	22,000	40.5	12.4	0
Hartford	10,200	20.6	11.2	(−)[c]
Newark, NJ	30,000	38.0	17.9	0
Vancouver		23.2[e]		(−)[c]
Miami	31,000	21.9	31.4	14
Chicago	58,100	18.1	16.2	17
Baltimore	32,000	17.1	11.9	0
Boston	28,000	16.4	13.7	0
Washington, DC	39,100	14.5	24.2	0
Atlanta	23,000	14.5	28.6	0
San Francisco	**23,000**	**14.3**	**40.7**	**94**
Philadelphia	51,400	11.9	21.0	0
Montreal		10.7[f]		(−)[c]
Detroit	35,000	9.9	13.8	0
Oakland	**21,000**	**6.3**	**19.3**	**100(+)**[c]
New Orleans	17,700	6.2	21.6	38
Houston	**65,200**	**6.1**	**27.0**	**83**
San Diego	**19,100**	**4.5**	**26.4**	**100**
Los Angeles	**88,000**	**3.8**	**22.6**	**93**
Denver	**15,700**	**3.8**	**24.8**	**100**
Phoenix	**16,000**	**3.4**	**13.6**	**100**
Dallas	**16,300**	**3.3**	**26.6**	**95**
Portland	**16,900**	**2.8**	**15.5**	**100(+)**[c]
St. Louis	**18,000**	**2.7**	**15.0**	**72**
Seattle	**17,000**	**2.4**	**14.1**	**100**
U.S. overall[d]	1,461,500	14.0	18.1	

Note: HIV = Human immunodeficiency virus (Italicized bold face type highlights cities where black tar heroin predominates). Empty cells indicate missing data.

[a]Source: Holmberg, 1996.

[b]Type of heroin. Source: Domestic Monitoring Program 1991–1993. Drug Enforcement Administration, U.S. Department of Justice. Because of observed broad patterns for drug distribution, data for type of heroin is extrapolated from the municipal to the state level. (See Fig. 1)

[c]Type of heroin determined by ethnographical experience in Hartford, Vancouver (P. Bourgois, D. Ciccarone), Oakland (D. Ciccarone), Montreal (P. Bourgois), and Portland (D. Ciccarone). Absence or presence of black tar heroin in those cities is denoted with (−/+).

[d]Does not include Canadian cities.

[e]S. A. Strathdee et al., *Aids* 11, F59–F65 (1997).

[f]J. Bruneau et al., *American Journal of Epidemiology* 146, 994–1002 (1997).

VOLUME 49, NUMBER 10 · JUNE 13, 2002

Review

The Great Room of Art

By Richard Dorment

Art on the Line: The Royal Academy Exhibitions at Somerset House, 1780–1836
edited by David H. Solkin
Yale University Press, 278 pp., $65.00

1.

In a famous letter to the committee of artists responsible for hanging the Royal Academy's annual summer exhibition in 1784, Thomas Gainsborough announced that he could not possibly allow his full-length group por- trait of the three eldest daughters of George III to be hung at a height "higher than five feet & a half." By attempting to dictate to the Royal Academy in this way, Gainsborough was asking for a radical dispensation from the rule that full-length or three-quarter- length portraits were hung above the "line," a projecting wooden molding running around the walls of the main exhibition space at a level of eight feet from the floor. Gainsborough, who was one of the founding members of the Royal Academy in 1768, was well aware of this regulation.

But, he explained, his was a special case. He had painted the picture of the three princesses "in so tender a light, that notwithstanding *he approves very much of the established Line for Strong Effects*,...the likenesses & Work of the Picture will not be seen any higher." In other words, he considered the rule about the line to be fair —for showing big, bold pictures that would register from a distance. But Gainsbourgh created the brilliant visual effects for which he was so celebrated in a darkened studio, often by candlelight, using the techniques of scumbling (the rubbing of a light and opaque color over a darker one) and glazing (laying a darker color over a lighter one). All the delicacy of color, subtlety of touch, and nuance of his paint surface would be lost to the public, he said, unless the academicians hung his picture at eye level. His demand for special treatment was, therefore, nonnegotiable. When the hanging committee refused to comply, he removed all eighteen of the pictures he had submitted to that year's show, resigned from the Royal Academy, and never again exhibited at Somerset House.

Art on the Line, subtitled *The Royal Academy Exhibitions at Somerset House, 1780–1836*, a book of essays published to coincide with a recent exhibition of the same title at the Courtauld Institute Gallery in London, helps us to understand the context and wider implications of this quarrel. The subjects covered in the book, edited by David Solkin, include Mark Hallett's study of eighteenth-century art journalism, Michael Rosenthal's look at professional rivalry among the academicians, Martin Myrone's masterly discussion of sensational imagery in the RA exhibitions, and Ann Bermingham's account of the relationship between high art and popular entertainments such as dioramas and panoramas in London during the later Georgian and Regency periods.

Though the exhibition closed in January, it will, I predict, come to be considered one of the most influential in recent years. What Solkin and his colleagues did was to reconstruct a Royal Academy Summer Exhibition as it might have looked in the late eighteenth and early nineteenth centuries in the actual spaces in which the exhibitions took place. This was much more than simply an academic exercise. Never again will we be able to show paintings of the British school without taking into consideration the lessons we learned from this show. What is more, it came at the very moment when Tate Britain opened a new extension for the exhibition of its historic British collection, and coincided with the unveiling of the Victoria and Albert Museum's long-awaited renovation of its primary British Galleries, which tell the story of British art and design from the Tudors to 1900. Had Solkin's exhibition and publication been available five years ago, curators, architects, and designers at Tate Britain, at least, may well have made different choices about the way paintings were displayed.

2.

Although no longer in regular use as an exhibition space, the main room in which the Summer Exhibition was held still exists at Somerset House, the former government building that today houses both the Courtauld Institute and the Courtauld Gallery. Known as the Great Room, it was designed by Sir William Chambers specifically for the annual display of pictures by members of the Royal Academy. High, square, and oddly proportioned, it is this space —and how painters responded to it— that provides the key to understanding why British pictures look the way they do. For ambitious British artists worked in the knowledge that their pictures would be seen under the specific physical conditions that prevailed at Somerset House—and adjusted their painting styles, their compositions, their use of color, and even their choice of subject matter to ensure that their work would be seen to advantage there. Unless you understand the hanging system at the Royal Academy, you simply don't appreciate how desperate artists were to grab visitors by the lapels with dramatic or topical subjects, strong colors, inventive compositions, or—in the case of portraits—famous, glamorous, or notorious sitters.

In 1768, a number of artists and architects, including the architect William Chambers, the pastellist Francis Cotes, and the American history painter Benjamin West founded the Royal Academy, under the patronage of King George III. Joshua Reynolds was elected the first president. The original premises were those of an auctioneer in Pall Mall, but in 1780 the Royal Academy's exhibition rooms and schools moved to the newly built Somerset House, a palatial building on the Thames, five minutes' walk from Covent Garden and Drury Lane, otherwise occupied by government offices, including the Excise and the administration of the navy.

Here, every year between 1780 and 1836, the annual exhibition took place.[*] The summer show opened on the last Monday in April or the first in May and ran for about six weeks. To see the show—one of the great spectacles Georgian and Regency London had to offer—visitors toiled up the narrow, vertiginous, winding staircase from the dark entrance level to the top-lit galleries under the roof. Nowadays most people take the elevator, but were we to follow in the footsteps of our ancestors we would arrive, panting and with a whoosh of exhilaration, into the noise and light of the Great Room, the dramatic heart of the annual exhibition. Here, the walls were hung from floor to ceiling and frame to frame with over two hundred paintings that, Solkin and his colleagues argue, were a mirror of the British establishment, and entirely in tune with the intensely competitive spirit of entrepreneurial capitalism of the period.

The places of honor in the center of each wall were reserved for full-length portraits, often of members of the royal family. Otherwise, there were no rules about who could or could not have their portrait shown at the summer exhibition. This was the place to see what people in the news looked like—society beauties like Georgiana, Duchess of Devonshire, or Lady Elizabeth Foster, actors like John Philip Kemble and Mrs. Siddons, and writers like Dr. Johnson, Lord Byron, and Sir Walter Scott. Then there were the adventuresses and courtesans who sometimes owed their livelihoods to the notoriety the appearance of their portraits at the RA bestowed upon them. According to the *Daily Universal Register* of May 1786,

> The French who visit our exhibitions are shocked at the indelicacy of placing the portraits of notorious prostitutes...close to the pictures of women of rank and virtue. In Paris, such portraits would on no account be admitted.

In a separate category were paintings of military heroes and their exploits, as well as depictions of topical events such as James Northcote's *Portraits Painted from Life, Representing Capt. Englefield with Eleven of his Crew Saving Themselves in the Pinnace, from the Wreck of the Centaur, of 74 Guns, Lost Sept. 1782*, exhibited in 1784.

In a world without photography or even cheap reproductive engraving, how interesting it must have been to see what the celebrities you had only read about in the newspapers or gossiped about in the coffeehouses actually looked like. Guessing who was who was part of the fun. The paintings didn't have labels, and by convention the identity of the sitters was concealed in the catalog under the title "Portrait of a Lady" or "Gentleman." The satirist Anthony Pasquin described the crowds at one summer show as "buzzing and fidgetting about the room...in the ardent wish to know who or what such a lady or gentleman is...." Clearly the RA summer show functioned as an early version of *Hello* or *People* magazine.

Right from the beginning of its existence, and long before the move to Somerset House, the Academy had introduced an admission charge of one shilling in order to keep out poor and uneducated members of society. In a controversy recently echoed in the successful struggle to achieve free entrance to all national museums in Britain, the academicians felt guilty about the imposition of this admission charge, well aware that, since the exhibition was under the patronage of the King, the RA could be criticized for not opening its doors free of charge to all

British people. But the RAs felt that

> they have not been able to suggest any other Means, than of receiving Money for Admittance, to prevent the Room from being filled by improper Persons, to the entire Exclusion of those for whom the Exhibition is...intended.

In fact the academicians' pious disclaimer told only part of the story. A pastel by John Russell, which serves as the frontispiece to this volume, reveals something else about the tensions that existed between the academicians and their public. At first glance, it simply shows an elderly porter taking admission tickets to the Royal Academy exhibition of 1792. But in the background behind this mild-mannered fellow, Russell sketches in what could be mistaken for a revolutionary mob surging through the entrance foyer and up the staircase at Somerset House. Despite the admission charge, these are not polite, well-mannered people. The men have not removed their wide-brimmed hats, and some of the women look as if they might be crushed in the movement forward. With the revolution raging across the Channel, the King and his Royal Academicians had good reason to feel nervous about encouraging large gatherings of "improper persons" in a gov- ernment building with strong royal associations in central London. Russell subtly suggests these fears by the way in which he juxtaposes the headless torso of a classical fragment in the middle distance with a form that looks exactly like a pitchfork—until you look twice and realize that it is in fact the metal rod of a standing lantern.

Like modern blockbusters, the summer exhibition was a highly lucrative business. The entrance fee made for the Academy a fat profit, which the academicians used to meet expenses for the rest of the year. In 1780, for example, the show attracted 61,000 people—about 2,000 per day in the six weeks it was open. Thereafter it averaged about 50,000 per season, reaching a peak of 91,827 in 1822. An amazed Samuel Johnson wrote Mrs. Thrale in May 1783:

> On Monday, if I am told truth, were received at the door, one hundred and ninety pounds, for the admission of three thousand eight hundred Spectators. Supposing the show opened ten hours, and the Spectators staying one with another, each an hour, the rooms never had fewer than three hundred and eighty justling [*sic*] each other.

The Great Room measured only 53 ¥ 43 feet, so visitors must have been packed in shoulder to shoulder. May and June can be very hot in London; the exhibition rooms were directly under the roof; people did not bathe frequently. No wonder that some ladies were "ready to faint, on account of the heat of the rooms, and the powerful perfumes of the odiferous company [with] which they are filled." Under such viewing conditions the Royal Academicians were forced to hang the pictures in such a way that they could be seen. Chambers's solution was a radical innovation in the history of displaying art, namely to establish the "line" of molding above which large and easily visible pictures could be hung, and which, in addition, served to bear some of their weight. The line established a common height for the largest canvases, and created a sense of symmetry that so higgledy-piggledy a collection of pictures badly needed. One of the strangest features of an RA hang is that pictures hanging above the line (and some below it) were canted outward from the wall at an angle of seventeen degrees from the perpendicular. This was simply a practical solution to an annoying problem. In what must be seen as a fault in Chambers's design, the light that flooded into the Great Room from the four semicircular windows in the lantern created glare on the reflective varnish of the pictures, which made them difficult to see unless they were tilted forward. This was relatively easy to do because the pictures were not nailed to the walls or hung from chains, but lashed by ropes to a wooden armature or scaffold erected every year specifically for the summer exhibition and concealed by a covering of green baize.

Smaller paintings, head-and-shoulder- sized portraits, landscapes, and genre subjects were all hung below the line, so that you could step up and see them close too. At the exhibition "Art on the Line" visitors found themselves in constant movement, looking at the pictures from two entirely different perspectives. When they wished to see the art below the line they moved close to the walls, but to see the pictures above the line they had to step back into the center of the room. Eighteenth-century visitors brought spyglasses and telescopes to inspect the pictures near the ceiling, and for the recent show the organizers provided binoculars, which, I must add, were not really necessary. But as you moved about the Great Room, you realized something else about the RA exhibitions: the way the pictures were hung encouraged visitors not only to look at what was on the walls, but also to look at each other. The theatrical space created by Chambers ensured that the people who came to the exhibitions were as important in creating a sense of excitement as the pictures on the walls. Surrounded by other visitors, you felt you were on a stage, looked down upon by the full-length portraits above you. It was a place to see and be seen.

After *Art on the Line*, museum curators must now take note of a few basic facts about British paintings. The larger ones—mainly history paintings and full-length portraits—work best when seen from below, and arguably best of all when tilted slightly forward. Take for example Thomas Lawrence's portrait *John Philip Kemble as Coriolanus*, which is usually shown at eye level. Seeing it for the first time from below, as we did last autumn, the actor's pose seems convincing for the first time. Suddenly we realize that we are meant to be looking up at him as though from the pit in the theater, as he steps forward to the footlights, just above our heads. The implications for the exhibition of all historic British paintings seem to me enormous.

The bizarre conditions under which art was seen in eighteenth-century Britain can be said to have changed the way British artists painted. This was the era before the rise of dealers in modern art, so another function of the summer show was as a shop window. To make his name at the Royal Academy summer show, an artist first had to be noticed. Larger pictures had to register their effect from quite a distance while hung frame to frame and against the ubiquitous green baize. This meant artists tended to favor broad, easy-to-read compositions and colors that would blend easily with the old masters their pictures were likely to hang next to, since only owners of great houses were generally able to commission or purchase art on this scale. Those who specialized in smaller genre pictures and landscapes faced slightly different problems. Since their pictures were more moderately priced, they sold more readily than large history pictures and full-scale portraits, but only if the landscapes showed familiar sites with pleasant associations, or, in the case of genre, if the subject were sentimental, charming, or edifying. It is hard to think of a significant still-life painter who exhibited successfully at the Royal Academy summer show during this period. Certainly there is no British equivalent to Chardin.

In the intensely competitive space of the Great Room, all artists—whether portrait, landscape, history, or genre painters—chose subjects that would either attract attention or readily sell. Joshua Reynolds, the academy's founding president, had hoped that by giving the British public access to the best works of contemporary art, his institution would instruct and elevate the nation's taste. Instead, the exact opposite happened. Desperate for attention and sales, artists became obsessed with novel subject matter or striking visual effects. As early as 1807 a monthly art periodical noted that "all pictures...must be coloured above nature, to prevent their being either overborne by the works of others, or overlooked by visitors in so large a room." Year by year colors became stronger, so that the soft palette of the eighteenth century was gradually transformed into the harsh, bright one we associate with the Victorian Era.

Not only were the landscape painters competing against one another, but they were well aware of the spectacular *son et lumière* panoramas and dioramas that were so popular during this period, and chose for their motifs subjects that would enable them to show off their virtuosity in reproducing ever more violent effects of light. This in part explains the attraction to subjects such as coal furnaces at night, volcanic explosions, or, in the realm of fantasy, theatrical extravaganzas such as John Martin's visions of heaven and hell or Francis Danby's *Opening of the Sixth Seal*. Another tactic was to attempt to overshadow the competition by the sheer size of your picture, as Benjamin West, Thomas Lawrence, and the landscape painter James Ward were apt to do. The point is that either an artist chose to compete in the marketplace by becoming a commercial entertainer or he faced obscurity.

Then too, the Academy became something of a snake pit, as painters competed against each other for critical notice and sales. In 1806 the Scottish painter Sir David Wilkie showed *The Village Politicians* to great acclaim. This spurred Turner to paint his "challenge," *A Country Blacksmith*, which was exhibited, as was Wilkie's *The Blind Fiddler*, the following year. In "Art on the Line" the last two pictures were hung together to demonstrate the development of an "anything you can do I can do better" mentality. More aggressive is the story of how Turner at the 1832 exhibition, seeing that his pale green seascape had been placed next to Constable's strongly colored *The Opening of Waterloo Bridge*, added a luscious crimson disc of pigment to the foreground of his picture, which later he shaped into a buoy. Constable, seeing how effectively Turner had diminished the impact of his picture, could only say after his rival's departure, "He has been here, and fired a gun."

But at the Royal Academy, any tactic was fair. The Swiss-born painter Henry Fuseli, seeing the great Sir Joshua Reynolds at work on the subject of the "Death of Dido" for the Royal Academy exhibition of 1781, made his reputation by showing his own version of the subject at the same exhibition. When the two pictures were exhibited that spring, visitors saw that one composition was horizontal, the other vertical; one showed the strong emotions on the faces of the protagonists, in the other the faces were hidden or half-hidden. Whatever Reynolds had done, Fuseli did the opposite. Through such stunts, unknowns and foreigners made their name in the bullring that was the London art world. Fuseli went on to a long career startling the gallery-going public with obscure, erotic, and violent subjects. Such pictures were broadly painted, with strong contrasts of light and shade, and little detail. Since their purpose was to create a sensation from a distance, there was little point in fussing over fine passages of painting. In July 1832 a critic writing in *Fraser's Magazine* complained that the "annual exhibitions are quite as much calculated to created a hankering after mere novelty...as to induce attentive study of what is most deserving."

For those of us who enjoy making parallels between the past and the present, you can see in these early tactics the origins of shows like "Sensation" (which attracted more than 200,000 visitors at the Royal Academy in 1999, before it traveled to Brooklyn to outrage Mayor Rudolph Giuliani in 2000). British artists know now what they knew then: that the way to bring the public in is to make art that is novel, attention-grabbing, and easy to understand. Poor Reynolds had to watch this happening:

> Our Exhibitions while they produce such admirable effects, have also a mischievous tendency, by seducing the Painter to an ambition of pleasing indiscriminately the mixed multitude of people who resort to them.

Fraser's Magazine complained that

> instead of educating themselves up to the level of literature and art, the people demand that both sink down to the level of their taste and comprehension.... Unless something occur to interpose a timely check to our present unnatural position, the million will, ere long, be the principle if not the sole arbiters in all matters of taste.

Then, as now, the British press fanned the flames of controversy, built up its favorites, and cruelly ridiculed the artists it didn't like.

3.

Paradoxically, Gainsborough fully accepted the reality of what the British public liked to look at and was willing to hang on its walls. Unlike his rival Joshua Reynolds, who stubbornly believed that that public could be educated to like subjects from classical history and Shakespeare, Gainsborough believed that an artist

> may do great things, and starve in a Garret, if he does not conquer his Passions and conform to the *common Eye* in chusing that branch [i.e. portraits] which *they* will encourage, & pay for.

After reading Sir Joshua's fifth discourse, calling for a national school of history painting, he wrote to William Hoare in 1773,

> But betwixt Friends Sir Joshua either forgets, or does not chuse [to] see that his Instruction is all adapted to form the History Painter, which he must know there is no call for in this country.... Therefore he had better come *down* to *Watteau* at once (who was a very fine Painter taking away the french conceit) and let us have a few Tints. Every one knows that the grand Style must consist in plainness & simplicity, and

> that silks & Sattins Pearls & triffling ornaments would be as hurtfull to simplicity, as flourishes in a Psalm Tune.

Gainsborough was not the only artist of his time unwilling to play by the RA's rules: George Stubbs, Joseph Wright of Derby, and John Singleton Copley all came into conflict with the Academy over these issues. Gainsborough made more of a fuss than any other academician because his pictures suffered more than most from the hanging system. Sir Francis Bourgeois remembered him saying that "chaste colouring was as necessary to a picture as modesty to an artist." As early as 1772 Gainsborough offered this advice to David Garrick, about the placement of the artist's recently completed portrait of the actor:

> If you let your Portrait hang up so high, only to consult your Room ...it never can look without a hardness of Countenance and the Painting flat: it was calculated for breast high, & will never have its Effect or likeness otherwise.

The following year Gainsborough reported, "I don't send to the Exhibition this year; they hang my likenesses too high to be seen, & have refused to lower one sail to oblige me." He did not exhibit again until 1777.

The subject blew up again in 1783 when Gainsborough sent to Somerset House his small oval portraits of the King, the Queen, and their thirteen children, along with a sketch showing how he wished the works to be hung—in rows of five as a single block, with their frames touching. Fearing that these exquisitely painted ovals might be hung above the line, he wrote to the Academy's secretary with the energy and wit that characterize all his letters:

> Mr Gainsborough presents his Compliments to The Gentlemen appointed to hang the Pictures at the Royal Academy; and begs leave to *hint* to Them, that if The Royal Family, which has been sent for this Exhibition (*being smaller than three quarters* [i.e., smaller than a standard 30 ¥ 25-inch portrait]) are hung above the line along with full lengths, he never more, whilst he breaths, will send another Picture to the Exhibition—This he swears by God.

Bearing in mind that each individual portrait was so small that it would normally have hung below the line anyway (it was only the block of portraits that presented the difficulty) and that the King was the patron of the Academy, the committee gave way.

The following year they put their collective foot down. The full-length portraits of the three eldest princesses had been commissioned not by the King but by his son the Prince of Wales for his palace, Carlton House. Now Gainsborough didn't have a leg to stand on. His picture measured 100 ¥ 70 inches—even larger than the standard full-length size of 90 ¥ 55 inches. To hang it at eye level would ruin the symmetry on which the integrity of the academy's hang depended. His argument that the line was established for broad effects but not for pictures conceived in a different spirit didn't really hold water.

So who was right? Until recently, all our sympathies were with Gainsborough. In designing four new galleries as part of the handsome new extension to Tate Britain, the architects John Miller and partners reproduced the proportions of the old top-lit, barrel-vaulted galleries designed by John Russell Pope in the 1930s. These gracious galleries reflect the aesthetic of their period, by allowing the eighteenth-century pictures to be shown only in single rows, hung at eye level and with plenty of space between them.

That never bothered me before, but tastes in hanging pictures, like tastes in art, change. After seeing "Art on the Line" I felt that pictures like Reynolds's *Three Ladies Adorning a Term of Hymen* and Gainsborough's *Giovanna Baccelli* really needed to be hung high enough on the wall to be seen from below, as the artists must originally have intended. It is only when we come to the long gallery in which eighteenth- and nineteenth-century history paintings are hung at Tate Britain that the paintings are double- and even triple-hung. Here, the height feels appropriate—but unfortunately very few pictures in this particular gallery are of the right scale to be seen from such a distance.

The new British Galleries at the Victoria and Albert Museum amount to a whole new museum-within-a-museum for London. Laid out in fifteen galleries on two floors with three thousand objects (two thirds of which have either not been seen before or were shown elsewhere in the museum), the £31 million project tells the story of British culture, style, and taste through displays of historic furniture, textiles, dress, ceramics, jewelry, silver, prints, paintings, and sculptures over a period of centuries. On the face of it, the installation would seem to have very little to do with the question of how British paintings are exhibited, because pictures are only one element in the display.

In fact, the approach of the curators Christopher Wilk and Sarah Medlam and the designer Dinah Casson harks back to the way pictures were shown at the Royal Academy in the eighteenth century. Instead of a spare, spot-lit installation which leads the visitor from one masterpiece to another, major art works are juxtaposed with lesser works, relatively minor pieces sit side by side with those of the utmost rarity, and every section is so crowded with objects that at times the visitor hardly knows where to look. And so, Bernini's bust of the dandy cavalier Thomas Baker (in my view, among the ten most important works of art in England) is shown next to a not particularly important copy of Van Dyck's triple portrait of Charles I.

The historical point this juxtaposition makes is of great interest (the original Van Dyck, now lost, was sent to Bernini as a guide to carving a marble bust of the King), but we have to read the label carefully to find this out. Of course, superb lighting and clear labeling help to bring outstanding works to the fore, but the overall effect, I think, is closer to what Somerset House was like in the eighteenth century than the new galleries at Tate Britain. As then, visitors find themselves working hard to focus on individual works of art. The wheel has come full circle. The job of separating the great from the good, and the good from the bad, is ours, not the gallery's.

Notes

[*] After 1836 the Royal Academy moved to the National Gallery in Trafalgar

Square before transferring to its present location, Burlington House, in 1856.

rials available to him in the toilet—basically for wiping purposes, though it is easy to see how readily it would have been used for diversion if circumstances demanded. During the Reformation in England, for example, quantities of monastic manuscripts found their way into the great houses, not for their content or their beauty but for their value as "tissue." In recent American history, outdated mail order catalogs and newspapers have served the same functions. Today, reading matter is not simply at hand, so to speak; it is institutionalized in the form either of a library to furnish diversion or distraction, or of business reading in order to accomplish more. In at least one instance, this proclivity has also been exploited with a fair degree of commercial success. The Starlite Tissue Company in West Germany is producing toilet paper with English language lessons printed on it for "those Germans who have always said they wanted to learn English but never had the time." In a somewhat analogous fashion, Dylan Thomas once suggested that novels might well be serialized on rolls of toilet paper.

Actually, this idea has been exploited in a number of other ways; for example: the "joke" tissue with sayings like "All's well that ends well" is presumably intended to divert in every sense. One of the boldest schemes has been to use the inside of public toilet stall doors for advertising messages, which is, in some ways, only an institutionalizing of graffiti. Although graffiti is a complex phenomenon (which will be explored in more detail in the section on public facilities), certainly one link, seldom noted, is with the psychologically diverting tactics associated with defecation.

In conclusion, it is obvious that elimination is even more burdened with a variety of attitudes that must be accounted for in the design of equipment and facilities than body cleansing is. In particular, the topic is considered at least in poor taste by many and obscene by others, and whatever attitudes one holds are apt to be pursued with vehemence, if not vengeance. As proctologists and some microbacteriologists are fond of pointing out, however, "It may be shit to you, but it's my 'bread and butter.'"

ANATOMY AND PHYSIOLOGY OF DEFECATION

9

THE PROCESS OF DEFECATION

Defecation is the process of emptying the bowels of feces—waste materials left from the digestive and other body processes. Literally translated from the Latin *defaecare,* it means "to cleanse from the dregs."

The feces of an average healthy adult varies in size from 100 to 205 mm (4 to 8 inches) long by 15 to 40 mm (½ to 1½ inches) and in weight from 100 to 200 grams. Both the consistency and odor as well as the size and weight may vary considerably, depending upon a person's diet, general state of health, and the particular state of the gastrointestinal system. The stool resulting from a basically vegetarian diet, for example, is generally larger, softer, and less odoriferous than the stool resulting from a high-protein diet. Food residues account, however, for only a part of the total bulk of the feces. The remainder is made up of bacteria, dead cells, and a variety of other internal body secretions and fluids. In terms of chemical composition, feces generally consists of approximately 65 per cent water, 10 to 20 per cent ash, 10 to 20 per cent soluble substances, and 5 to 10 per cent nitrogen.[1]

The process of defecation may be said to begin when the basic digestive processes have been completed and the residue of the food leaves the stomach and passes through the small intestine into the cecum, or pouch, at the beginning of the colon. Here the fecal mass progresses more slowly, taking about twelve hours to pass through the colon. Normally the fecal mass does not pass into the rectum until the act of defecation is about to occur. The sigmoid, or the end of the colon, is where the fecal matter is stored, and the rectum is merely a passageway and "notification chamber." The entrance of the fecal mass into the rectum is brought about by a massive peristaltic movement, a propelling motion set up by the muscles. The desire to defecate is initiated by a distention of the rectum caused by this peristaltic movement. If defecation does not follow the original notification, the rectum again relaxes and the desire to defecate does not return until a further increment of fecal matter enters the rectum and distends it further, or unless the action is consciously initiated.

Although this phase of the defecation process is largely involuntary, the final act of expulsion itself can be controlled and regulated through a conscious use of the related musculature. It is this control, which we refer to as "toilet training," that every child is required to learn. In its simplest terms, this voluntary or controlled phase consists of: 1 / contracting the diaphragm and the muscles of the abdominal wall, which we commonly know as "straining"; 2 / assuming a doubled-over or squatting posture so that the thighs provide support, or resistance, to the abdominal wall, and thus permit a greater intra-abdominal pressure; and 3 / relaxing the external sphincter muscle that controls the opening of the anal canal. Normally, the anal canal is totally collapsed by the contracted sphincter, and the anus is a mere slit. Control of this sphincter muscle is particularly vital, as everyone who has ever had to postpone defecation is aware.

Of all the various sets of neuromuscular actions over which man must develop control, defecation is probably one of the most difficult and complex, largely because of the potent psychological factors involved in its development. In spite of the degree of conscious control we may achieve over this process, there still remains a substantial area that may be considered to be controlled purely by reflex action. For example, a person with highly regular habits accompanied by a set of ritual procedures may experience all sorts of difficulties if this routine is disturbed, even though that person may be quite adult and otherwise capable of controlling his actions. Most abnormal defecation—constipation and diarrhea—results directly or indirectly from an interruption of the reflex pattern or disturbance of the normal mental set, such as extreme fright or anxiety. In still other instances, disturbances of the central nervous system or of cerebral functioning can lead to abnormalities of defecation.

DIARRHEA

Diarrhea is the passing of feces in a liquid or semi-liquid state, often with great expulsive force. Usually it is brief in its duration and ascribable to a toxic, infectious, or dietary origin, and occasionally to nervous disorders. The most common form is often described as "emotional diarrhea" and may in some instances persist together with other functional anomalies for some time. For the purposes of this discussion, it is a condition that can be neither assisted nor aggravated by the physical accommodation provided for defecation, but it is a condition that occurs universally and frequently enough to be considered in the ultimate design of any fixture.

CONSTIPATION

"Constipation may be defined as an abnormal retention of fecal matter in the intestinal canal or an undue delay in the discharge of excreta from the rectum."[2] The most significant point with respect to constipation is that it means simply an unusually long interval between bowel movements, whether caused by a retention of feces or by a *lack* of feces. This last point is largely unknown as far as the public is concerned, with unfortunate results. It is still a popular belief that at least one bowel movement daily is essential to proper health and that a person must purge himself of the toxins and poisons that build up in his system. One recent survey in the United States indicates that "two-thirds of the American population have the belief that a bowel movement every day is necessary for health, and nearly a third believe it is appropriate to do something regularly to help with bowel movements." Of the latter group, 76 per cent use laxatives, 22 per cent eat special foods, 5 per cent use an enema, 4 per cent a suppository and 3 per cent some other aid. The total is more than 100 per cent because some persons use more than one method. Approximately 2 per cent of the total population, some two and one-half million persons, use some aid every day.[3]

While this practice may have a general usefulness, it cannot be regarded as an absolute rule to be applied to all circumstances, as any number of normal everyday changes in routine or diet may cause a temporary change in a person's accustomed pace. Many people become disturbed, however, at any such change and resort to laxatives or purges that clean out the colon so thoroughly that several days are necessary before a sufficient bulk is built up again for normal functioning. It is not realized that the feces is less than half food residue. The remainder is made up of dead cells and other residues of the body's constant regenerative activity. During this "normal" and necessary period of inactivity, there is little or no desire to defecate or even the possibility of doing so, but the person often becomes increasingly anxious and starts off a vicious circle of events by again purging himself.[4] These persons will often develop gastrointestinal disorders because of the strain such practices put on the system.

The other, and probably more common, form of constipation, and the one most directly related to fixture design, is the undue retention of feces. This may be brought about by a variety of factors: 1 / an insufficient reflex activity of the intestines; 2 / a lack of normal stimulus; 3 / a resistance to movement of the feces by an internal obstruction, a disorder of the anal sphincter, or a pressure on the bowel such as from pregnancy; or 4 / what might be termed "the effects of civilization," that is, unbalanced diet, laxative abuses, improper posture, and weaknesses of the diaphragm, abdominal wall, and intestinal musculature. In many such instances the fecal mass has become so large and so hard that it must be broken up manually before it can be finally ejected. So long as feces remains in the gastrointestinal tract the body continues to draw off water, with the result that the longer the feces remains in the body the drier and firmer it becomes. This condition is further aggravated by the average American and western European diet, with its high protein content.

The significance of this form of constipation, which has often been referred to as "the Great American Disease," is not to be taken lightly. Many physicians regard habitual constipation as one of their commonest and most difficult problems.

> This condition is reflected in the statistics of illness in industrial populations where the functional gastrointestinal disorders rank second only to the common cold as a cause of absence from work due to illness. Patients affected by functional colonic disorders seldom are very sick but, in the aggregate, they constitute a large medical and socio-economic problem.[5]

Chronic constipation, whether organic or self-induced, is also often a major cause of hemorrhoids, or piles, which essentially is a dilating or rupturing of the blood vessels of the rectum and anus brought about by habitual hard straining while one is attempting to defecate or by the use of strong laxatives. The ultimate irony is that persons who develop hemorrhoids as a result of their imaginary constipation often then hold back defecation because of the pain of defecation and develop organic constipation. Estimates of the incidence of hemorrhoid sufferers run as high as one-third of the population.[6] In addition, habitual hard straining also stresses the heart through the Valsalva maneuver and can result in serious complications for persons with heart conditions.[7]

POSTURAL CONSIDERATIONS

Virtually every physician and physiologist who has ever troubled to write on the subject agrees that there is a natural and physiologically sound posture that encourages the defecation process. As the following sampling from the literature illustrates, there has also been substantial agreement over the years that contemporary water closets do not permit this posture.

> The ideal posture for defecation is the squatting position, with the thighs flexed upon the abdomen. In this way the capacity of the abdominal cavity is greatly diminished and intra-abdominal pressure increased, thus encouraging the expulsion of the fecal mass. The modern toilet

> seat in many instances is too high even for some adults. The practice of having young children use adult toilet seats is to be deplored. It is often necessary for them to sit with their feet dangling. Unless the toilet seat is low enough that the feet may rest firmly on the floor and some flexion of the thighs is possible, the accessory muscles which aid in defecation normally have little opportunity to fulfill their function.[8]

> Man's natural attitude during defaecation is a squatting one, such as may be observed amongst field workers or natives. Fashion, in the guise of the ordinary water closet, forbids the emptying of the lower bowel in the way Nature intended. Now in this act of defaecation great strains are imposed on all the internal organs. . . . It is no overstatement to say that the adoption of the squatting attitude would in itself help in no small measure to remedy the greatest physical vice of the white race, the constipation that has become a contentment.[9]

> It should be mentioned in this connection that a very common cause for unsatisfactory results . . . is improper height of the toilet seat. It is usually too high. An ideal seat would place the body in the position naturally assumed by man in primitive conditions. The seat should be low enough to bring the knees above the seat level.[10]

> . . . the high toilet seat may prevent complete evacuation. The natural position for defecation, assumed by primitive races, is the squatting position. Dr. Hurst points out that in Japan there are no seats in the toilets. The pan is sunk in the floor, and in some places a shoe is fixed to the ground on each side of it so that a firmer foothold may be obtained when the individual squats. When the thighs are pressed against the abdominal muscles in this position, the pressure within the abdomen is greatly increased, so that the rectum is more completely emptied.
>
> Our toilets are not constructed according to physiological requirements. Toilet designers can do a good deal for people if they will study a little physiology and construct seats intended for proper defecation.[11]

In view of the fact that proper posture can contribute substantially to ease of defecation, it becomes obvious that if we are going to continue to accept the notion of a water closet as we have known it, we must begin to pay attention to the basic functional problems associated with its use. While on the one hand, there is little doubt about the fact that a full squat posture would be difficult and uncomfortable for most of us to assume and maintain for any length of time—primarily because of a lack of exercise, we must also bear in mind that while we regard the use of the water closet as natural, we represent only a relatively small percentage of the world's population, and a percentage that may be said, in an absolute sense, to be wrong, insofar as we have allowed civilization to interfere with our biological functioning.[12] The irony is that virtually all of the "squat plates" used elsewhere in the world are manufactured in the "civilized" Western countries. In a World Health Organization report dealing with hygiene facilities for developing nations, no mention is made of water closets, but elaborate instructions are provided for the construction of squat plates.[13]

The problem we are ultimately faced with is that the body must have a certain amount of exercise in order to keep functioning optimally; the musculature that needs to be brought into play during the defecation process is also the musculature we need to use in assuming or rising from a squatting position. The unfortunate consequences of this circular dilemma is perhaps best illustrated with reference to the aged, many of whom suffer from functional constipation caused in large measure by a natural decline in muscle tone. H. L. Bockus observes that:

> A marked atrophy of the oblique muscles of the lower abdominal wall and diastasis recti are quite common in elderly constipated patients. . . .In patients with an atony of the abdominal musculature, certain procedures may be recommended which may increase the efficiency of the voluntary aids to defecation. Among the more important is the advice to assume the normal posture for moving the bowels. In some instances constipation may be corrected by the simple expedient of having the patient squat over a bed chamber rather than use an ordinary toilet seat. Another alternative is the use of a block of wood or stool on the floor

> in front of the toilet fixture so that adequate support may be available for the feet while permitting the thighs to be flexed upon the abdomen.[14]

Our actual practice is, however, generally to the contrary. This weakness of the abdominal and upper leg muscles also tends to make it difficult for such persons to accommodate themselves to a normal sitting height, let alone to a squat position. The "solution" has been to provide "seat extenders" of various sorts to raise the height, or to provide higher water closets. This has obviously helped the problem of raising and lowering the body, but it has just as obviously aggravated the functional and physiological problems of defecation. Admittedly the problems of the aged will not be resolved by expecting them to use a squat plate, but it is not unreasonable to suppose that the use of a substantially lower water closet over the years will provide us with some of the exercise we need and keep the problems from assuming such major proportions in our later years. In short, the more apathetic we are about making the necessary effort the more difficult it will become.

DESIGN CRITERIA FOR DEFECATION

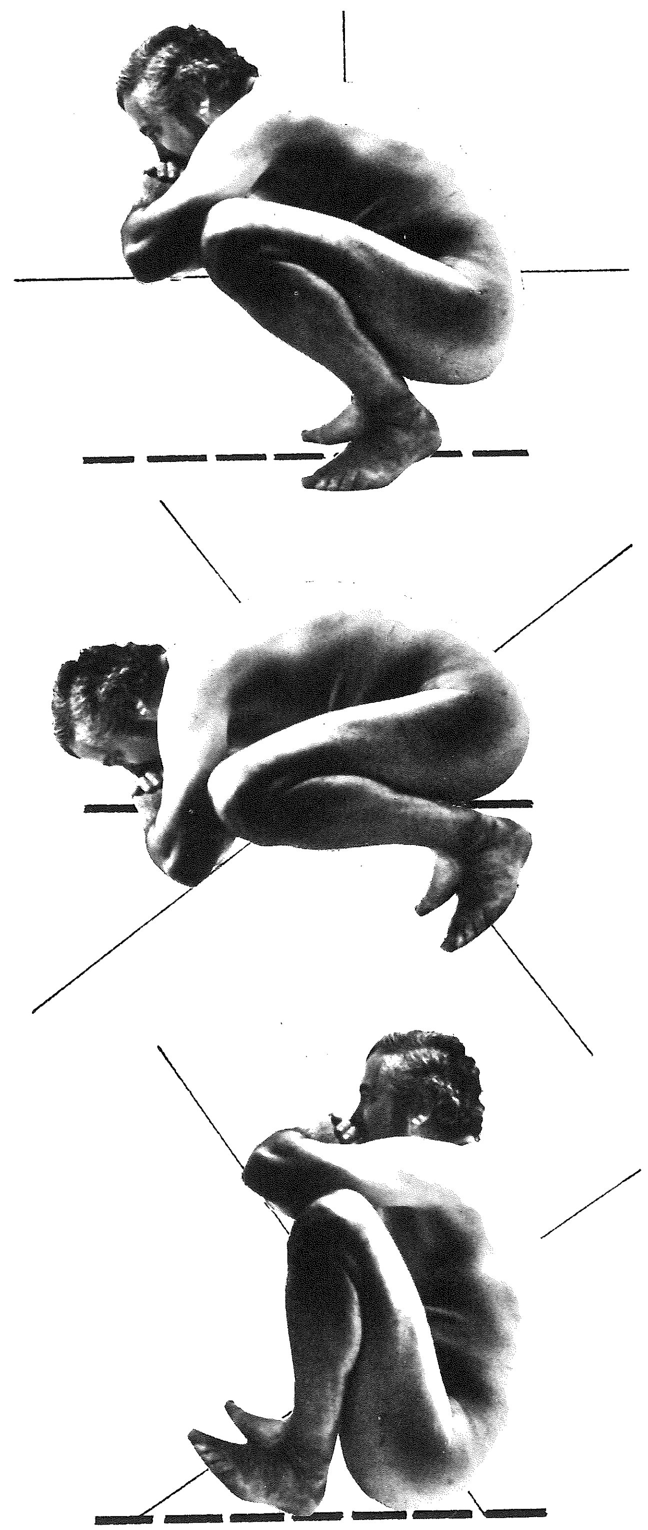
10

DESCRIPTION OF ACTIVITY

For the purposes of analysis we can divide the activity of defecation into several distinct phases: adjustment or removal of clothing; assumption of a squatting or sitting position; initiation of defecation by straining and relaxing the anal sphincter; defecating, urinating, and farting; wiping, cleansing, flushing; and readjusting one's clothing. While this sequence may be more or less complicated by incidental activities such as reading or by personal idiosyncrasies such as repeated flushing or examination of one's feces, these do not essentially alter the basic pattern. From a design standpoint, the major variable in this activity is the particular posture assumed for which the fixture will have to be designed. The management of clothing, wiping and flushing procedures, and so on are secondary, and the treatment of specific applications will vary, depending upon the precise posture and fixture involved.

The ideal posture for defecation is the full squat, which provides the stomach muscles with the proper support during the straining/expulsion process, as contrasted with the familiar casual "sitting-on-a-chair" posture so often assumed on a standard water closet, particularly by those who read there. In this posture the individual is essentially passive and unable to do much to aid the body's natural mechanisms. Conversely, it is possible, by doubling over, folding the arms across the stomach, and drawing the feet back, to assume a fair approximation of the desirable squat posture even on a standard closet. As Figure 32 illustrates, the squat posture can thus be assumed in a variety of attitudes, particularly since the most important relationship is that between the trunk and the upper leg. Stressing the legs is also helpful and can be achieved by drawing the legs back, which automatically transfers some of the body's weight from the buttocks to the legs. In this respect the concept of a water closet as a "throne" is particularly invidious in terms of its suggestion about sitting posture, particularly when the water closet has been encased in a *chaise percée*, a decorative marble shroud, or some other bit of disguising cabinetwork, since in all such instances one is prevented from drawing one's feet back.

Although we can postulate that the free full-squatting posture for defecation practiced by most of the world's population is ideal from the viewpoint of physiological functioning, this posture is an unaccustomed and difficult one for most Western people to assume, let alone maintain, for any length of time, particularly if unaided. While this difficulty can easily be overcome with sufficient practice and exercise, a more serious impediment to its adoption is the problem posed by contemporary Western clothing. In large measure the problem of clothing is responsible for the gradual shift in Japan from their traditional squat closets to Western-style water closets. Having largely adopted Western dress, they also have to adopt the Western closet. In societies where some version of a loose gown is the traditional costume for both men and women, this can easily be gathered up around one's waist, out of the way, leaving the legs free. Western clothing, in contrast, almost invariably involves some form of under or outer garment—trousers, shorts, girdles, panty hose, and so on—that must be lowered around one's legs or else removed entirely. If simply lowered, as is now customary, these garments would interfere with the assumption of a properly balanced stance, which involves spreading the feet. Furthermore, they would be in danger of being soiled in the fixture or by one's own actions. Urination, for example, would be impossible in this situation. In addition, one is liable to lose the contents of one's pockets, as many unfortunate souls can wryly testify. In short, the only satisfactory way to use the traditonal form of hole-in-the-floor squat closet is to disrobe completely from the waist down. Given our accustomed habits and in light of the practical, not to mention the psychological, difficulties,



32 / POSSIBLE ROTATIONS OF SQUAT POSTURE

it is probably reasonable to conclude that the adoption of the squat closet is not a proposition that can be seriously entertained in the present context. From a design viewpoint, however, it possesses the one great virtue of being foolproof: it simply leaves the user no choice except to assume the proper posture.

If we leave aside a full squat posture and the fixtures based on it, there are still, however, a number of other modified approaches to the problem to be considered, as shown in Figure 33. These range all the way from a modified squat closet to a modified seat for use on existing water closets. In each instance the premise is that a reasonable approximation of the desirable posture can be achieved. The problem is to make appropriate modifications that will actively encourage the assumption of such a posture with reasonable ease and comfort. While we recognize that none of these modified approaches are foolproof, in the sense that an uninformed or determined user could still ignore the proper posture, they do promise to find wider acceptance and to be of benefit to a larger population.

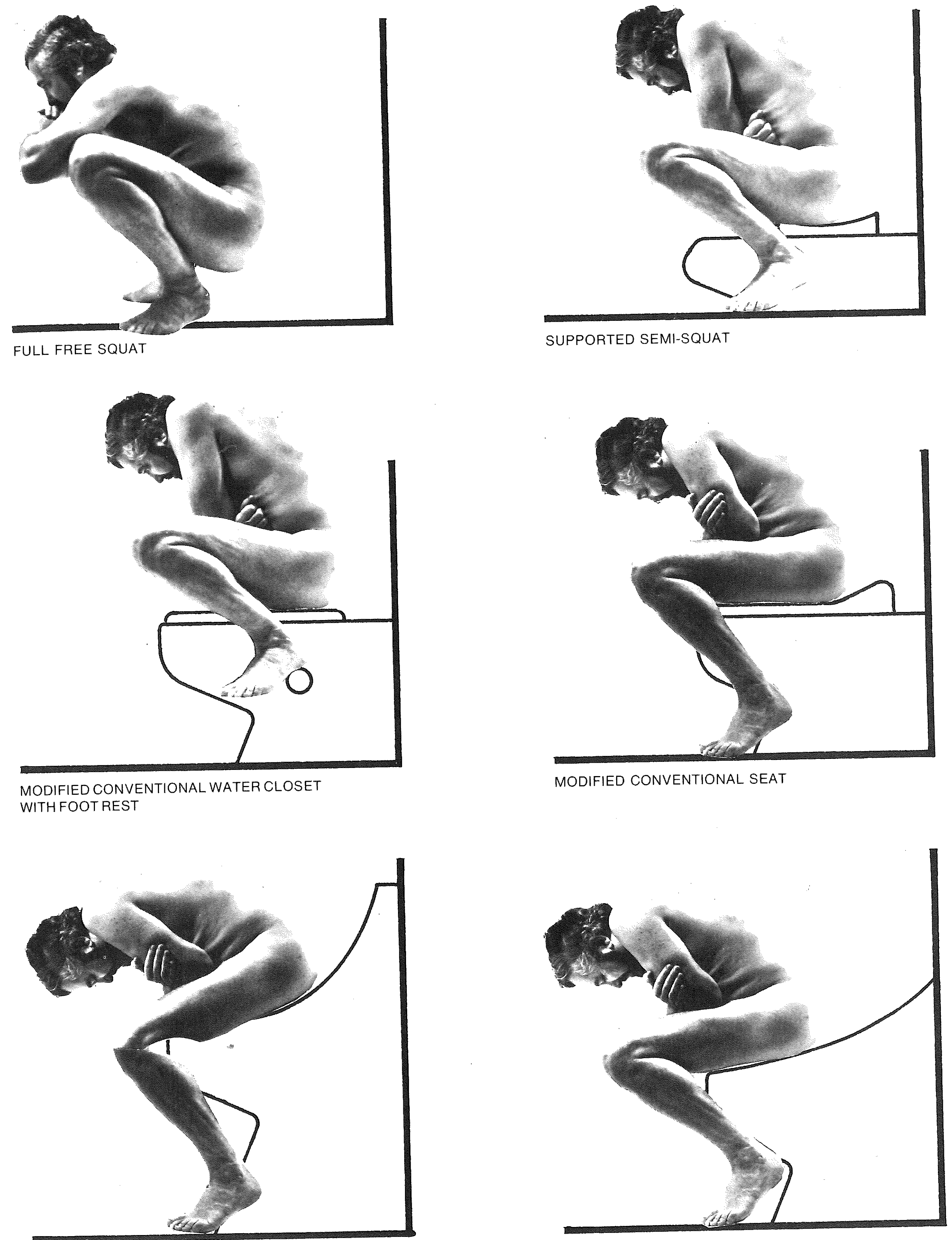
Modified Squat Closet When we assume a full free-squat posture, the degree of that squat is very much a matter of body size, agility, joint flexibility, and so on. Most Western people unaccustomed to the posture cannot begin to achieve the position commonly achieved by most primitive peoples or by people who are accustomed to the posture. In this connection, age appears to have very little, if any, relationship to one's ability to assume the posture, as the experience of primitive peoples testify, though in the Western world one of the commonest problems of elderly people is their inability to cope with low seating.

In the assumption of a full squat, the height of the buttocks from the floor varies from 150 to 205 mm (6 to 8 inches), and though these heights can be reached with reasonable comfort, they place a considerable strain on the calf muscles and are difficult to rise from, particularly because of the problems of balance. This is true even when support is available at those heights. In part this is because the legs need to be spread far apart in a flat-footed stance so that the trunk can lean far enough forward to maintain balance. Rising from such a position is complicated because the entire body must be brought into play in order to maintain equilibrium. Put another way, there is a considerable horizontal force component to be overcome.

A somewhat modified squat, which we are more accustomed to assume, ranges in height from 230 to 280 mm (9 to 11 inches) and involves balancing on the balls of one's feet. In this position, lowering and raising the body is considerably easier, as is the problem of maintaining balance, since the force exerted by the legs is essentially vertical and the distance to be traveled is less. In the assumption of either posture, however, that posture must be assumed initially by straddling the fixture or support and then lowering and raising the body from that position. It is exceedingly difficult first to sit, as we do normally, and then pull one's legs back or, conversely, to undo the posture once one has assumed it. For all practical purposes a height of approximately 255 mm (10 inches) would represent a workable figure for this approach.

One of the most important design consequences of assuming any such squat or semi-squat posture is that the conditions of support are drastically different from a normal sitting posture, requiring a radically different seat or supporting structure. The essential difference, illustrated in Figure 34, is that in a squat posture only "point" support is possible and this support is at the ischial tuberosities, the bottom-most protuberances of the pelvic structure. If one sits normally on a standard seat, the bulk of the body weight is placed on the thighs. Sitting on a standard seat in a squat position is acutely uncomfortable and causes one to fall partway through the seat opening, since the sole support for the body is at the sides of the buttocks and on unaccustomed points of the pelvic structure. For proper and comfortable seating,

33 / POSSIBLE APPROACHES TO A MODIFIED SQUAT POSTURE

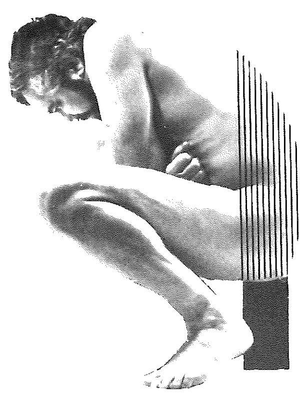


therefore, the seat must be designed to offer support directly to the tuberosities. Since the ischial tuberosities are quite close together, this design requires a narrowing at the midpoint of the seat opening, a radical change from the shape of the present seat.

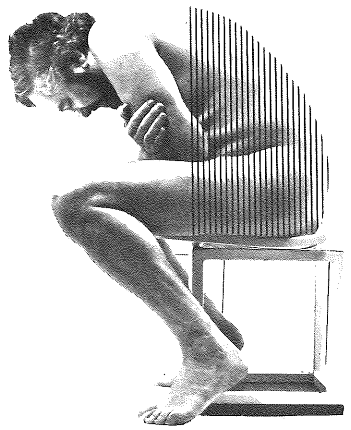
Point support of this sort necessarily involves, however, a greater concentration of force than is normally the case. Some preliminary experiments on normal seating and buttock loads conducted by Hertzberg at the Wright Air Development Center indicated "that loads under the tuberosities can run in certain body types to as high as 4,218 g/cm² (60 pounds per square inch), and perhaps higher. . . ."[1] Figure 35 illustrates the approximate pressure pattern that develops. With respect to the modified squat closet the loading is, however, nowhere near as great, primarily because of the different distribution of weight that results from the squat posture. As we saw in Figure 34, a semi-sitting squat results in less than half the total body weight's being supported by the seat, compared with 70 per cent or more when one is sitting normally. Nevertheless, it is apparent that to provide the greatest comfort, the seat should still be shaped in such a way as to distribute the total load over a somewhat greater area than would be possible by a perfectly flat seat.

This support of the tuberosities is in many ways preferable to the kind of support provided by conventional seats. With a conventional seat the body weight is largely borne by the thighs, often in an unsatisfactory manner, as we shall learn in more detail later. In this position, tissues are compressed, and, in addition to some potentially injurious effects to muscles and nerves, considerable discomfort results. A very common complaint, particularly among the elderly and those with constipation problems, is that their legs go numb from having the circulation cut off. Consequently, people are inclined to sit far forward in their seats in

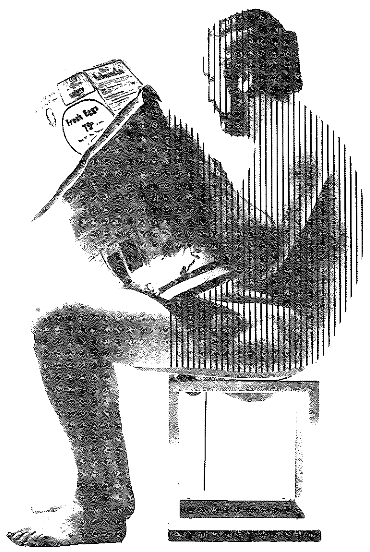
34 / VARIATIONS IN SUPPORT AND WEIGHT DISTRIBUTION BY POSTURE ASSUMED



POINT SUPPORT: 40–50%



BUTTOCK SUPPORT: 60–70%



BUTTOCK AND THIGH SUPPORT: 70–75%

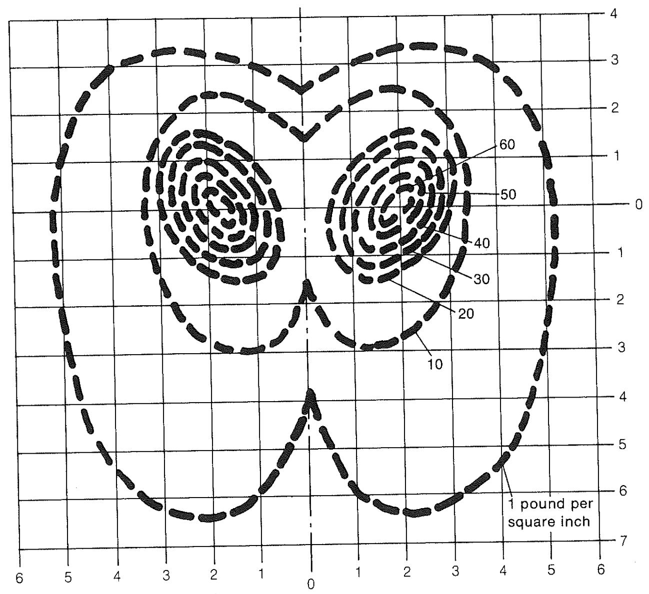
order that the ischial tuberosities and not the thighs may support their weight. A better solution is to modify the design of seating:

> To avoid undue pressure on the sensitive soft parts of the thigh, the chair must be so constructed that the weight of the body is borne on the ischial tuberosities. The thigh should be able to hang freely or only rest gently on the seat. The height of the seat should therefore be less than the length of the lower leg. If the chair is low enough to suit those with short legs, these conditions will also be fulfilled for those with longer legs.[2]

In considering the actual dimensions for such a support and its related opening, we must turn to the anthropometric measurements illustrated in Figure 36, which show that the ischial tuberosities are approximately 25 mm (1 inch) in width and 40 to 50 mm (1½ to 2 inches) in length and vary in span from 120 to 160 mm (4¾ to 6¼ inches), females tending to have a larger bi-ischiatic diameter than males, generally by about 20 mm (¾ inch). This suggests that the maximum opening on the transverse axis of the tuberosities can be no greater than 75 to 90 mm (3 to 3½ inches) since, for comfort, the tuberosities must be fully supported. Partial support, so that the pelvic structure is supported at the sides, is acutely uncomfortable.

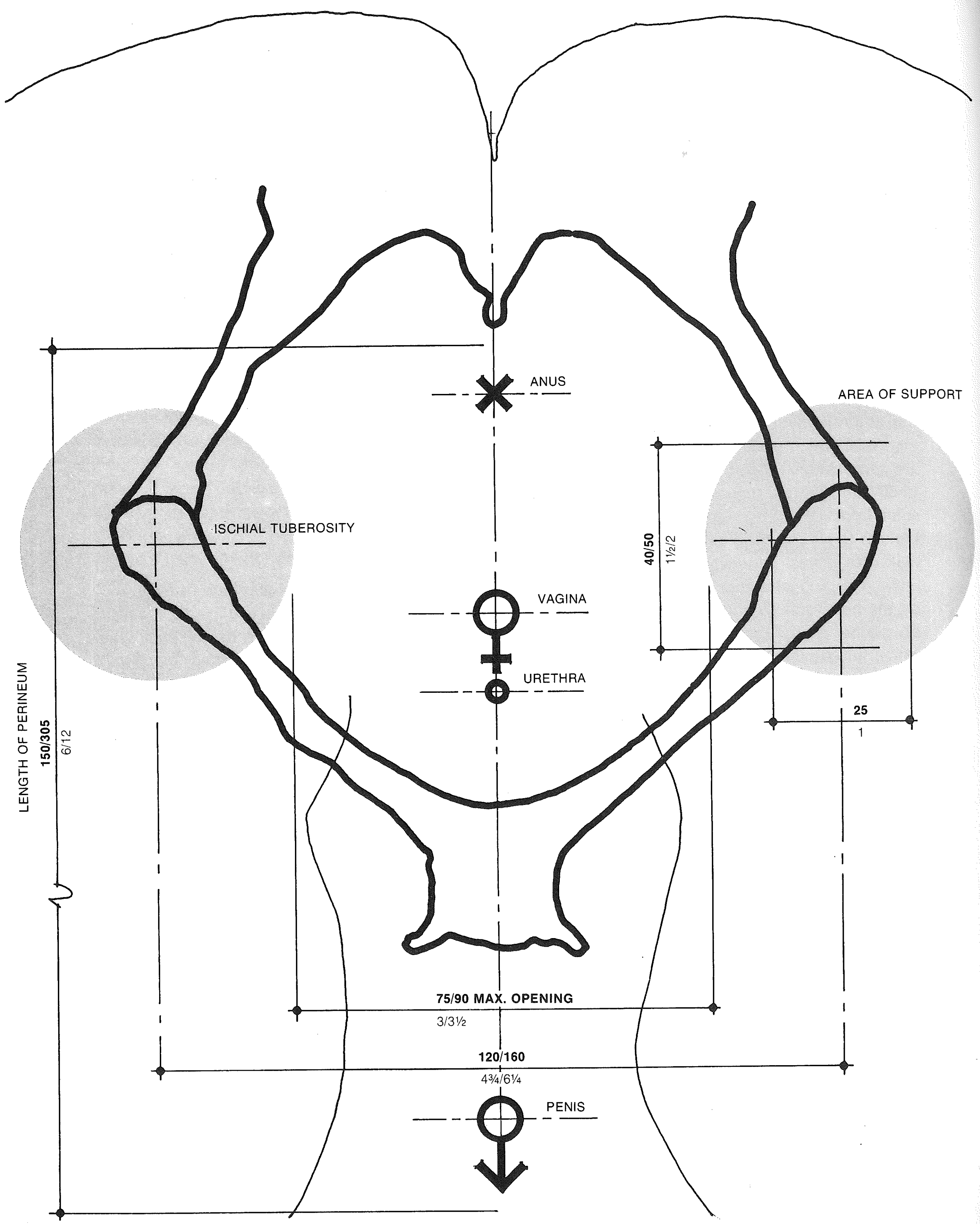
It will be noted from Figure 36 that both the anus and urethra lie within what must necessarily be the narrowest point in the seat opening if the tuberosities are to be supported. Although extremely narrow, compared with the average 205-mm (8-inch) width of opening in present seats, this opening width of 75 to 90 mm (3 to 3½ inches) is sufficient to prevent soiling the seat, particularly since this proposed type of seat has the effect of "locking" one into position. While minor movement is possible, radical changes of position from the one intended are not really possible without considerable discomfort.

Other considerations necessitate, however, enlarging the opening both in front and back of the point of actual support. One is for purely psychological purposes—to remove the fear of soiling the fixture. When the openings in front and back appear more or less normal in size, the user is likely to feel more secure in this respect. One doesn't realize that the anus is actually located at the narrowest point. The practical, and more important, reason for enlarging the opening is to provide room so that the hand can reach in to wipe both the anus and urethra, since wiping the anus must be done from the back instead of the side as is common on conventional water closets. In a squat or semi-squat posture it proves exceedingly difficult to perform this operation from the side because one cannot shift one's weight so as to reach under from the side. For this, clear openings of approximately 150 to 205 mm (6 to 8 inches) in diameter are necessary both front and back. Another reason is to keep the penis from touching the fixture. In this respect, most conventional water closets are woefully inadequate. This is particularly true of the standard round or egg-shaped bowls in which the front-to-back dimension is usually so small as to necessitate a delicate balance in finding the proper location on the seat—to avoid soiling the back of the seat and to



FIGURES REFER TO INCHES

35 / PRESSURES DEVELOPED IN SITTING



36 / ANATOMICAL PLAN VIEW OF PERINEAL REGION

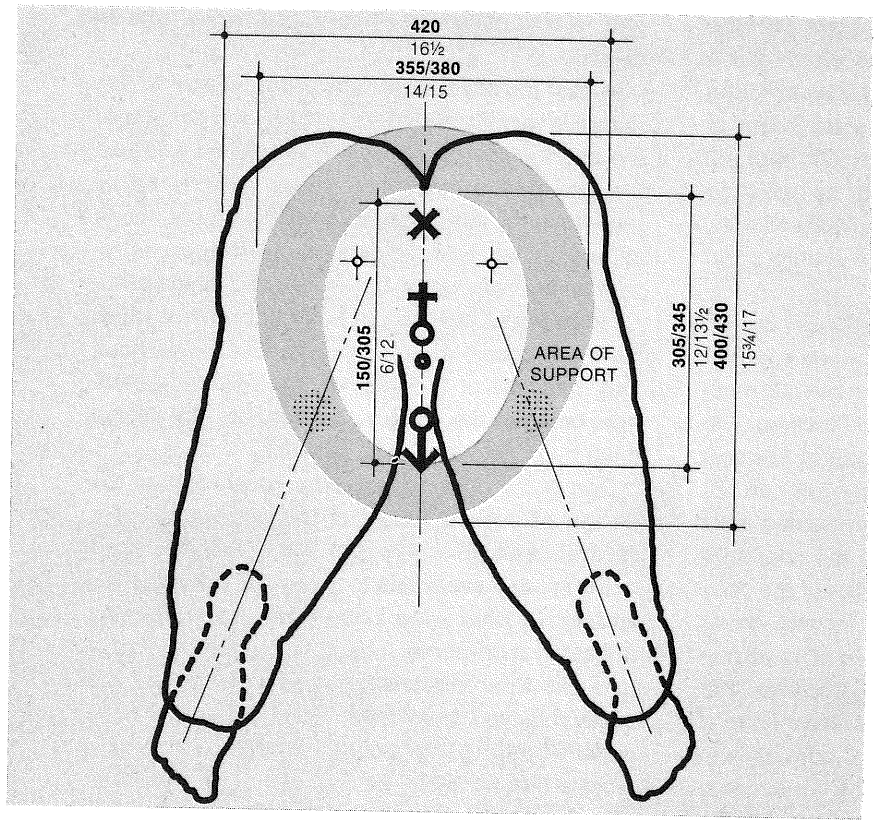
avoid touching the front lip of the bowl with the penis. As indicated in Figure 36, the length of the perineum from the cleavage of the buttocks to the front of the genitals ranges from a minimum of approximately 150 mm (6 inches) to 305 mm (12 inches).[3] Conventional water closets, as indicated in Figure 37, often meet these dimensions exactly, with no room for error, wiping, or a slightly erect penis.

As suggested in Figure 38, the minimum overall inside front-to-back dimension should be approximately 455 mm (18 inches) and the overall minimum width approximately 205 mm (8 inches). The height from the floor to the low point of the seat can be established at approximately 255 mm (10 inches). In order, however, to contain seated male urination properly, the front lip of the container should be raised higher than the level of the seat. A sitting height of 255 mm (10 inches) being assumed, the front lip could be raised 50 or 65 mm (2 or 2½ inches). At the same time, however, that the front lip is raised, it must also be lengthened in order that the thighs be kept free of contact with the fixture. The vertical profile of the container is directly related to the posture assumed and to the height of the seat—largely because of the problem of clothing. In this regard, the distance from the center of support to the front of the container should not exceed 305 mm (12 inches) in order that clothing gathered around the knees be kept clear of the container. At this 305-mm (12-inch) distance the front lip should not be more than 150 mm (6 inches) wide in order that the fixture may be comfortably straddled. The front of the fixture must also taper back both vertically and horizontally in order to accommodate the desirable posture. In addition (on the assumption that such a fixture would be floor-mounted), the support for the fixture must be cut back a minimum of 150 mm (6 inches) to accommodate the feet—both while one is mounting the fixture and while one is seated.

In this instance, particularly, the shape of the fixture is very largely determined by the posture of the user, which, of necessity, is relatively inflexible, and by the basic dimensions established by this posture. Although these recommendations should not be regarded as absolute, they do represent an approximation of the size and shape criteria to be met. Obviously, a number of minor variations are possible, but it should also be obvious that, because of the high degree of interrelatedness between the various components, the design must be considered in its entirety.

One of the possibilities to be considered in the design of such a fixture is the elimination of a separate seat. In this regard, the primary purpose of the separate seat was to provide a warm seating surface in what were once upon a time largely unheated bathrooms. Indeed, some hardy individuals like Churchill are reported to have had the seats removed from their water closets, a taste acquired during their public school days when the closets had no seats. Be that as it may, the separate two-piece seat and lid represent a major cleaning problem and could be considered superfluous in these days of heated bathrooms. (In Japan, where central heating, or heating at all, is still rather rare, one of the most popular luxuries is an electrically heated toilet seat.) An additional justification for eliminating the separate seat is that the actual contact area is so small. If, however, a seat were felt to be necessary from the viewpoint of user comfort, it might well be no more than two permanently affixed pads of a material warmer to the touch than china. Two contemporary fixtures designed to be used without a seat are shown in Figure 39.

Modified Conventional Closet Referring back to Figure 33, we can examine several other possibilities with reference to modification of a conventional water closet. Essentially, these approaches are based upon the premise that, while a squat or semi-squat posture is desirable, it poses major difficulties in getting up and down relative to a low seat height. It would appear obvious,



37 / PLAN VIEW OF SITTING POSITION ON A CONVENTIONAL WATER CLOSET AT A CONVENTIONAL 380/400-MM (15/16-INCH) HEIGHT

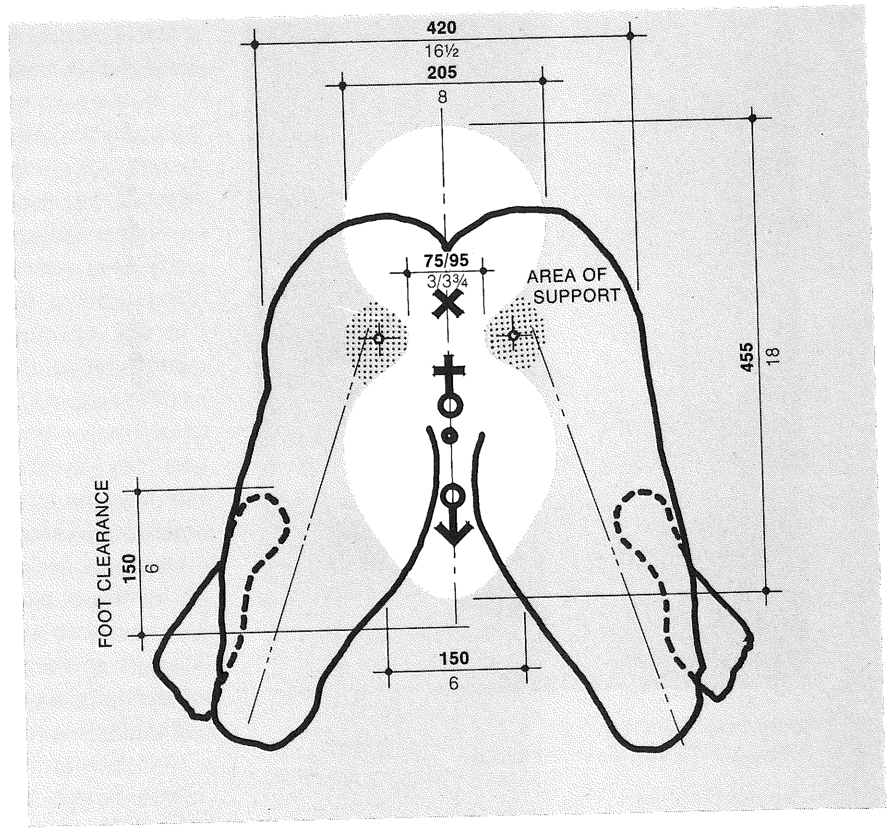
therefore, that, by beginning with a conventional-height water closet and equipping it with some form of footrest so that one could assume a squat posture after being seated, one could resolve both problems. Indeed, several such fixtures have been tried over the years, as illustrated in Figure 40. Unfortunately, however, the matter is not so simple, as we shall see in a moment.

If we assume a conventional fixture approximately 405 mm (16 inches) high equipped with integral footrests at a height of 205 mm (8 inches), similar to the illustration in Figure 40, we have theoretically established the proper relationship for a full squat as described earlier. It proves to be virtually impossible, however, to assume such a squat posture after one is once seated—except perhaps for a trained gymnast. In order to assume the posture and lift one's feet up into position, it is necessary first to lean back, bring up the feet, position them properly on the footrests, and then try to rise from a leaning-backwards position forward to a balanced squatting posture, without tumbling over forward in the process. Even when possible to execute, this maneuver is more awkward and strenuous than getting up and down from a simple squat posture, particularly in terms of balance.

Clothing again provides a further complication. With one's feet effectively bound by lowered clothing one must either try to lift both legs simultaneously or else disrobe in order to lift one leg at a time. Shoes aggravate matters even more since the feet must be lifted higher in order to hook the heels behind the footrests.

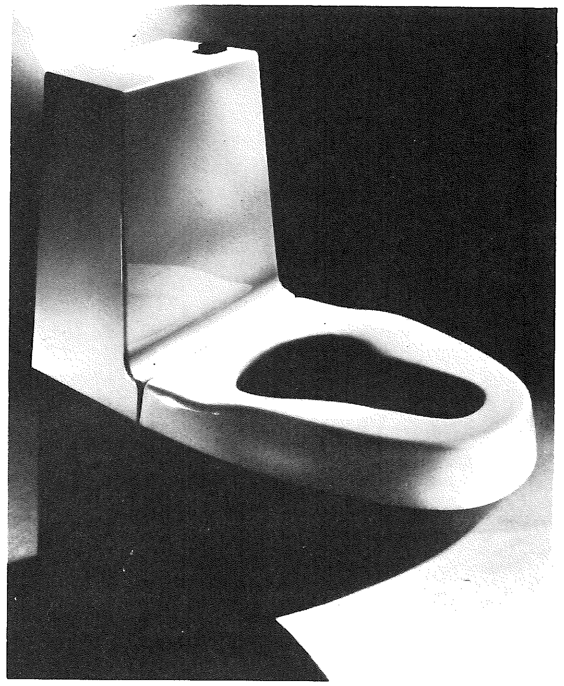
Another approach in this situation is not to assume a horizontal squat position but to remain

38 / PLAN VIEW OF NECESSARY DIMENSIONS AND CLEARANCES FOR A SEMI-SQUAT WATER CLOSET AT 255-MM (10-INCH) HEIGHT



in a rotated 45 degree squat position. This, in some ways, is similar to the position assumed, for example, when one is using some temporary form of footrest in conjunction with a conventional closet, as has sometimes been recommended, or when one is using some of the "health closets," which are based on the idea of simply having the knees higher than the pelvis. Theoretically these are all based on the assumption, as illustrated in Figure 32, that once a tight squat posture is assumed it can be rotated. This is true but only within limits. The difficulty with many of these rotated positions is that they are rendered ineffective as a defecation posture because one's energy and musculature are working to maintain one's balance and cannot at the same time be employed for expulsion, nor can one relax the anal sphincter while in a strained posture. In order for the posture to be effective one must first be relaxed, comfortable, and properly balanced; only then can one fully bring into play the musculature for defecation. The most important relationship is between the trunk and the thighs; the positioning of the legs is of secondary benefit.

One rather extreme way in which such a rotated posture might be made to work is to provide a full backrest that would effectively lock one into position so that maintaining one's balance and posture is no longer a problem (see Figures 41 and 32). In such a case, one would not only be locked into the proper posture but would also be free to devote one's efforts to the task at hand. For greatest ease of assuming such a posture, the backrest could be made adjustable so that one would sit, lean back, place one's feet on the footrests, and then raise the back, bringing one into proper posi-



"ROHAGLAS" WATER CLOSET:
RÖHM, DARMSTADT, WEST GERMANY

"ARMADA" WALL-HUNG WATER CLOSET:
VILLEROY-BOCH, METTLACH, WEST GERMANY

tion. Conversely, the backrest could be fixed and the footrests made adjustable.

In the assumption of any such drastically rotated position, the conditions of support change once again—this time to the coccyx, or to the ischial tuberosities, depending on the degree of rotation. While this can be resolved in the case of any new fixture based on the possible approach just outlined, it is one more problem with respect to any attempt to assume such postures on a conventional closet equipped with a conventional seat designed for a flat horizontal sitting posture. Rotation of the body also means that the body orifices are rotated relative to the horizontal plane so that the front of the bowl would have to be raised and shaped to contain the expelled body wastes properly. In this respect, there is obviously a limit beyond which such rotation ceases to serve any purpose and becomes a hindrance. Wiping can, moreover, become difficult if one cannot shift one's body weight to one side or if the opening is not sufficient for free hand access.

Another variation, shown in Figure 40, uses footrests that are sloped back and up from the floor, allowing one to work one's feet into the proper position without the strain and difficulty encountered previously. If, however, the footrests are too far forward, one is faced with much the same problems as before—either of adjusting one's posture once seated or of remaining in a rotated and strained position. Conversely, if the footrests are sufficiently far back so as not to involve a postural change once seated, they do offer some support or resistance to the feet, but once one has pulled one's feet that far back under the body the footrests are largely superfluous. In either case, they interfere with the placement of the feet if we assume that the water closet will continue to be used for standing male urination. In addition, the determination of a suitable angle is complicated, if not rendered insoluble, by the dif-

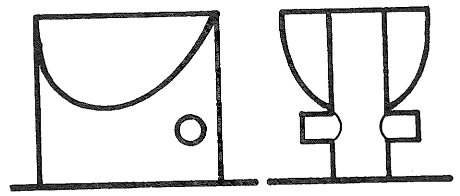
39 / CURRENT APPROACHES
TO WATER CLOSETS

ferences between what would be optimal for a bare foot and what would be suitable for a woman wearing high-heeled shoes.

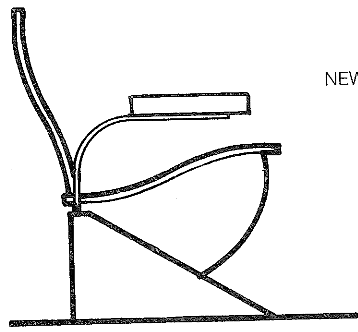
In virtually every instance, the adaptations examined are sufficiently drastic in terms of altering all the normal relationships between the body and the fixture that they cannot be seriously considered as workable approaches. The complexity of these relationships is such that it is necessary to begin from the beginning.

The last and least drastic, not to mention least foolproof, approach, suggested in Figure 42, is to provide an improved seat for use on existing conventional closets. The premises underlying this approach are that a reasonable approximation of a semi-squat posture can be assumed even on a 405-mm (16-inch)-high conventional closet and that a properly designed seat can assist in the comfortable assumption and maintenance of that posture. While this is obviously a compromise, it does have the merit of permitting existing installations to be more effective and comfortable. Most conventional seats (and closets) have two major shortcomings: a front-to-back opening dimension that is too small for hygienic use and a configuration based more on appearance than seating comfort. Although one obviously cannot stretch existing fixtures, it is possible to lengthen substantially the actual effective opening, which is determined by the seat rather than the bowl. In the common range of residential water closets, particularly European ones, the front-to-back dimension of the bowl opening varies from 305 to 345 mm (12 to $13\frac{1}{2}$ inches). Almost all seats have, however, a considerably wider rim than the rim of the bowl and overhang the inside of the bowl by 20 to 40 mm ($\frac{3}{4}$ to $1\frac{1}{2}$ inches), thus further reducing the effective opening to as little as 255 mm (10 inches) in extreme cases. By cutting the rim away and eliminating these overhangs at the front and back, we can at least have the maximum opening possible on any given closet, even though it still falls short of the 370 to 400 mm ($14\frac{1}{2}$ to $15\frac{3}{4}$ inches) desirable.

With respect to configuration, most often one finds seats that have a doughnut-like constant



POSTURE CLOSET BY LE CORBUSIER
FOR POZZI, MILAN, ITALY



POSTURE CLOSET
BY J. GARDNER
FOR ADAMSEZ, LTD.,
NEWCASTLE-UPON-TYNE,
ENGLAND

40 / APPROACHES TO SEMI-SQUAT
WATER CLOSETS

convex cross-section or, worse yet, a sharply tapered constant cross-section (see Figure 42). Such configurations obviously have little to do with seating. They are essentially decorative, and, as we are seeing increasingly, such seats have become a major fashion item for the bathroom. Because of one's seated posture, however, as shown in Figure 34, such configurations provide only point support for the thighs, where some 70 per cent of one's body weight is concentrated. In the bolt-upright posture this is not so serious since the buttocks receive support as well, but the more one leans forward and pulls one's feet back to assume a semi-squat posture, the more one concentrates one's weight on the thighs. In fact, the moment one leans forward at all—enough, for example, to part the buttocks as one must—one has shifted the weight to the thighs. When that happens, the weight is essentially concentrated on two pivot points approximately midway along the femur with the familiar result that one then complains about one's legs' "falling asleep" from having the blood circulation cut off. The possible pressures at those two pivot points can exceed those encountered under the ischial tuberosities.

In order to accommodate such a seated posture properly so that the weight is borne by the thighs, it is necessary to maximize the area of support. This can be accomplished by providing a constant and slightly concave cross-section all along the axis of the thighs, as shown in Figure 42. In some respects, the overall configuration would be much like a cut-away tractor seat with a complex continuously variable cross-section based on the sitting posture rather than the symmetry of the bowl.

In order to encourage further the assumption of a doubled-over semi-squat, the back of the seat might be raised slightly (rather than lowered). With a high back and a slight forward tilt the seat is somewhat awkward and uncomfortable to use if one attempts to sit on it in a bolt-upright position but is very natural in the proper position. Such a high-back seat could also incorporate the mechanism necessary for a washing/drying function within the back seat cavity.

One new seat based on these principles, designed in this case for use on existing elongated closets, is shown in Figure 43. Obviously, similar seats can be designed to accommodate other sizes and configurations of conventional closets. Seats can also be designed to accommodate other approaches such as the proposal for the ultimate reading seat shown in Figure 44.

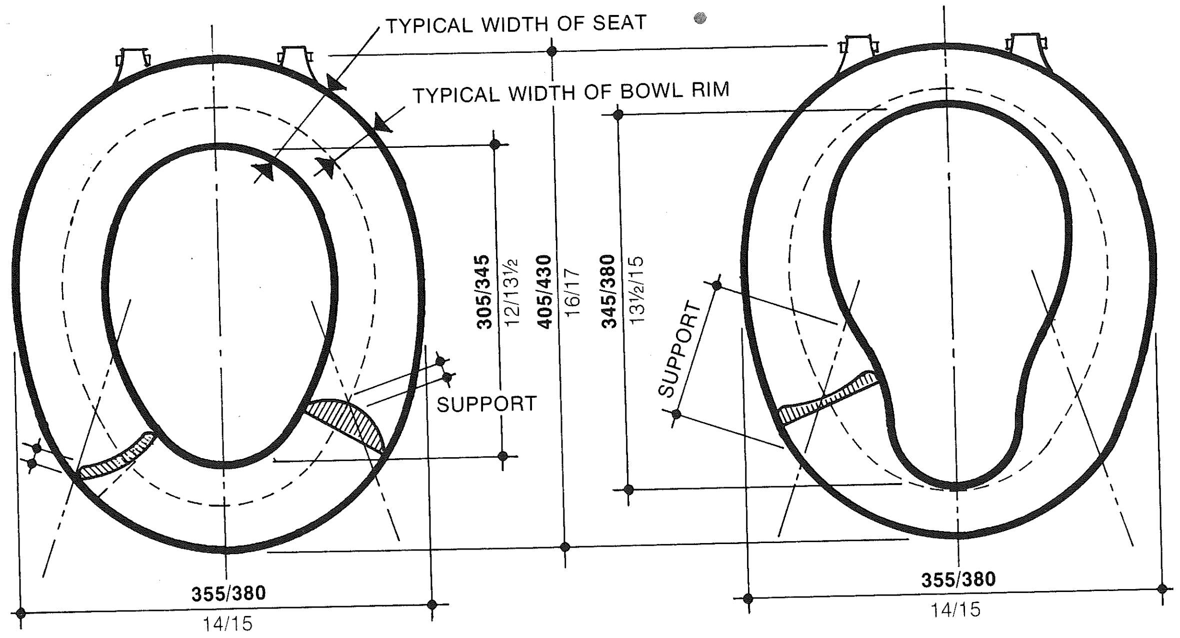
High-Rise Closet The approach just described could also be applied to the design of an entirely new fixture as well, as suggested in Figure 33. In this case, the features could be incorporated in optimal form into the fixture itself and, as suggested earlier, the seat itself might be eliminated altogether (see Figure 45).

Another variant on this approach would be to raise the point of contact sufficiently so as to require a lean-on posture, to approximately 610 mm (24 inches), as suggested in Figure 33. In this case the only way the fixture could be used would be in the doubled-over position, which is neces-

41 / POSSIBLE APPROACH
TO A FORCED-SQUAT POSTURE

sary in order to remain in place. Such an extreme approach poses problems, however, with the management of clothing, which must be dropped to the floor, proves tiring after a short while, and thereby violates the principle of a relaxed posture. In addition, it is unsatisfactory for urination from a seated position unless there were a pronounced projecting front lip similar to some of the early female urinals, which, in fact, such a fixture would closely resemble, though it would serve admirably for standing male urination. On balance, however, this approach does not appear as promising as others reviewed.

Anal and Urethral Hygiene The techniques for anal wiping will vary according to the posture assumed on a given fixture—from either the side or the back. Urethral blotting is simply done from the front regardless of posture. In either case, however, the ideal location for the toilet tissue is in front of the seated user where it can easily be reached and folded. A side-wall location is acceptable if the dispenser is located just even with the front of the fixture. Unfortunately, all too



42 / POSSIBLE MODIFICATIONS OF
EXISTING WATER CLOSET SEATS

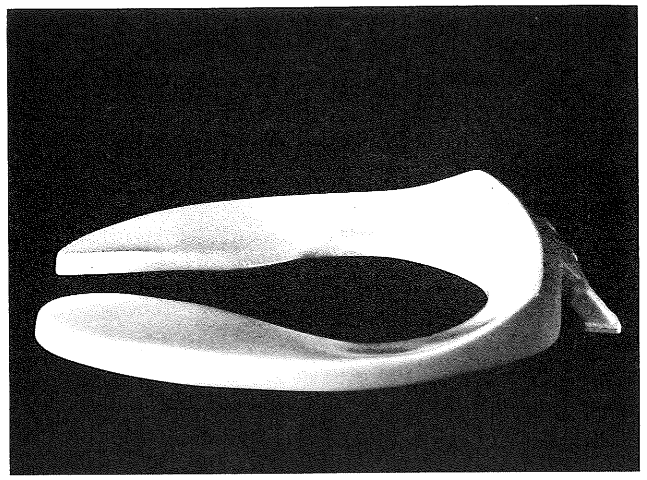
often it is located behind the user, necessitating considerable contortions to reach it. Indeed, it is curious that we persist in regarding paper-holders as "accessories," an attitude that implies they are optional rather than essential.

For proper cleansing the anus should, at the very least, be wet-wiped if not actually washed. Most commonly, such washing is done at the bidet. For maximum convenience, however, such cleansing would logically take place as a natural part of the wiping process while one is still seated on the water closet. The most direct approach to this is to build a washing and drying function into the water closet itself, as is illustrated in Figure 46. This requires a fairly sophisticated design: separate controls, tempered water, an aerated stream, a precise and dependable stream trajectory, a warm-air drying jet, and other features. The precise dimensions and mechanics involved are obviously determined by the posture assumed and the precise configuration of the fixture in question, and they cannot be detailed further here. Another variant on this approach is to build such functions into the toilet seat, which offers the advantage of use on existing water closets. Several such seats are on the market in the United States and in Europe.

Several early models of water closet used a washing jet relying only on water from the tank. If the closet had not been used for some time and the water was at room temperature the result was reasonably satisfactory, but when fresh cold water from the mains happened to be used, the result was a very pronounced shock to the system. No such fixtures are on the market any longer.

In some circumstances it is possible, of course, to wet the tissue by using another nearby water source such as a lavatory or a bathtub, which is the stratagem used by most hemorrhoid patients. It would be far preferable, however, to make this operation an integral part of the elimination and hygiene processes by making it as convenient and natural as possible.

For the benefit of the sceptics or the guilt-ridden, note that the sensation resulting from our common habit of dry-wiping—once one has become ac-



"POSTURE MOLD" SEAT: AMERICAN STANDARD, NEW BRUNSWICK, NEW JERSEY, U.S.A. (COLLECTION THE MUSEUM OF MODERN ART, NEW YORK)

43 / CURRENT APPROACH TO A POSTURE WATER CLOSET SEAT

customed to washing—has been compared by some to the sensation resulting from not brushing one's teeth from one week to the next.

One other relatively simple and adaptable approach is to use one of the proprietary cleansing foams (commonly marketed as hemorrhoidal preparations) with regular tissue, a usage that would require only a new combination holder for the foam and the paper. Most such preparations have the additional advantage of containing some lanolin-like substance and consequently are extremely soothing, almost voluptuously so. Still another variant is the use of packaged prewetted papers, again commonly marketed for hemorrhoidal use.

Odor and Noise Privacy in the bathroom, particularly for elimination functions, is a major concern for some people. All manner of verbal and social conventions are resorted to in an attempt to disguise or mitigate our activities, such as running water to mask the sound of other running water. While such stratagems are more or less successful, they all ultimately fail with respect to telltale odors—a source of considerable concern and embarrassment for many. Probably the oldest, and still the commonest, approach to this problem is to use some other, presumably more pleasant odor to mask the offending odor. As a solution this is rarely satisfactory: either the chemistry of the combination is unfortunate, or the masking odor is so strong as to be repellent in its own right, or the masking odor becomes so familiar, as in public urinals, as to be equally as offensive as the original.

A more recent though actually quite old, and certainly more satisfactory, approach is to use some form of positive chemical or hydromechanical exhaust system. There are a number of commercially available systems today that perform satisfactorily. Some are built into the water closet and use a syphonic action to induce venting of the bowl itself. Others rely on a mechanical exhaust to the outside and still others use a chemical, generally charcoal, filtration system.

One of the more promising recent approaches employs a chemical attack on the problems not only of odor but also of noise. It is based on the deployment of a complex chemical foam blanket on the pan water, which reduces the noise and splash back of urination and of defecation and is a deodorant, a germicide, and a cleansing agent.[4]

Controls If we examine the matter of controls from the viewpoint of user comfort and convenience rather than of mere mechanical convenience, it becomes readily apparent that most flushing mechanisms are poorly located. The most common arrangement in the United States, for example, is a lever handle mounted on the top lefthand face of the tank. This is convenient only if the user flushes the closet after rising and turning around. A sizeable number of persons prefer, however, for one reason or another (odor, peace of mind, and so on), to flush the closet while seated and after each bowel movement and must engage in some contortions to do so.[5] Since the water closet is presently also used for standing male urination, this might be regarded as a justification for its location. Conversely, many of the modern low one-piece water closets have a flushing mechanism that can be readily reached from the seated position but that is awkward to reach from a standing position.

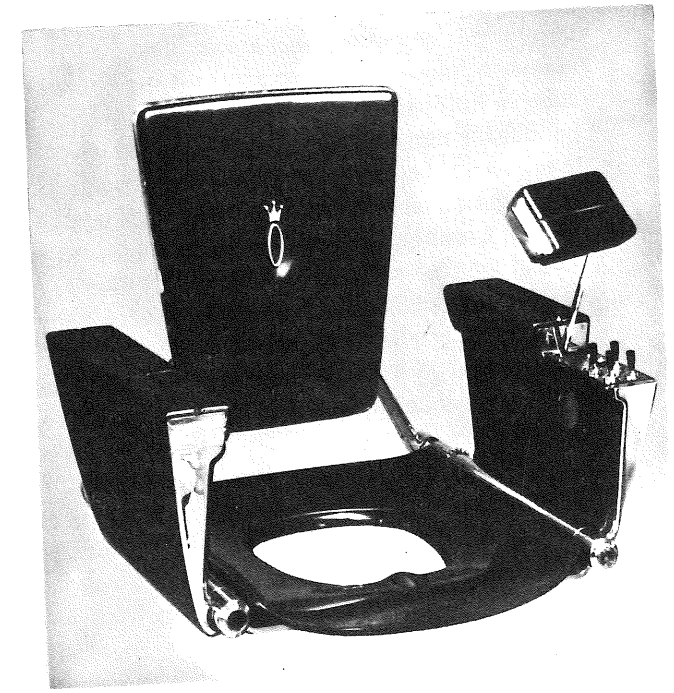
Throughout much of Europe the most common arrangement is a pull-up knob located on the top of the tank. While this has the virtue of mechanical simplicity, it effectively precludes use from any but a standing position. It also has the not inconsiderable drawback of negating the use of the space over the tank for much needed storage or for a counter.

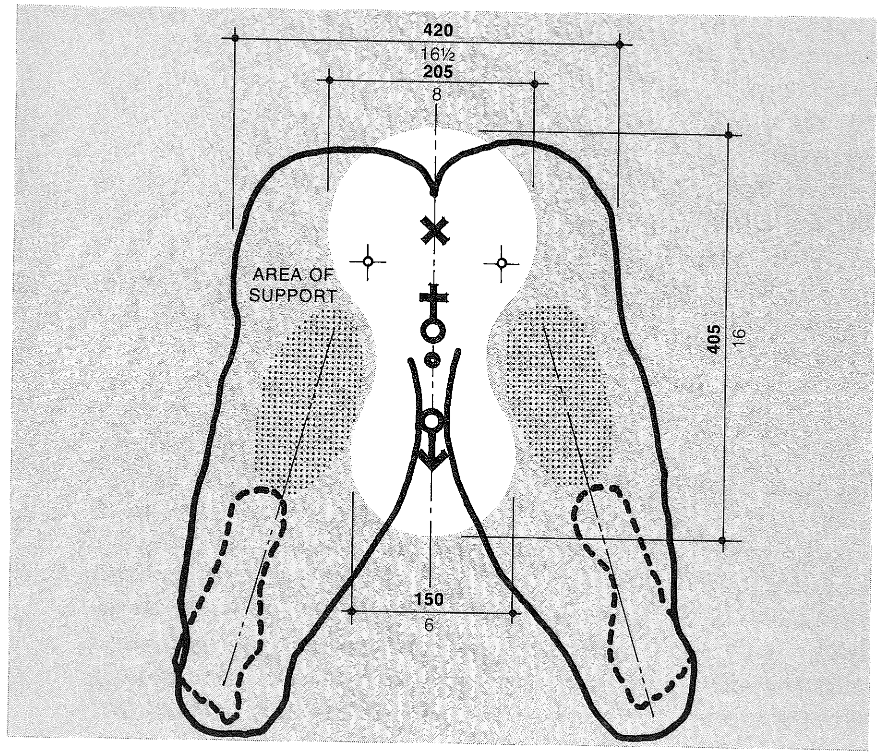
Although most commonly found in public installations, possibly the worst arrangement from a functional viewpoint is the ambiguous positioning generally used for flushometer installations where the flushing lever is approximately 455 mm (18 inches) off the floor and where it is not clear whether the device should be operated by hand or by foot and where either method is ultimately awkward and unsanitary.

More suitable arrangements are, of course, possible and have, in fact, been available for many years. In 1891, Thomas Crapper patented a "Seat Action Automatic Flush," for example. Another venerable solution with much to recommend it is the use of a spring-loaded flushing button set into the floor just in front of the fixture, an arrangement that permits use from either a standing or sitting position, before, during, or after elimination. Contemporary versions of this floor-mounted button now operate electronically rather than

44 / APPROACH TO A WATER CLOSET READING SEAT

"AD 2000 COMFORT CONTROL CENTER": OLSONITE, DETROIT, MICHIGAN, U.S.A. (FEATURES INCLUDE READING LIGHT, ASHTRAY, RADIO, T.V., TIMER, TILTING MASSAGE BACK, BIDET)





45 / PLAN VIEW OF NECESSARY DIMENSIONS AND CLEARANCES FOR A HIGH-RISE WATER CLOSET

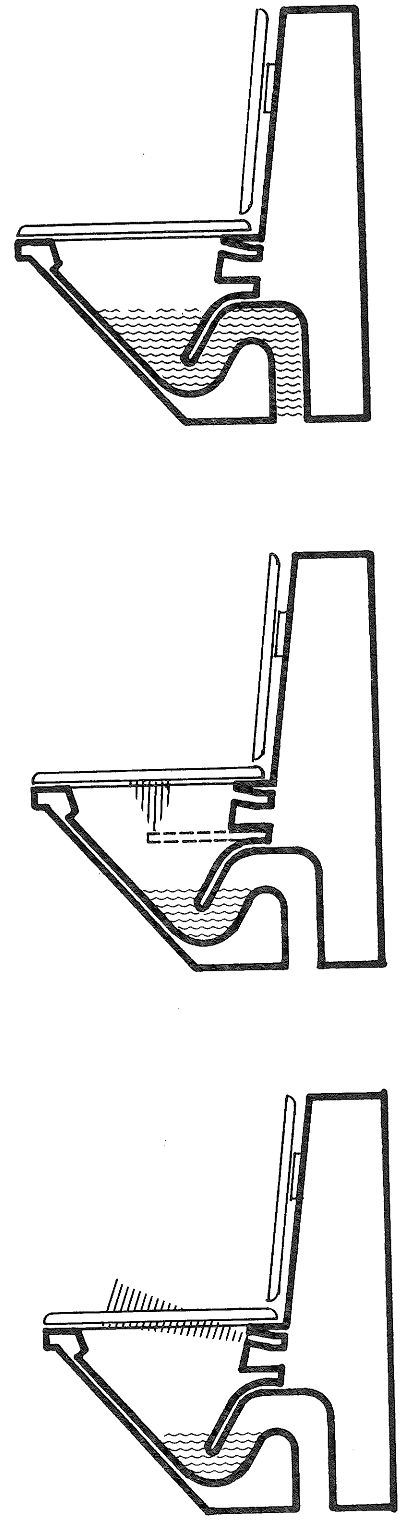
through a complex mechanical linkage. The increasing use of electronic control devices, especially in Europe, is extremely promising in that it permits the triggering switch to be located virtually anywhere that is appropriate without regard to where the actual flushing mechanism itself may be located. These new devices are also a much cleaner and simpler solution than the "electric-eye," which has been in use in public facilities for many years. Another effective approach to this problem is illustrated in Figure 46. In this instance the flushing mechanism is activated by the pressure bar extending horizontally across the face of the tank, an arrangement that permits operation by leaning back or hitting the bar with an elbow from a seated position or with a hand from a standing position if desired. This approach is particularly well suited for use by the elderly and by the disabled.

When the fixture is also equipped with built-in washing and drying functions, odor control devices, etc., the same criteria should apply: that the controls be easily accessible from the posture of primary use, that they be easy to identify and distinguish one from another, and that their operation be apparent.

Soilability and Cleanability Although cleaning is not strictly speaking a personal hygiene activity, it is, nevertheless, an aspect that must be taken into account in the design of any fixture. This is particularly true of the water closet since it and its immediate surroundings have long presented an aggravated and especially disagreeable cleaning problem, both in and out of the home.[6] Not only is the water closet more readily

46 / CURRENT APPROACH TO A BIDET/WATER CLOSET

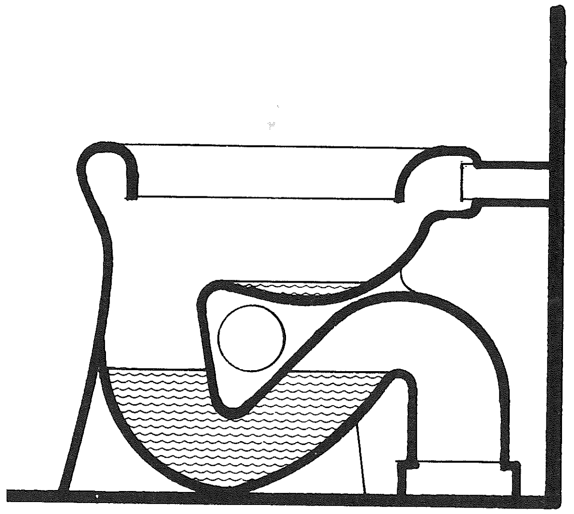
"ATLANTIC" BIDET/WATER CLOSET: CLOSOMAT, ZURICH, SWITZERLAND



soiled than most other fixtures, but its overall design and configuration also commonly make cleaning needlessly difficult. In addition, the particular nature of that soiling is such as to provoke in many people a strong psychological revulsion.

While there are a number of common soiling agents—urine, feces, vomit, spittle, blood, etc., as well as water deposits and precipitates—the single biggest problem is probably that of standing male urination. As we shall examine in more detail in the next section, such urination inevitably results in urine's being deposited on the rim of the closet, the seat, floor, and walls. If not cleaned frequently, such deposits produce an odor, attract dust, and can ultimately cause rotting or blistering of seats and adjacent nonceramic surfaces (see Figure 65). Insofar as this problem can be alleviated by the provision of a separate stand-up urinal or by specifically adapting the suggested high-rise water closet to accommodate such urination, an enormous improvement will have been made. The semi-squat closet suggested earlier is not intended to be employed in this fashion, for it would only aggravate the problems further. A separate urinal would be a necessity. As long as existing water closets are employed for urination, there is, unfortunately, little or nothing that can be done to improve the situation.

The soiling problem due to the adhesion of fecal matter or menstrual blood is, in most instances, a question of the design of the particular type of water closet. In most contemporary top-of-the-line syphonic-action water closets the standing pan water surface is sufficiently large to eliminate the problems of adhesion to a dry surface. With many older types of water closet, however, such fecal adhesion is a common problem. This is most pronounced in the case of the flat-pan wash-out closets still favored in many parts of Europe, where the fecal matter is deposited onto a barely wet shallow ledge, ostensibly so that the feces may be examined before flushing (see Figure 47). With such closets, fecal adhesion is virtually guaranteed, a fact that is recognized in that every such closet is inevitably accompanied by a scrub brush or "Johnny Mop." Even the fixture manufacturers



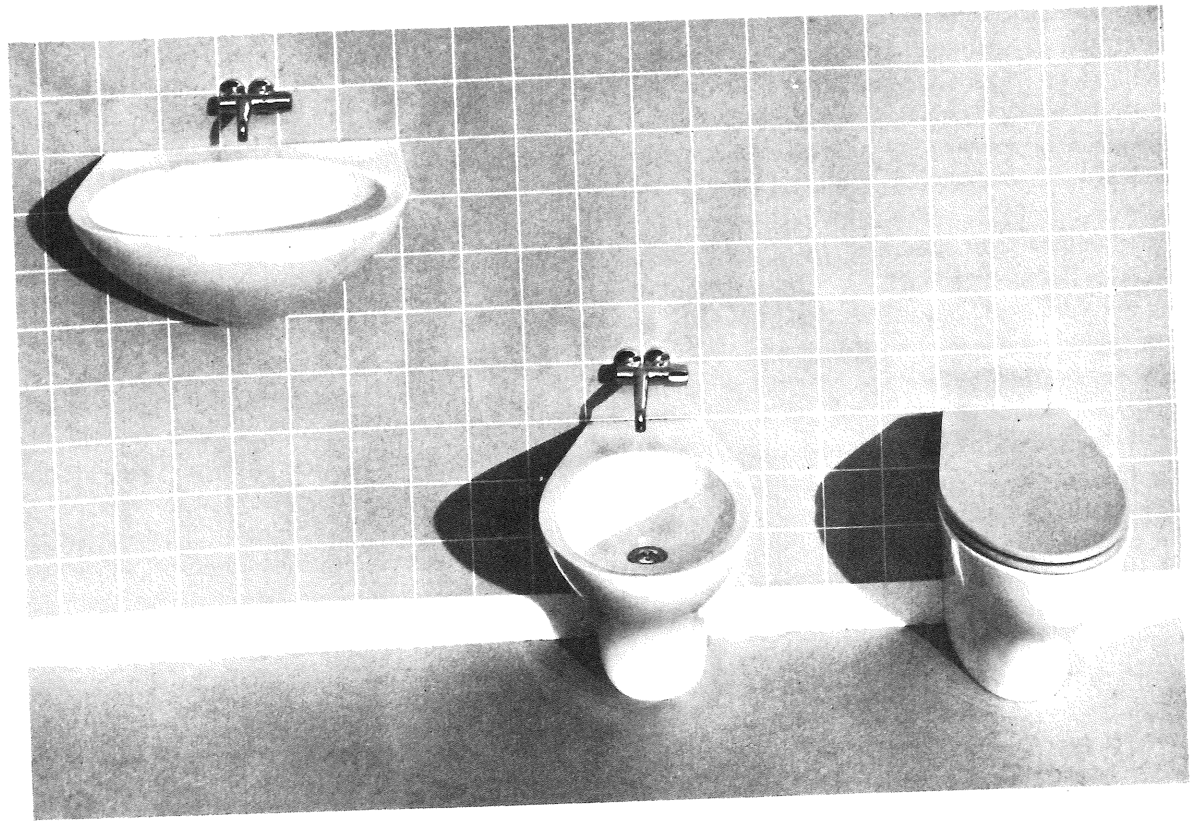
47 / CROSS-SECTION THROUGH TYPICAL EUROPEAN FLAT-PAN WASHOUT WATER CLOSET

recognize this, since many of them manufacture matching ceramic brush holders! Even if one accepts the desirability of such examination, the examination is essentially visual and can just as well be accomplished with a syphonic closet without the eternal necessity to scrub the fixture after every use. The flat-pan type of closet was originally designed for hospital use, for the collection of uncontaminated fecal samples—a use to which it is still put in many hospitals around the world. As a home fixture, however, it remains an unsightly, if not unsanitary, anachronism. Perhaps the strangest view of functional design is to be found in parts of Latin America and Spain where an unfortunate combination of coarse, heavy toilet paper and water closets with traps too small to pass the paper has resulted in the rather unpleasant institution of the waste basket next to the water closet into which the used paper is put. This is not simply a rural phenomenon but is also common in the cities and can be found in hotels and public toilets.

The most common, and inescapable, soiling arises from the water's leaving behind various

48 / CURRENT APPROACH TO MODULAR FIXTURE DESIGN

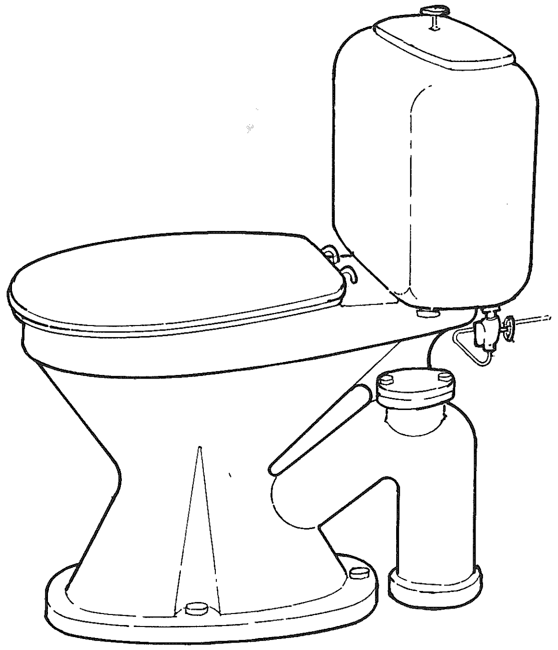
"MERIDIAN ONE" FIXTURES:
ADAMSEZ, LTD., NEWCASTLE-UPON-TYNE, ENGLAND



organic and inorganic deposits and precipitates. Depending upon the relative hardness of the water and whether or not it has been treated in a municipal treatment plant, the precise composition of these deposits will vary rather substantially, hard water resulting in the greatest proportion of inorganic scale build-up. In addition, many of the compounds in the water react chemically with the urea in the urine and produce ideal conditions for bacterial growth. This is particularly true in areas such as the underside of the rim, which retains splashed urine, is alternately wet and dry, and is rarely cleaned. Allowing urine to stand in the bowl in order to conserve water, as is sometimes done, accelerates bacterial growth, which in turn provides more base for the adhesion of inorganic solids. There appears to be little that can be done in these respects except to clean the fixture regularly, preferably with a cleaning product formulated for the purpose.

From a design standpoint, however, there is a good deal that can be done to improve the cleanability of fixtures as far as configuration and accessibility are concerned. Placement of the fixture can also be important since often a fixture that appeared easily cleanable in isolation proves virtually impossible to clean after installation because of inaccessible spaces in and around the fixture.

Perhaps the most important criteria are the simplicity of the shape and the simplicity of the joints the fixture makes with the wall and the floor. In this respect, the wall-hung fixture possesses the one considerable advantage of eliminating the most bothersome juncture. Its use is, however, limited by the additional costs involved in providing the special mounting arrangements necessary. One of the simplest, most direct, and effective approaches to this problem is illustrated in Figure 48, where the base and back of every fixture is straight-sided and modular with respect to common tile dimensions; this permits the fixtures to become integral with the wall and floor surfaces instead of being applied to the surface as is commonly done. A further refinement of this approach would be to provide a cove at these junctures so

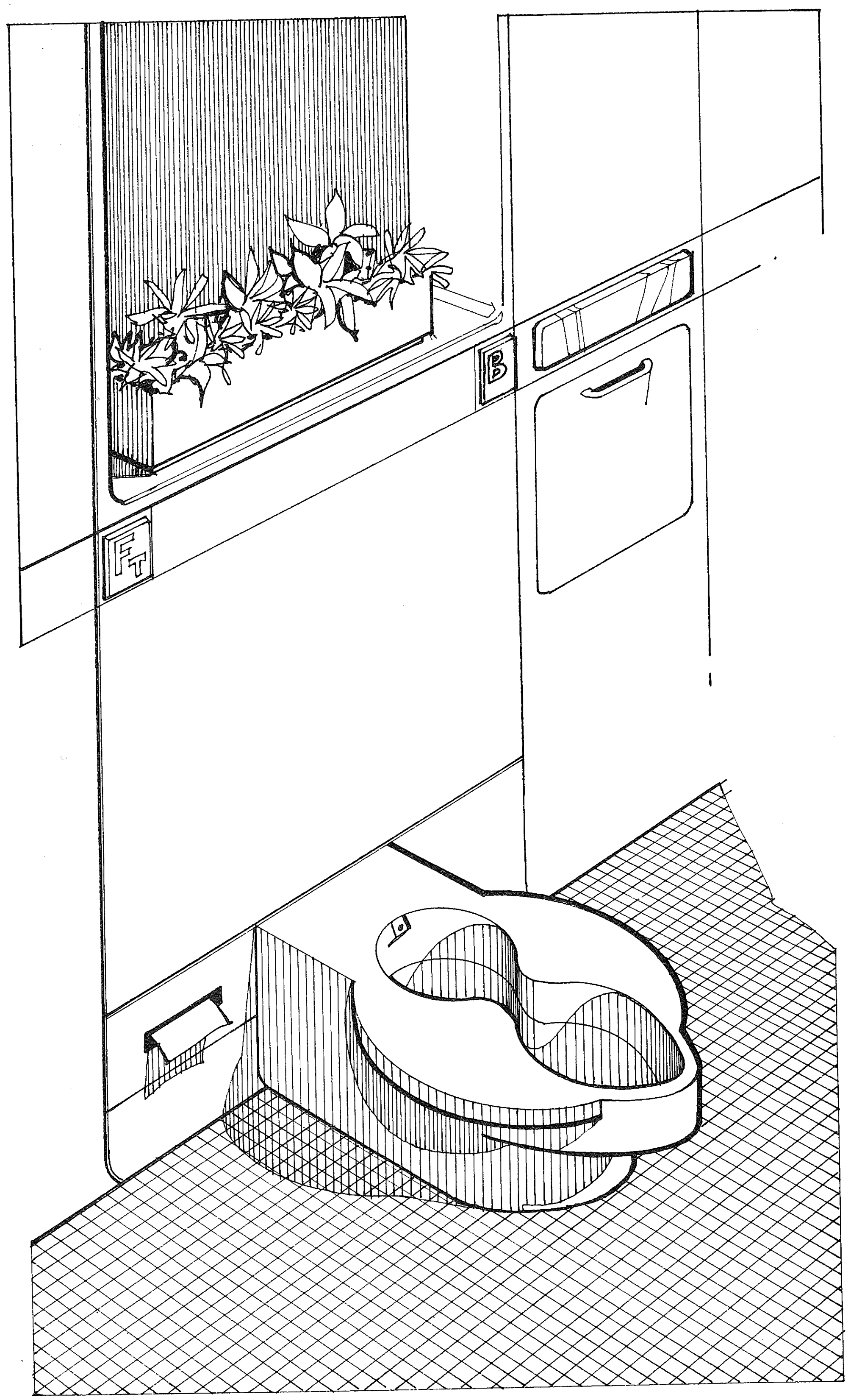


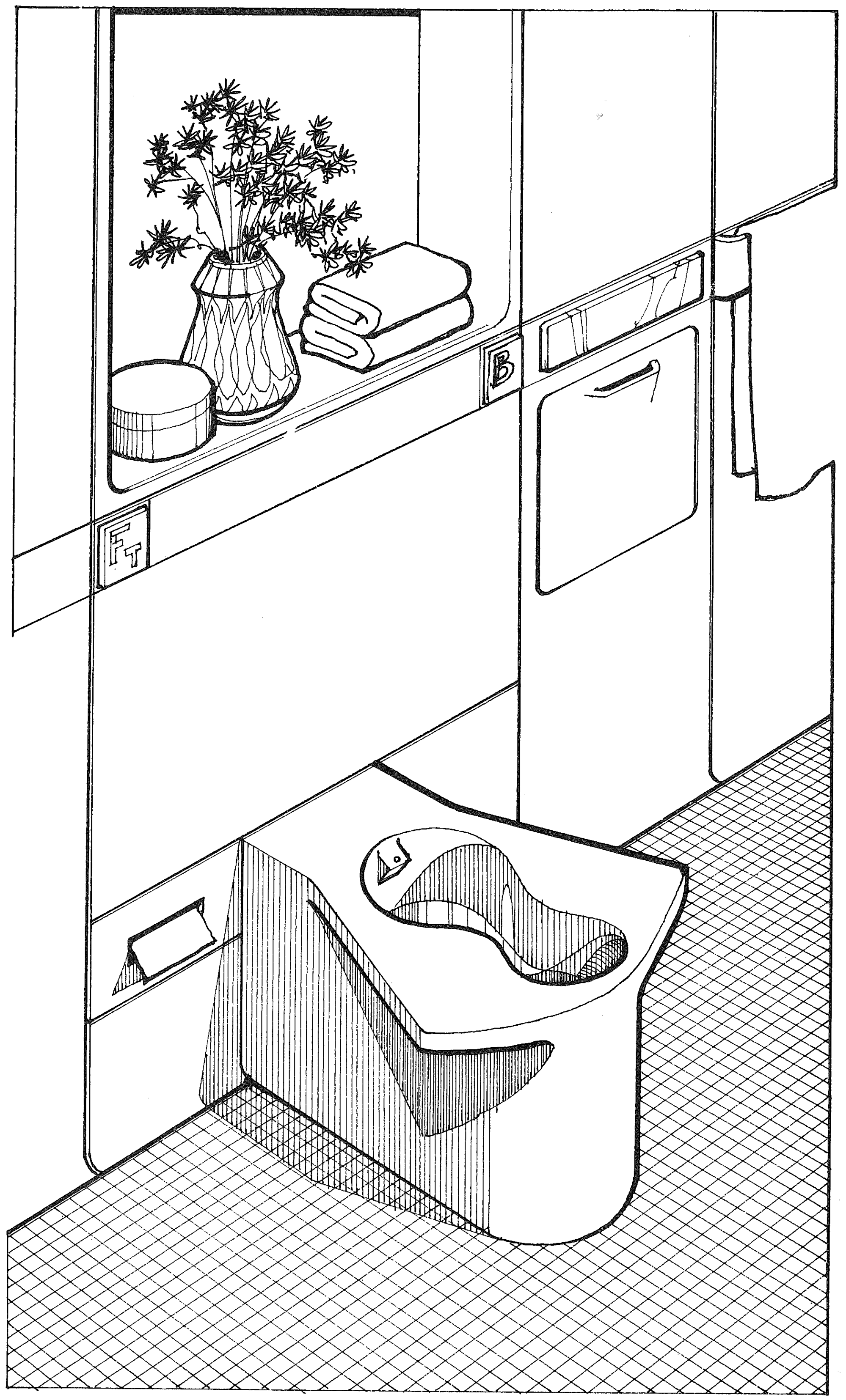
49 / TYPICAL DIFFICULT-TO-CLEAN EUROPEAN MODEL OF WATER CLOSET

that the finished joint is rounded and the surfaces flow smoothly together.

Probably the worst configurations as far as cleaning is concerned are those found in the inexpensive European water closets that resemble Rube Goldberg contraptions where every turn of the trap is left exposed (see Figure 49). The complexity of these shapes renders them virtually impossible to clean properly. Such fixtures are actually no more complex than others; the problem is that the enclosing shroud normally used has been omitted, presumably for cost reasons, leaving the functioning parts exposed. While this may raise, in the minds of some, the aesthetic arguments about form following function, it can

50 / POSSIBLE APPROACH TO A SEMI-SQUAT WATER CLOSET





be argued that cleanability is also a function and, in the minds of many housewives, is equally important.

Similar difficulties are also commonly encountered with toilet seats and lids, particularly with the hinging and mounting mechanisms, which are often complex and provided with clearances so small that many parts are essentially unreachable. While there are many approaches to simplifying these mechanisms, the most foolproof is, of course, to eliminate the seat entirely.

51 / POSSIBLE APPROACH TO A HIGH-RISE WATER CLOSET

Summary — Design Considerations From a functional viewpoint, any fixture designed to accommodate defecation should observe the following criteria: encourage and permit the comfortable assumption of a squat-like posture; provide appropriate and adequate support; have an opening adequate to permit hand access to the anogenital region; incorporate provisions for anogenital cleansing; have functional, easily accessible controls from both a seated and standing position; minimize soiling problems and permit ready cleaning; and be conceived of as a part of a comprehensive modular system of personal hygiene facilities.

(Several possible approaches described earlier are illustrated in Figures 50 and 51.)

CHAPTER 3

Below the Subway: Taking Care Day In and Day Out

with Noah McClain

The New York subways are obvious sites of security concern; many measures get taken as a result. Such concern is not folly and nor is the disposition to try and address it. Attacks on the trains in London, Tokyo, Moscow, and Madrid have unleashed, each in their own time, death and destruction. New York has 468 subway stations, each with multiple entries. Depending on time of day, some crowd together hundreds, if not thousands of people in compact spaces. It surely dawns on most who are ever there that these are rich targets.

One way into understanding how security actually works in the subway is to talk with the people who are there a whole lot—the subway workers. They are employees of the Metropolitan Transit Authority (MTA) and most spend many years at the job. Noah McClain and I studied just what workers see and what they do as part of the routines of their jobs. We aimed our interviews to learn, in particular, how they responded to a range of troubles, from minor ones to those at the level of disaster. There is a connection, we came to think, between what goes on during routine work and what happens—or will likely happen—when genuine emergency strikes. Disaster studies, oddly enough, seldom attend to the way in which people work through more ordinary troubles. But doing so, it seemed to McClain and myself, was a good way to understand what individuals might do should a bigger threat develop. Given what we learned in our research, we concluded that workers' encounters with ordinary troubles shape response to difficulties of whatever sort and scale.

Our interviews, conducted mostly by McClain although occasionally with me trailing along, involved a total of eighty workers—station agents (the people inside the glass-enclosed booths), station cleaners, train conductors, and train operators (formerly known as "motormen"). We also spent much time in the subways observing, taking photographs, and doing other kinds of field research. In some cases, those we interviewed, usually in their homes or some neutral place like a coffee shop, invited us to their work location, where they could point out to us exactly how they operated their equipment, called for aid from police or fire, or managed to innovate on their own. From these various maneuvers we were able to see how their "eyes and ears" —what they see and hear—link up with concrete actions. To learn from those with a more overall responsibility for the system, we had extensive and numerous interview conversations with union—Transport Workers Union (Local 100)—and MTA officials, including present and past heads of security.[1]

Reacting to the terror attacks, officials in charge of the system experienced the obvious need to "do something." Like the security planners in Boston observed by Kerry Fosher (see chapter 1), they responded to Department of Homeland (DHS) color alerts—a system that several of our informants indicated began within the MTA itself, or at least within New York jurisdictions. As alert levels rose, more plainclothes police were put on platforms and in train cars. There was also a beefing up of uniformed Transit Police placed on duty. Whatever the source of heightened concern, armed troops are sometimes put on patrol in heavily used spaces, like Penn Station and stations serving Times Square. Police also conduct random inspections; they have the right to check bags and belongings of passengers, a program begun in 2005 following the London Underground bombings. Out of concern for civil liberties and to lessen likelihood of profiling, anyone singled out for inspection has the right to decline by exiting the train station. But those who resist may, of course, face disconcerting follow-up once they get above ground. A Brooklyn hospital manager of Kashmiri descent, whose case was taken up by American Civil Liberties Union, complained that he had been stopped twenty-one times in the subways—most New Yorkers have never been stopped at all.[2] So another cost of doing something is indeed profiling, and the costs and resentments to which that leads.

A not-incidental additional cost occurs in the ridicule these measures stir up among workers who take note of the more absurd aspects of the programs. One is the concentration of troops and equipment in the densest stations, like the one at Times Square, which no doubt rank high in official risk assessment analyses. But the heavy security squads at central Manhattan stations make sense, some workers reason, only under the presumption that bomb makers would drive with their bombs into the densest part of the city and park, and then descend to the subway station, where their suspicious behavior would be detected and lead to their arrest. These measures don't just provide mirth among workers, they also shape how workers interpret other policies announced as security related.

Part of the MTA "do something" involves deployment of new technologies that would replace or at least supplement human guards and inspectors. The federal government, augmented by state and local financial support, makes security dollars available for advanced technologies and equipment (but not, for example, routine training and maintenance). At a city council hearing I attended with McClain in early 2005, council members took turns lambasting MTA officials for failure to spend the money "to make our city safe." Almost $600 million in federal and state security funds had been left untouched for years after 9/11. Partly in response to these confrontations, the agency made its first funds commitment to Lockheed Martin in August 2005 for $212 million, a contractual commitment that eventually hit $453 million—with still more millions to follow in future contracts—for the subway's surveillance system.

These were not to be ordinary surveillance cameras. They would detect a human separating from a suitcase or package, something that *then* would trip off an alarm that *then* could trigger emergency response. At minimum, the system would generate "intelligent video" that would integrate moving images from across the system to certain centralized locations. It was clear to us, based on our interviews with MTA security officials, that the agency had not been "foot-dragging" in avoiding commitment, but instead were discouraged by genuine concerns that, even within its own terms, the pattern-recognition aspect would not work (there had been no demonstration of the promised technical capacity). Indeed, New York subway stations are particularly inhospitable to even ordinary surveillance photography, given the unevenness of lighting, plethora of nooks and crannies, and vulnerability to vandalism. But authorities were under great pressure to spend, and the new cameras provided a means to do it. Some MTA officials were particularly skeptical, both of contractors' motives as well as capacities. As the head of transit security told us at about this time (in June 2005), "Everybody wants to sell you what they say is the best security technology in the world. And then you find out they're trying to sell you closed-circuit TV cameras. We don't want a *picture* of the event."

The contracts went forward with the sad outcome of, in the words of a state comptroller's office press release, the program falling into "disarray."[3] Lockheed sued to withdraw from the project after the city complained of performance failures. The MTA countersued Lockheed arguing "that Lockheed had provided faulty technology that did not pass basic operational tests."[4] As a transit official testified to the City Council about the legal and financial morass, "The technology does *not* work in our subway system." She continued, "We piloted the technology in a subway tunnel–like environment. It's dark, there are too many columns, there are too many people and there were too many false alarms."[5] It had been, it seems at least in retrospect, a sci-fi pipe dream, a Reaganesque Star Wars for the subway.

But what if the surveillance equipment had worked? Worked at what? We still would have faced the same problem: we don't know what kind of terrorist is going to strike what kind of target and use what kind of method. A suicide bomber, as we know, never separates from his or her bomb. We still would have been stuck with a stupid "working" thing tripping off all sorts of false alarms as folks forgot their backpacks or teenagers tossed belongings to one another. Or a New York fly got on a lens. Without the right kind of attack under just the right anticipated conditions, the system would have done little but screw up the routines of workers' and passengers' lives.

Other subway initiatives are designed to further the vigilance of passengers themselves as surveillance instruments. Most famously, people are instructed, "If You See Something Say Something." Under contract with the MTA, the ad agency Korey Kay & Partners designed signs and posters, starting in 2003, that the agency placed throughout the transit system—in subway cars, on buses, in stations, and on the agency website. It comes in different languages (¡*Si Ves Algo, di Algo*!) and its use has spread to authorities around the world.[6] There are also TV spots. “Since the phrase was introduced by the MTA in 2003,” the MTA website indicates, “dozens of municipalities in this country and around the world have asked permission to use it in their own anti-terrorism campaigns.” The slogan has been officially licensed by the MTA to the DHS for a national campaign. The loudspeaker systems within subway cars and at station platforms recite the same admonition. Given the poor quality of the public address apparatus, the messages add still another garbled set of noises to the din of screeching trains, yelling kids, and buskers asking for money, selling candy, and performing musical numbers. Striking a higher-tech note, New Jersey Transit invited passengers to use their cell phones to text authorities (“text against terror”) as the media reported it.[7]

Another front of the MTA drive to enlist public support was a campaign to assure people that their calls did indeed matter, and that they are not alone in making them. So, as part of a three-million-dollar program, posters on buses and subways, as well as ads on TV, proclaimed, “Last year, 1,944 New Yorkers saw something and said something.” The ads include a hotline number to call: 1-888-NYC-SAFE. According to a *New York Times* article by William Neuman on the subject, neither police nor transit officials could say where the number 1,944 came from.[8] Indeed, in our own interviews with transit officials (conducted both before and after the *Times* story), including the heads of public information and security, McClain and I similarly could come up with no method or sources that could confirm this number. Our informants did not claim confidentiality, they just did not know. Putting up the number, the precise number, enables the MTA to show that it is indeed competently dealing with the problem, “doing something” in fact, and that many New Yorkers, through their active participation, believe it as well.

In all, according to the *New York Times*, the hotline received 8,999 calls in 2006, including calls that were transferred from 911 and the city’s official nonemergency help line (311). The next year, the total shot up to 13,473, explained perhaps, according to the police department spokesman, by the ad campaign touting the readiness of New

Yorkers to participate. No calls, however, involved terror or terrorist threats against the subway or any other target, although there were false reports and phony allegations. Some people used the hotline to turn in people against whom they evidently held a grudge. A total of eighteen arrests over the two-year period (2006–2007) could be traced to the calls, but for offenses like making up a phony ID or possessing an unregistered handgun. Several calls led to deportations. Eleven reported on individuals said to look like Muslims, who were taking photos of train tracks or "counting people." Muslims use hand-held counters to tally their prayers as they make them, the way Catholics use rosary beads to keep track of the number of Hail Marys they say.

A total of 816 of the calls in 2006 led to follow-up investigation by the police intelligence division or its joint terrorism task force with the FBI. Of the calls about suspicious packages, most were about backpacks, briefcases, or other items accidentally left behind by their owners (who were grateful, one suspects, for their return through security). None had bombs. Neither newspaper reports (we have searched all data bases) nor our workers' accounts provide any cases where "Saying Something" revealed a bona fide attempt to inflict terror in the subways.

In response to the *Times* reporter's query, the police spokesman explained, in reference to the "See Something" campaign, "It's just one small part of the initiative the Police Department has to capture any information that might prevent another 9/11 or another catastrophic attack on the city. One call one day may be the one that stops an attempt to destroy the Brooklyn Bridge." The Director of MTA Security commented that officials in Madrid said that several passengers interviewed after the bombings remembered seeing the unattended knapsacks that turned out to contain the bombs, but for whatever reason decided not to alert anyone. The head of the MTA (Katherine Lapp at the time) indicated that this is something that needed to be avoided at all costs, even if it meant dealing with a deluge of false alarms.[9]

There have indeed been plots against New York since 9/11, most recently a failed effort to detonate a car bomb at Times Square in May 2010. The device did not explode; smoke detected by those on the streets alerted authorities, who closed down Times Square and gathered clues from the intact vehicle. It was an amateurish Rube Goldberg assemblage of home-made ingredients, according to police officials—by a miscreant who had "more desire than ability."[10] Faisal Shahzad, a U.S. citizen born in Pakistan, was readily caught and pleaded guilty.

The most serious plot against the subways, per se, was a 2009 scheme by an Afghan American, Najibullah Zazi, to set off simultaneous bombs (with two accomplices) at the Grand Central and Pennsylvania Railroad subway stations. The plan, or at least an early version of it, was evidently first detected by British intelligence, which intercepted communications between Zazi and his purported al-Qaeda handlers in Pakistan.[11] U.S. authorities trailed his activities as he bought bomb-making materials in Colorado, where he was working as an airport shuttle bus driver (the airport–downtown Denver route), and then as he made his way to New York to commit his crime. Before he could act, but with plenty of evidence accumulated against him, he pleaded guilty. His accomplices were also charged and face severe sentences. As per pattern, detection came not from the "See Something" campaign, from on-site inspections, or from any of the security operations set up by the subway system. It was external intelligence that found the threat.

The other big attempt to commit subway mayhem, a plot ended by arrests in 2004, was frustrated through a different scenario. It was a conspiracy to set off an explosion at the 34th Street–Herald Square station (the busy station that serves Macy's—the "world's largest store"). The plotters never reached the point of purchasing any materials for making a bomb, or going beyond amateurish drawings and speculations about what it would take to do destruction. The charge of conspiracy was also complicated by one of the plotters repeatedly saying, as captured by police recordings, that it was important to him that people not be hurt.

The NYPD's paid informant in the case had a central role, as made evident by the tapes played at the trial of one of the plotters. The informant was double the age of the two young men charged with the crime, James Elshafay (nineteen at the time) and the only somewhat older Shahawar Siraj. The young men regarded the police informant, Osama Eldawoody, as a father figure, and he called Elshafay "son."[12] The younger two seemed anxiously deferential toward his wishes, expertise, and what he convinced them were his impressive connections with Islamic scholars in the Middle East. Eldawoody had in fact been trained as a nuclear engineer, although he made his work in the United States by driving a taxi and performing other miscellaneous occupations. The three of them were all in frequent and lengthy conversations over many months. The interactions of the informant with Siraj are recapitulated in an impressive book on "snitching" by the journalist Ethan Brown, who presents the informant as clearly leading the young men toward schemes they do not seem capable of hatching (alas, Brown does not consistently provide sources for his paraphrases of the conversations).[13]

A defense lawyer for Elshafay argued in court there was little evidence that his client—young, uneducated and clinically schizophrenic—"had any ability whatsoever to carry out any kind of plan."[14] The informant received about $100,000 from the NYPD over a little more than a two-year period, with his weekly stipend rising as his reports became more forceful (he was to complain later of underpayment). The case resembles another highly sensational plot (not involving the subway, but also out of New York) where it turned out that some of the defendant's incriminating statements were made while both he and his NYPD informant were smoking pot. The alleged plotter, Jose Pimentel, later emerged as "unstable," having twice attempted to circumcise himself, among other unusual acts. And, again, we have an informant whose aid was crucial given the incapacity, evident from still other details of his biography, of the man charged.[15] In this instance, the issues of entrapment were sufficiently strong for the FBI to refuse to be part of the prosecution effort, considering the case unwinnable in court. Taking a very different stance, Mayor Michael Bloomberg, Police Chief Raymond Kelly, and Manhattan District Attorney Cyrus Vance all appeared together to eagerly announce the plot when it was first "revealed."

Finally, I turn to the example of one of the very few bombs ever to go off in the subway system and the only one to do so in the post-9/11 period. It was a pipe bomb explosion at the Times Square station just before the start of the Republican National Convention in 2004. It was reported by New York subway police officer Joseph Rodriguez, who himself was injured in the explosion. The event caused a lot of tension in the city in the immediate aftermath and led to an intensified police presence in the subway. Police brass and the media lauded Rodriguez for finding the bag and taking the hit when it went off. But the next day police changed their account and filed charges against Rodriguez for planting the bomb himself. Police found incriminating evidence on his person and in his personal computer, including instructions on bomb making. He had not called 911 when the bomb went off (he was not seriously hurt), a basic part of police protocol, and had other incriminating materials in his Manhattan apartment. Further, there was a history of psychological disturbance; Rodriguez was about to be retired from the force at age 27 on a psychological disability pension.[16] There was no hint of motive, other than what has been called a "hero complex," in which individuals yearn for acknowledgment for having saved the day.[17] Perhaps wars on terror create not only heroes, but also hero wannabes who create their own kind of disturbance.

Looking back at these various attempts at grand violence, patterns emerge. They have been amazingly few in number. None has been successful. All are amateurish, performed by people lacking the relevant skills, not always having even the healthy psychological disposition that anyone needs to be effective at any task (as in the Rodriguez pipe-bomb case). The security apparatus of the subway system did not foil their plan—although intelligence from communication interception was crucial in the Zazi case. The city was spared destruction from miscellaneous things going wrong, such as mishaps of machinery (as in the Shahzad car bomb), and from an apparent lack of seriousness or capability on the part of the plotters (as in the Elshafay "plot"). All such elements imply relevant lessons for thinking about the nature of threat and how possibly to deal with it.

Arsenal in Place

Against the background drama of bombing conspiracy, injury and death do take place in New York and in its subways. Despite the city's substantial decline in crime (murder rates are now one-third of what they were in the early 1990s), crime still occurs, and this is true of the subway as well as the streets. There is the occasional murder (only one in the subways in 2006), but quite a few robberies—819 in 2006. There are reasons to think this an undercount because of sharp and inexplicable shifts in numbers from various years and because of ambiguity in determining whether to attribute a crime as occurring at a particular subway station or at the corresponding aboveground location.[18]

Whatever the actual numbers, one feature of crime victimization is people's fear of it. Criminologists devote a great deal of research to understanding "fear of crime," partly because of the direct costs it exacts from citizens. There is also, however, a causal relationship between fear and victimization that criminologists focus on: if people fear crime they will be less likely to go out, and that, in turn, increases vulnerability for the increasingly few who do. Subway workers protect passengers from this kind of fear first by simply being there in uniforms, which likely reassures the public, especially with all the signs that advise customers to "alert" a train operator or station agent if they see something. But workers do other, more concrete, things, to make passengers feel they are safe, and—no small matter—that makes them safer in fact. Combined with actions by passengers and others on the scene, they enact day-to-day solutions that hold routine mayhem at bay. In essence, many people are already deeply involved in making things secure, albeit in ways not evident from looking at posters or newspaper reports on security. They involve garden-variety mechanisms in some nonobvious ways.

First among these mechanisms are the passengers' own tendencies to solve problems. Despite the so-called bystander effect, in which people supposedly do little to intervene to save others from assailants, individuals actually do help one another out—a whole lot (as is consistent with the bystander effect follow-up literature).[19] We encountered sixty-two different instances in our interviews, and a few more as reported in the press, in which passengers (not counting off-duty police or subway workers) intervened, usually spontaneously and not infrequently at some risk to themselves. Their actions included jumping down into the tracks to rescue people who had fallen in by accident or on purpose, as in an attempt to commit suicide. These rescues sometimes meant close calls for the passengers in the face of oncoming trains.

In other instances passengers subdued a violent wrongdoer. A twenty-six-year-old named Jonathan Cohen, after seeing a disturbed man (who later turned out to have been psychotic) push a woman to her death in front of an oncoming train, pursued the assailant up escalators and through corridors, eventually to hold him in what the police called a "bear hug" until they could arrive to take over.[20] In another case, a passenger was stabbed as he tried to break up a fight among people he did not know. Still another good samaritan waylaid a man after his attempt to molest a child, physically holding him until police eventually came. Following a passenger's apparent diabetic seizure, and after a conductor announced it to others who might not otherwise have seen that the passenger needed medical assistance, seven people (doctors and nurses) made their way to his car to help.

There are other examples where passengers stayed behind to staunch others' wounds, in one instance, alas, where a woman nevertheless bled to death. Evidence of spontaneous helping runs consistent with what researchers have documented over and over again.[21] Instead of succumbing to passivity and inaction, survivors quickly move to do what can be done. We sometimes see it in TV news footage, whether the responders be Israelis or Palestinians; Hindus, Sikhs, or Muslims; or tornado survivors in the American Midwest. Most often, however, it seems from a close inspection of media coverage that official responders are the ones depicted as the helpers. Even though this is most certainly not the prevalent pattern on the ground, media coverage does often depend on official responders arranging journalistic access. Or maybe it is the pattern because journalists favor the uniforms and rescue equipment that make for good theatrical narrative.[22]

Presumably there are times when people do not come to the rescue and when indeed some are content to let the "hero" take care of the problem; they are passive. But we could find few instances from the subway where somebody who was in a position to do something failed to initiate aid when there was imminent bodily threat to another (some of our workers do report, however, that police have walked off, leaving it to good samaritans, for example, to deal alone with fracases involving teenagers). An exception of workers coming to others' aid seems to be when violence occurs among people who are in some sense out-laws in common, as when two young men who were engaged in the illegal fare-beating system known as "selling swipes" (see later in this chapter) got into a knife fight over access to customers. Under such circumstance, people will call for the police, but do seem—sometimes appropriately—unwilling to physically step in. There was a case where private contractors working in the station did not physically stop a crazed assailant attacking an innocent man with two sabre saws (one in each hand) that he grabbed from a worksite.[23] It also seems, as more of a matter of routine, that passengers do not intervene in minor rule breaking, as when people hold open train doors or, as one woman conductor complained during our interview with her, when men exhibit themselves.

Subway workers indeed *presume* customers will help out, at least by calling police or, at times, taking a more active role. One example became clear to us through McClain's participation in MTA's fire safety and evacuation training, with real-time simulations of train and tunnel darkness. As explained by the MTA veteran who was directing the training, only with the assistance of passengers could there be effective movement of people out of cars when stopped in poorly lit or smoky conditions. People would be needed to assist at every turn to prompt customers to move right or left or to go up steps or across thresholds—for example out of trains and on to catwalks. McClain estimated that a minimum of four helpers would be needed for a simple evacuation. Without the assumption that passengers lend a hand, the whole training protocol would make no sense. Our interviews similarly revealed assumptions among workers that passengers help out when needed—like doctors and nurses who come to the aid of the injured, or those who spontaneously arise to deal with a miscreant.

While customer help is strictly voluntary, subway workers are charged in various ways to maintain order, including guarding against passenger injury. Injury can derive from direct assault or, for example, from a fire (perhaps inadvertently started by a pile of free newspapers catching a spark). Passengers can become ensnared in equipment. Or there can be flooding, which, although not necessarily a danger to life and limb, does threaten the smooth operation of system equipment and trains. In warding off any such threats, workers know and, to a substantial degree, do attend to the formal organizational rules, while at the same time enlisting ad hoc work-arounds as situationally warranted—a central theme of McClain's own research conclusions.[24]

One of a subway worker's intense worries is passenger suicide. Being part of a suicide or some other fatality becomes a trauma. Workers know about coworkers' experiences with such incidents and other types of danger through news reports but more vividly through shared word of mouth. The union, through its own safety initiatives, also provides information regarding such incidents, and how to avoid them—and how to recuperate once they occur. Subway workers find themselves dealing with what they sometimes refer to as "knuckleheads," "crazy people," or, also in their argot, "EDPs" (emotionally disturbed persons). Many worker interviewees report having themselves been victimized. In the most extreme case, a woman worker was subject to one attempted and one actual rape.

Workers' routines, including observance of official regulations, expose them to danger as a matter of course. Very specific work protocol requires conductors to put their heads at risk. They must stick their heads out of their cab windows in ways that do not allow them to see an assailant intent on hitting them on the head, perhaps with an object. Located in a middle car (the train has eight or ten cars), conductors open train doors by depressing two buttons simultaneously, each opening half of the train's doors on to the platform. But they close them in two separate operations. Conductors look first to the right (head out the window) to make sure it is safe in that direction. All doors to the right are then closed with the push of a single button—head remaining out the window and looking right. Then they repeat the same maneuvers but to the left, pressing a different button. Their heads must remain out the window during both maneuvers, sometimes repeatedly at a single stop, often for long moments and looking in only one direction at a time, focused on the car doors. As the train exits the station, their heads remain out the window until it has moved a prescribed seventy-five feet forward.[25] Taking advantage of these conditions, according to our informants, miscreants have punched conductors' heads, slashed them with blades, thrown glass bottles at them, hit them with various objects, and spit at them or doused them with other liquids.[26] But in the mean-

Figure 5. Signs and notices block views at station agent booth. Photo by Noah McClain.

time, conductors' assiduous looking (and risk taking) means that few if any customers are caught in the doors and hence dragged through the tunnels. This maneuver also ensures that children are not separated from a care giver—a high priority for conductors, as they brought up repeatedly in our interviews.

Station agents are almost always alone in their booths and, depending on location and hour of the day, also alone in the station (a minority of stations also have commercial news kiosks, which are open for at least some of the hours of train operations). Station agents answer travelers' questions, an activity that itself is sometimes helpful for learning about problems in the system (how come the right exit is blocked?) as well as providing answers that enhance passenger safety. The glass-enclosed walls behind which they sit make it possible for them to visually survey their surroundings. They have access to an intercom button to call for outside help. Their views of space surrounding the booths are often blocked by signs that the MTA puts up over the glass, some of it ironically warning passengers to be on alert for unattended packages, that their stuff might be searched, and, of course, that if they see something they should say something (figure 5).

One worry that station agents have is that miscreants could spill or spray a flammable liquid or gas through the opening at the base of their booth's window, followed by a match to set the booth afire. A number of station agents were killed or maimed in the 1980s through the mid-1990s in such attacks, accompanied by robbery attempts. Ironically, prior efforts to increase booth security contributed to the problem. To deal with station agents' vulnerability to shootings, authorities made the booths bullet-resistant. But specifics of the design had the effect of decreasing ventilation, which made explosion (rather than fire) more likely. More recent add-on remedies, particularly a fire-retardant system, have had an apparently corrective effect.[27] But since 1995, there have been seven additional attacks reported in the press, which indeed support the stories that workers have told us and which they continue to discuss with one another. In four of the seven accounts, attackers used an unidentified liquid; in the other instances, one used anti-freeze, one gasoline, and one lighter fluid. Besides dangerous items hurled through the opening, a man inserted his erect penis to shock a female agent—another account that made the rounds. If they sense trouble in time, workers can quickly shut the aperture to their window (ouch!).

Some of the station agents' safety routines are not for self-defense but for defense of others. As school lets out each day, hundreds of preadolescents or teens may descend into a station. They jump on and about one another as they approach the turnstiles, risking human crush. Some station agents have told us that at these times, they disengage the turnstiles allowing anyone to go through with a simple push. The agents fear that someone attempting to go through the turnstile will discover they have no card, that their card will malfunction, or that they will just fumble and take too long. Rather than risk a pileup, station agents disengage the system—something not permitted for workers to do under MTA rules, but done anyway.

Similarly, station agents are not to leave their enclosures to service the needs of customers, but if a child is in dire trouble, for example, they simply do it. If there is a difficult person at a station, the station agent can give him or her a free admit through the turnstile to enter a train, albeit then perhaps to become a problem at a different station. Workers on the trains also have options, not in the rulebooks, to deal with safety concerns. One strategy is to lock wrongdoers out of trains or, depending on circumstance, lock them in. Conductors have the capacity to key lock car doors individually, both the doors that lead off to platforms as well as doors between the cars themselves. A worker reported to us that on one of his runs he realized kids in a car were shattering bottles against windows, ripping seats from the train floor, and trying to throw them through the glass. Among other hazards, their actions could have resulted in objects on tracks with potential for derailment or sparking a fire. He also saw they had knives. The conductor—who happened still to be in his apprenticeship—took the impromptu strategy of turning the last car, where the disruptive kids happened to be, into a holding cell. He closed all train doors. He then made his way back to the penultimate car and got the remaining innocent people out of the car still containing the troublemakers. As he explained in our interview,

> So I went to the ninth car. I motioned to the people [who could see me through the glass window], "come this way come to me," but I didn't want to give it away because they [the miscreants] would have probably attacked me. Locked what they call a storm door, which is at the end of the car.

This trapped the offenders where they could be held while police were called.

When the police did come, they were on the wrong side of the platform. Some crossed over the tracks (contrary to policy because of the danger of electrocution and exposure to moving trains), while others took the stairs up, crossed over on the street above, and then came down again (several flights each way). By securely penning the offenders, the conductor saved the day regardless of the wait-time for the police. While the capacity to key lock individual cars stems from the need to protect them while in storage or in yards while awaiting repair, the feature comes in handy when workers need to make an ad hoc security application.

Figure 6. "Shoe slipper" used by conductors to disconnect a train from its power source. Photo by Noah McClain.

Similarly the public address system on a train can be oriented toward safety, albeit in an unauthorized way. A cleaner working inside a train reported,

> I'll go to the intercom system where the conductor gives his announcements, and I'll go, "All right, people, be aware. There's a pickpocket on the train." And I'll make an announcement over the, uh, you know, the PA system.
>
> Now if my supervisor caught wind of me doin' that I'd be in trouble . . . [asked why by the interviewer] because I'm not supposed to be in the conductor's cab makin' announcements cause I'm not a conductor.

Another instrument at hand that train workers can put to unauthorized safety use is the hardwood "shoe slipper" or "shoe paddle" as it is also sometimes called, which is yellow in color (figure 6). According to official procedures, train workers may use it to disconnect the train from its power source, by inserting the shoe slipper between the shoe and the third rail. They may need to do so when, for example, someone is under a train and could be electrocuted by train components in live contact with the third rail.

Here is a conductor interviewee showing us the shoe slipper's common use as a security instrument:

> We're not supposed to carry that for any other reason but to do that function [disengage the train]. But, when I go out there in the middle of the night, I don't go out there without something [to protect myself].

One conductor describes how he was able to divert a ranting drunk away from passengers. He employed the shoe device to prop open the door to his compartment (which otherwise closes on its own) to divert the angry man's attention away from passengers and toward himself:

> It was a homeless person. He was getting up in peoples' faces while they were sitting down. He was screaming at them, cursing at them. Daring them to get up and fight him. He was really really out there. And I was telling him, "Sit down and shut up already. Stop bothering these people." And he'd come and scream at me, and I said, "Sit down. You want to sit here?" I propped my door open with a shoe paddle.[28] I said, "Go ahead. You can scream at me all you want 'til the cops get here and take you off!"

While rules do not permit MTA employees even to touch a passenger, workers use the shoe slipper to wake up sleeping passengers sprawled across the seat or to nudge them along from one location to another. Other instruments as well have served as defensive weapons, according to reports—for example, a detachable train break handle. In a very uncommon report, a worker used his brake handle to hit a violent person.

Train operators also have horns they sound to herald their coming to anyone who may be in the way of the train around the next bend, including impatient passengers on platforms leaning forward to see if a train is coming. The horns are very loud, and hence operators sound them only as short bursts and only prior to entering passenger platform areas. They can go to full volume, which is literally deafening,[29] to warn of danger, as in cases where workers or others are on the tracks.

In many instances, such as an assault or robbery on board, train operators are instructed to "blow for police en route." Blowing the horn in a particular pattern (long-short-long-short) sends a signal to police who might be within earshot, even several stations away. But horn blasting can also deal directly with perpetrators. A train operator once used his horn when he noticed a mugging attempt as he came into a station. He blew his horn loud and long as he came on the scene and kept it blaring. As he explains,

> So the woman was getting mugged on the platform. Avenue H I think it was. So I blew the horn. Try to get him to leave her alone. I was slowing down in case these guys think about pushing her in front of the train. . . . And I started chasing them with the train but I couldn't go any further cause I came to the end of the platform.

The operator had hoped the sound of his horn, indicating an approaching train with an employee on board who could see what was going on, would dissuade the mugger. Or that the continuous blast, by sheer aural force, would disrupt the activity. As it happens, it did neither; the perpetrator got the woman's purse and made a successful escape. But the operator's reaction shows how the horn, and indeed the moving train itself, can be operated in an extemporaneous way to address a problem.

Besides locking in perpetrators, workers can use the individual car key feature to protect passengers from a repulsive and bloody mess, and hence the fears and anxieties that such visible residue of violence might induce. During one of McClain's sessions observing train crews, a teenage male was assaulted in one of the rear cars of a train. After police and ambulance responded to the call for help, the car had blood on the seats and floor, and rush hour was fast approaching. Lacking equipment or personnel to safely clean the train car, the MTA worker "keyed off" all the doors in that car—ensuring its doors would not open at station stops. The train was thus able to remain in service without further alarming passengers.

Admonitions to report all problems to supervisors pervade the *Rules and Regulations* book (MTA, 2003) and all deviations from strict procedures require permission from supervisors. Workers do strictly adhere to some rules, like the instruction for conductors never to close train doors without simultaneously having their head out the window, as explained earlier. But in many other cases, the rules (as with the equipment) are treated in a more ad hoc manner.

Unlike many other kinds of workers, such as members of a construction crew or clerks in a government office, subway workers do not typically have the resource of calling on adjacent workers for tips and real-time help.[30] When they do communicate with supervisors—known colloquially as "Control"—it is via radio or phone. Such communication is an example of what has been called "interaction in isolation."[31] Subway workers are not allowed to call police and firefighters directly. Although not all such contacts between a station agent and Control is for a major emergency, the contacts are voluminous—on the order of over 150,000 calls each year.[32]

Summoning help always carries some risk of being misunderstood by the call-center operators or generating delays in action being taken. Examinations by social scientists (conversation analysts, as they are known in this case) of "calls for help" to 911 in the United States, clarify how, quite beyond the subway context, systematic difficulties can arise from phoned-in requests for aid.[33] In the 911 case, emergency dispatchers require specific information to fill in data fields before sending assistance, sometimes insisting that their questions be answered in a required order. This can lead to misunderstandings, delays, and disastrous consequences, including—in one instance yielded in the research by Don Zimmerman and his colleagues—death.[34] (A woman died as her frantic adult son was unable to follow the proper information sequence demanded at the other end of the line.) In responding to forest fire reports in Southern California, dispatchers cause response delays with questions aimed at determining who or what started the fire.[35] Emergency aid dispatchers, as in fire or crime reports, have two goals: gaining clues to who may be an arsonist or criminal and also attending to the need for aid. It can be a tricky balancing act trying to decide how much time and effort should go toward one goal versus the other, given the urgent need for fast action.

In the subway, workers must satisfy the information needs of Control (for example, "Are the people fighting on your train male or female?").

They may be prevented from following what for them, at the moment, might seem a more common-sense narrative that would mobilize police or fire immediately. By the time help is on the way (and sometimes promised help just never arrives, according to our informants), the damage may be done, a culprit gone from the scene, or the problem already solved by workers and/or passengers. Control may not agree with workers' evaluation in any event, or the type of response may not correspond to what the worker thought necessary (for example, fire fighters arrive when it was police who were needed). For all sorts of reasons, then, workers may "take care of it" on the spot, as best they can. The arrival of help can also disrupt train service, and the request for this help, if not considered the proper response to a bona fide problem, can reflect poorly on the employee.

Especially when situations are highly ambiguous (they are always ambiguous to some degree), it comes in handy for those at the top of organizational hierarchies to decentralize responsibility for emergency aid "as far down the organizational line as possible."[36] If things go wrong, those below can take the blame. As Robert Jackall explains, "When blame is allocated, it is those who are . . . politically vulnerable or expendable, who become 'patsies', who get 'set up' or 'hung up to dry' and become blamable."[37] As much as possible, blame is set up to go downward.[38]

Perhaps the most instructive instance of a subway worker under high stakes and organizational uncertainty comes from the World Trade Center site itself on the date of 9/11. One of those McClain interviewed was the operator of a train headed right into the World Trade Center–Cortlandt Street subway station as the building was about to collapse (the only such train reported to be doing so at that time and place). The train driver defied his orders to *not* stop. As he said during the interview,

> They said continue in service, but bypass Cortlandt Street–World Trade Center. So I made the station stop at Whitehall and right around then is when they said, "All trains, northbound southbound, bypass Cortlandt Street." What happens is Command Center has to go by what they're hearing. They're in the office somewhere in Brooklyn: Jay Street. They

> can't see what's actually happening, so they just run their playbook. They were running it as a "smoke condition." And as a rule of smoke conditions we bypass the station. You don't want to put customers onto a platform with smoke. You don't want to take smoke into the train.

But the train driver saw a waiting crowd, about four persons deep. The station was indeed in a smoke situation. As he entered the station, the train driver locked eyes with a woman anxiously waiting to board. He recalled,

> And I had never seen that kind of fear in my life. I couldn't see her facial expression. All I could see were her eyes, and it was fear.

This cue of "her eyes" apparently convinced him to stop the train and take on the new passengers. As he told McClain,

> Cause, um, I was disobeying orders. My order was to bypass the station. I made a judgment call. I made a judgment call based on experience, basically.

The train driver was intensely concerned about the immediate circumstance, but also how he was going to report the fact that he stopped the train against standing procedures as well as the specific instruction on this occasion:

> Well, in my mind I'm thinking, "I know I'm going to have to hear about this one." So I was trying to figure out how I was going to write the . . . we call it a G2, and it's in the report. I was trying to [figure] out how to word it. Where, um, I minimize myself and my conductor getting into trouble.

As it turned out, the train operator made the right call; he surely saved many lives.

But he was right to have concern. Workers are wary of calling attention to themselves in general. It may force them to stay beyond a shift to make written reports or meet with agency officials. This may disrupt personal schedules, like complicating child-care or causing a worker to miss a ride home or commuter train. In these after-shift sessions, supervisors might scrutinize workers' accounts (or workers fear this will happen), as they look for rule violations, including actions not even related to the reported incident. Perhaps the workers were standing in the wrong place at the time of the occurrence or they were wearing an unauthorized type of clothing or insignia.

By far the most prevalent reason for workers to resist rules, especially security dicta, is the risk of stopping the whole system. Not everything can be reported and corners must be massively cut in this regard. Given the range of things that go on in the subways, there is huge potential for false positives. Someone may tell an MTA employee that a suspicious bag has been left on a station platform. A worker tells us she "just kicks" bags to see what might be in them. Calling it in to Control risks a station closure, just like a conductor calling in a box left on a seat means that the train goes out of service. Major delays then occur and a worker who does not use common sense is, in Harold Garfinkel's apt term, a "judgmental dope."[39] That could mean discipline, including being "targeted," as the workers say, for some other potential infraction.

Some instructions aimed against terrorist acts are seen as not capable of being of any use in an actual emergency. For example, the MTA issues masks, known as "escape hoods" to protect against chemical or biological attack to conductors, train operators, and track workers. The masks hang from belts at hip level in a plastic pouch. Station agents have a single mask in their booths. The MTA issues no masks at all to cleaners (who are on the lowest-paid rung of the job scale). Some workers think the masks add security to their work yet note they are meant for only fifteen minutes' protection, maybe enough time to get themselves out of tunnels but not to aid others in doing so. Evacuations, they say, require great patience and the power of speech (which is muffled by the mask). The actual evacuation that McClain encountered in fieldwork in September 2004 (as opposed to the simulated one in the training reported earlier) occurred at midday on a lightly populated subway. A conductor and train operator collaborated to move all passengers into the rear car, which was very near a station platform. Passengers stepped from the train directly onto a two-foot-wide catwalk and proceeded fifteen feet to the platform. This calm and uncomplicated evacuation (as no one had to walk through tunnels or climb emergency ladders) took about thirty minutes. In the experience of other evacuations, between 2003 and 2006, it took an average of one hundred minutes from incident to "open sky."[40]

The transit union has publicly complained that the emergency training offered to workers is merely perfunctory—the MTA distributes pamphlets to workers and offers a brief course that is especially inadequate in terms of showing how to assist the public. Workers are basically told, as one of our interviewees complained, to "cut and run"—something we know they don't in fact do, judging from the examples already described. But given the paucity of equipment and other practical exigencies, cut and run may at times be the only realistic course. And beyond a mask shortage, or the proliferation of masks that hinder communication, there are deep problems of infrastructure, like inadequate lighting and signage, indecipherable audio messages, and emergency platforms that are only marginally accessible—they are often too narrow for a significant group of people to make their way out.

The details of turnstiles, otherwise so routine, become critical during evacuations. Early New York subway turnstiles provided relatively little obstruction for those needing to exit in a hurry, as evidenced in a 1943 photo (figure 7). But that meant they also did little to obstruct people from entering without paying their fare. Whether young or old, frail or obese, people could climb over or crawl under and get a free ride. Out of concern to "protect the fare," the modern turnstile emerged with sloped entry panels, bars placed just above head level, and angled flanges on the turnstile itself (figure 8). Although claimed by its creators to have been "designed with reference to the human body," it was based, as made clear in the patent application for the equipment, on the bodies of U.S. military servicemen. Even within that constricted population, only the bodies of the middle 90 percent of the sample were taken into account. The design method traveled from the military to the subway turnstile without adjustment for sizes and physical capacity of the civilian population. Some people have crutches, guide dogs, babies, strollers, large packages, or luggage. Bulky coats, worn in a New York winter, add further girth.[41]

The most ambitious turnstiles for warding off fare beaters are the so-called high-entry (HEET) models, which are harder to cheat, at

Figure 7. Subway turnstile, 1943. Image courtesy of New York Transit Museum.

least by using the usual techniques. Observing with McClain and me, sociologist Christine Nippert-Eng, while spending a few days with us as a consulting field researcher, named it "the garlic press" (see figure 9). It prevents people from jumping over or crawling under (although not from coupling up and squeezing through on a two-for-one entry). Whatever has been gained or lost in paid fares, there is a potential security cost. Entry or exit takes more time. Especially when a crowd is trying to exit, people can routinely be seen in line waiting their turn, many dozens of them sometimes. When the MTA originally installed the modern garlic press, they were in noncompliance with the state fire code, which requires greater capacity for exit. Any evacuation will take longer than in the past when the very vulnerabilities of the machinery for fare beating provided a kind of emergency safety valve.

Figure 8. Contemporary turnstile with sloped side panels and other features to ward off fare beating. Photo by Noah McClain.

In one vivid example of security backfire, police were delayed in responding to a platform shooting because they lacked MetroCards. They had to borrow fare cards from others to enter.[42] A man bled to death in the interim. Getting medical equipment in is made more difficult as is getting an ailing person out (imagine a gurney trying to pass through the turnstiles). In partial compensation, the MTA announced it would provide police (and fire officials) with MetroCards, and I've seen them use them.

Responding to the evident problem of the HEET turnstiles, the MTA installed additional emergency exits at platforms (see figure 10), untested in terms of evacuation numbers but certainly an improvement (although in some exit corridors the HEET models remain the only way out). The emergency exit doors have now been thoroughly tested in terms of unauthorized use. People hold open the exit doors (one kid pays to go in and holds the door open for his friends). Or

Figure 9. HEET turnstile, a.k.a. "garlic press." Photo by Noah McClain.

people happening on the scene take advantage of someone making his or her way out and grab the door to go in. Either way, non-payers gain free entry, making the whole system more permeable. As people push open the emergency doors, an alarm sounds for about thirty seconds. These alarms, contrary to early specifications, are connected to nothing. During busy times in busy stations, the opening of these doors creates a continuous piercing noise, adding an unwelcome sound to an already noisy environment—but also rendering incomprehensible, or potentially drowning out all together, any instructions made over loudspeakers or by emergency personnel.

System redesign may not end cheating or add to security, but it does alter the ways to cheat and the kind of people who do it. It turns out that a spent MetroCard can be bent in such way that fools the machine. People with the right skills stand at the turnstiles and sell swipes, usually well below the price of a normal ticket (one dollar instead of two).

Figure 10. Emergency exit. Photo by Noah McClain.

They can also guarantee themselves plenty of business by jamming up nearby MetroCard vending machines with chewing gum in the money and credit card slots, making their swipe business the only alternative for some customers. Subway cheaters have been pivotal in the city's anticrime quality of life campaigns; arresting them is thought an important part of the "broken windows" strategy to head off big crimes by stopping small ones. Perhaps because of the nature of the culprits (poor and black or Hispanic young men), authorities have increased penalties of fare evasion. By explicitly alleging that they were hurting efforts to fight terror,[43] the state legislature increased the penalty from simple violation to a misdemeanor to a crime with penalties of up to three months in jail for those selling swipes. In 2005 prosecutors began interpreting the crime as felony forgery, with a potential of seven years imprisonment. In the first six-month period, sixty-one individuals had such charges leveled against them.[44]

In any event, the cheating problem remains unsolved despite the legal and mechanical interventions. At one recent point in time, the authority estimated as many as 10 percent of passengers to be free riders. For the year 2008, the MTA had estimated lost revenue of $7 million; a recalibration led to a reestimate of losses at $27 million.[45] Unlike when phony tokens deposited in fare boxes could be counted up and proper estimates made of that mode of cheating, newer technologies eliminate such precision. Card-bending leaves no trail. As McClain elegantly summarizes, "Through these progressive iterations against fare evasion, the MTA in effect chased the problem out of accountability."[46] And this also means, as does happen with increasingly complex technical apparatuses, the organizers are hoisted on the petard of their own technical sophistication.

What to Do?

Our master guiding remediation principle for this chapter on the subways is the need to acknowledge everyday work practices and their instrumentation as guiding emergency response, in the past and for the future. The key ingredient is people's ordinary sense making that, in Karl Weick's term, must not be pressed into "collapse."[47] That means there needs to be integration of security precautions with the actual jobs that workers do and the manifold tasks they undertake. And, yes, this may mean accepting some fare beating and other inconvenient outcomes. Some of the specifics follow on.

- *Attend to workers' practices.*

Authorities should know and respect workers' repertoires for dealing with ordinary problems, including those based on their experiences with outside agents. If workers think their supervisors have bad information, they will not treat instructions given to them as bona fide. If they think their supervisors do not take into account job exigencies at hand, they will discount their directives. If they think bosses' initiatives are silly, they will deride rather than follow them. Remedies for dealing with such disjuncture are either to change the routines of the work situation (easing, for example, the vulnerabilities to human and mechanical challenges) or to make sure instructions take those contexts into account. Understanding the mundane is the best route for addressing the spectacular.

This orientation aligns well with that of the master security analyst and distinguished psychologist Elliot Aronson. From his studies of how people respond to campaigns to use condoms or conserve water, he concludes that they will ignore advice they do not consider trustworthy.[48] Besides their own routine experience with how the system operates, our subway workers know that the "See Something" campaign is preposterous and the signs that tell how many New Yorkers "Said Something" are no less ridiculous. Unless advice, following again Aronson, is concrete, doable, and credible, it will yield nothing—something judged to have been the case with the DHS color code warnings and the much-derided advice to have a "safe room" sealed off with duct tape.[49] An alternative to all of it is to abandon these earmarked and official security measures and concentrate on the daring alternative of just making things better—creating conditions that have the effect, quite counter-intuitively, of indeed adding to security. I list some of them for the subway system.

- *Improve ventilation.*

If there is a chemical attack or just a pile of newspapers set on fire, better ventilation will save people from smoke inhalation or visibility problems. In terms of workers' long-term health, less steel dust from wheels on tracks will reach their lungs. And, until an attack happens, everyone will breathe easier; perhaps also, if the MTA plays the infrastructure right, stations will be less stultifying in summer and cold in winter.

- *Fix communication.*

In an emergency, people need to be in touch, two (or more) ways if possible. During the World Trade Center disaster, almost all systems collapsed: radio, television, and even cell phones. The notoriously primitive loudspeaker systems of the New York subways (a routine joke among passengers), as well as the so-called emergency communications system for workers, leave much room for improvement. Indeed, the workers have only the most primitive means of communicating with one another and with their supervisors. Think of the poor subway conductor approaching Ground Zero, trying to get accurate information.

Passengers need to know if they are being advised to go to a particular exit, to remain in train cars or to vacate, to walk to the right or the left in the tunnel, away from the third rail. Like those in the factory pulling at the locked doors, they will, without being able to understand what is being said, perish. But again, good sound brings a bonus for the everyday. People will be less tense because they know which stop is theirs and which train has been rerouted to what track. New York's subways, at this writing, still do not have cell phone service in most places, in part out of authorities' fear that cell phones could be used to trigger a bomb. With cell phones, of course, individuals could also warn one another as well as authorities of real-time danger. Another way to improve communication would be to decrease noise levels. The antiquated cars and tracks screech and squeal; that creates a communication problem as well as aural discomfort. Modernizing them would be another everyday amenity that would also enhance safety.

• *Improve signage.*

Even on a good day, passengers have trouble figuring out which way to the street and, where there are multiple exits, to a particular street. If passengers are to learn that a particular exit is closed, they need clear markings so they can know, without hesitation, the best alternate route. And they need it, in a polyglot city, in a way that does not require knowledge of English. "Way finding," and there is something of a science of it, is another aspect of emergency infrastructure, and if it is good, it is good 24/7. It's a simple as this: people are safer when they know where they are. And the learning occurs in the routines of everyday use as individuals notice out of the corner of their eye the markings for exits and routings they ordinarily have no reason to use.

• *Make better stairways, corridors, ledges, and platforms.*

Because of their origin as separate competing private companies and also years of deferred maintenance, the subways are mazes of make-do add-ons. Anything that clarifies and simplifies helps with safety—and again makes things more pleasant day in and day out. Stairways should not be limited by their "behind the scenes" functionality like toilet stalls left in architectural silence. They should be made visible from floor areas to enhance safety as well as to let in light to enhance pleasantness. Vertical columns now press bodies close to the platform edge, complicate evacuation, and obscure vision. They need realignment or elimination.

Also making matters worse for quick exits, escalators break down and are blocked off routinely—not just put into disuse, but left with empty voids where landings need to be. So in a power failure they can not even be physically climbed. Lack of an escalator does not just create inconvenience and pain for the aged, injured, and those with luggage, it yields hazard by pressing passengers into a narrow stairway. Routine maintenance and fast repair thus not only make life better, but also make things more secure. In some instances, escalators are blocked off for years because the developers who pledged to maintain them as a condition for their high-rise building permits fail to do so. Unlike friends and neighbors, these developers are less likely to spontaneously jump in to help; they need to be forced into compliance.

• *Lose the HEET.*

Designed to protect against those who would cheat, turnstiles evolved to become increasingly difficult for all users, and the HEET models must go. They reflect a continuous narrowing of the range of human beings who can pass with comfort and safety. They embody the opposite of "universal design" —the idea that in serving the needs of everyone, including the disabled, a good device enhances benefits for all. One of the most famous lines of universal design products are Oxo Good Grips kitchen tools, first developed to aid people with limited hand capacity but then found to be better for everyone. And they now dominate in their industry.

An even more universal alternative is to get rid of all turnstiles completely. An honor system, as practiced in various parts of the world, sometimes involves random checks with penalties for nonpayment. In other cases, I have been told Norway is one such place; people there consider transportation enough of a government responsibility that authorities are not too exercised about some cheating. In the United States, much of central Portland, Oregon, is a free ride zone as are ski resort areas. It is life enhancing to just walk in and walk out, spared the ordeal of finding the card, making sure it has enough money on it, and assuming the swiping pose (while balancing packages, babies, and suitcases). Some business models have a built-in degree of theft ("shrinkage"), sparing employees from scrutinizing every customer. Not just luxury stores but also low-end chains and web-based merchants allow returns "no questions asked." Some museums have only "voluntary" contributions, but a great majority of patrons do choose to pay the "suggested" fee with the result that admission income is no less than 75 percent of what would result if everyone paid full price.[50] Meanwhile the museum can let up on fare policing, while expanding its audience and serving those with financial difficulty, which is useful for private fund-raising as well as public relations. At Apple stores, as with Singer sewing machine centers generations ago, people get free advice regardless of whether or not they buy. But buy they do; Apple's midtown NYC branch had the highest per-square-foot gross of any retail space on Fifth Avenue.[51] Starbucks lets pedestrians come in off the street to use the toilet and provides sofas and electricity for laptop users. Many if not most private employers let workers use paper, copy machines, and other goods and equipment for "unauthorized" private purposes. It is all part of the slack.

For subway workers, doing away with the turnstile would also mean a nicer day. They would have less exposure to customers' anger about cards that don't work, machines that malfunction, and rides that were paid for but not taken. Staff could devote more of their time to other aspects of the job, including watching out for real trouble. Homeless people, who ride the rails all day, could more easily exit and reenter the stations, which in turn would allow them to come up and relieve themselves in a more appropriate place rather than urinating or defecating in the subway system.

• *Light better.*

Enhanced lighting and durability, particularly at exits, would pay off in a moment of crisis, as well as improving conditions at most all other times as well. Will there be disorienting glare or will lighting be crisp? What objects are lit, not just on the platforms but also in the system's innards that people may have to negotiate? And what is the system redundancy, not just to handle catastrophe but routine troubles of power failure? Improve on all fronts.

• *Clean it up.*

Ordinary construction crews know that an unkempt workplace causes accidents. In the subway, dirt and grime obscure signage and may cause people to slip and fall. Helter-skelter environments facilitate chaos in times of emergency. After exhaustive investigation, officials judged track debris the culprit for the 2001 Baltimore freight-train-tunnel derailment and fire that lasted five days and virtually shut down central Baltimore. In the chemical conflagration 2,554 gallons of hydrochloric acid were released from one of the tank cars, with additional toxic chemicals spilled from others.[52] The solution for Baltimore, as for subways in general, is to implement regular inspections and a higher standard of order and cleanliness.

Greater cleanliness brings a much-used public facility into conformity with how people routinely live in their private spaces. Societies differ in the standards of comfort and pleasantness between the public and private. In the United States the discrepancy seems extreme, with the standard for public facilities, especially the subway environment, falling far below the level of cleanliness that residents achieve in maintaining their own living space. This means that people notice the low standard of the public facility and might well support a higher one, and pick up some nonobvious security benefit in the bargain.

• *Keep humans.*

Finally, what about the MTA employees? We learned that real security is provided by context-sensitive inventiveness by people not charged with doing security—without benefit of special budgets and without armaments or much coercive authority over others. In the language of security organizational analysts, they have "high mission valence"[53]—they are committed to keeping people safe. And they do countless things for others along the way, including giving advice to tourists, uniting families with lost children, and compensating for decrepit and faulty artifacts. They are security bulwarks. Humans alone scan, know, and remedy. Somewhere and somehow their presence and goodwill seem key.

Whether in respecting the role of the workers or upgrading the passenger experience, the best way to enact security is by doing things that do not look like security at all. Improving subways to make them more efficient and pleasant can make them more resilient and safe. Such improvements are not *alternatives* to investing in security, but the most important aspects of it. Officials can lessen danger by implementing measures that are inconspicuous, passive vis-à-vis both customers and workers. It is akin to what has been called "crime prevention through environmental design,"[54] but more generous in spirit, delivering benefits that can stand even if there are no attacks, which is the overwhelming likelihood for any given context. Improvement in everyday infrastructure bodes well for resilience, and at relatively low cost. A gentler subway provides for people and enables them toward solutions of their own safety and survival.

The Cover Design

ART, TECHNOLOGY, AND SCIENCE: NOTES ON THEIR HISTORICAL INTERACTION

CYRIL STANLEY SMITH

Introduction

It is misleading to divide human actions into "art," "science," or "technology," for the artist has something of the scientist in him, and the engineer of both, and the very meaning of these terms varies with time so that analysis can easily degenerate into semantics. Nevertheless, one man may be mainly motivated by a desire to promote utility, while others may seek intellectual understanding or aesthetic experience. The study of interplay among these is not only interesting but is necessary for suggesting routes out of our present social confusion.

Humanists have shown a widespread disregard for technology's role in human affairs, but if they had seen technology as an eminently human experience, they could have better guided society's choice of objectives and controls. Civilization has been an ecological process with interacting contributions coming from an infinite diversity of individual human characteristics and social institutions. As historians have turned away from their older concern with the great movements headed by kings, generals, or businessmen, they have naturally emphasized the role of people like themselves (scientists and other intellectuals), and they have, until recently, largely disregarded the rather messy technology that has been associated with virtually every important historical change and which continually impinges directly upon everyman in his daily

DR. SMITH, institute professor emeritus at the Massachusetts Institute of Technology, is famous as both a metallurgist and historian of science and technology. He is a former president of the Society for the History of Technology and a recipient of the Leonardo da Vinci Medal. This article is based on a paper read at the University of Oklahoma Symposium on the History of Science and Technology, April 1969. The author is grateful to his colleagues Arthur Steinberg and Heather Lechtman for discussions in the general area of this paper, and to Janet Abramowicz, Lynwood Bryant, and Julia Phelps for critical comments on an earlier draft. Dr. Smith's work in the history of technology is supported in part by the National Endowment for the Humanities, grant no. H68-0-86.

life. Neither religious conviction nor institutional conservatism has, until today, sensed in technology a peril sufficient to prompt an examination of its nature and its growth. Certainly, at the extremes, the concepts of the cosmos and of the ultimate nature of matter developed by philosophers and scientists are of overriding importance, for they have basically influenced man's opinion of himself: Men have gone to the stake for their ideas on the nature of the universe, and all men know of it. Ideas on ultimate atomism have aroused bitter philosophic debate. Conversely, however, anyone who considers the nature of materials, advocates a new way of making pottery, or advances a new theory of the hardening of steel meets with both intellectual and popular indifference. Yet the voyage to the moon depends on men making metal as well as on computations based on the theories of Newton and Einstein.

Art, Techniques, and Materials

The present paper is an outcome of my realization, some years ago, that many of the primary sources I had selected for a study of the history of metallurgy were objects in art museums. Though materials are not all of technology, they have been intimately related to man's activities throughout all of history and much of prehistory and therefore provide an excellent basis for a study of some of man's most interesting characteristics under greatly different social and cultural conditions. A materials-oriented view of history may overemphasize the association of technology with art; yet it was precisely the artist's search for a continued diversity of materials that gave this branch of technology its early start and continued liveliness despite an inner complexity which precluded scientific scrutiny until very recently.

Several writers have discussed the manifest interactions between artistic expression and the basic view of the world embodied in contemporary scientific or religious concepts. Such interactions certainly exist at the highest level of insight, but artists have had far more intimate and continuing association with technology than they have had with science. In turn, the attitudes, needs, and achievements of artists have provided a continuing stimulus to technological discovery and, via technology, have served to bring to a reluctant scientific attention many aspects of the complex structure and nature of matter that simplistic science would have liked to ignore. The antecedents of today's flourishing solid-state physics lie in the decorative arts. One must conclude that creative discovery in any field is a matter for the whole man, not his intellect alone. Though it occurs in an individual mind, it is strongly interactive with society and tends to seek out the least rigid parts of a community structure.

Leonardo da Vinci said in his treatise on painting: "Those who are in love with practice without science are like a sailor who gets into a ship without rudder or compass and who never can be certain where he is going."[1] At the same time, Leonardo strongly opposed the view that knowledge that is both born and consummated in the mind is enough: "It seems to me that all sciences are vague and full of errors that are not born of experience . . . , that do not at their origin, middle or end pass through any of the five senses."[2] And, of course, all his extant works reflect continual interplay between sensual experience and intellectual analysis. The same view is to be found in the writings of many scientists, though for most of the last three centuries science has rightly been more concerned with the unreliability of the senses than with their essential contribution to whatever knowledge human beings can acquire.

When discussing the new routes to the understanding of nature in the preface to his *Micrographia* (1665), Robert Hooke remarks:[3] "So many are the links upon which the true Philosophy depends, of which, if any one be loose, or weak, the whole chain is in danger of being dissolv'd; it is to begin with the Hands and Eyes, and to proceed on through the Memory, to be continued by the Reason; nor is it to stop there, but to come about to the Hands and Eyes again, and so, by a continual passage round from one Faculty to another, it is to be maintained in life and strength." Hooke believed that the advancement of knowledge depended upon both the senses and the intellect—upon the mind, the hand, and the eye in cooperation. His writings repeatedly reflect his obvious enjoyment of natural phenomena and his intuitive understanding of them. However, Hooke's slightly younger contemporary Isaac Newton was engaged in demonstrating the great power of mathematical science and setting the stage for three centuries of superbly unfolding knowledge based on the belief that the senses are unreliable and that science advances best if, at any one time, it is limited to those small areas in which rigorous methods can be applied. Though the domain accessible to such science is steadily expanding, there remain many important aspects of natural and man-made systems that are too complicated for complete analysis. The present-day political and intellectual unrest reflects increasing awareness that the scientist's understanding of things "in principle" is not enough. The more holistic view of the Renaissance artist may be returning—though whether it will be put into

[1] Leonardo da Vinci, *The Literary Works of Leonardo da Vinci . . .*, ed. Jean Paul Richter and Irma A. Richter (London, 1936), 1:119.

[2] Ibid., 1:25–26.

[3] Robert Hooke, *Micrographia* (London, 1665), preface (unpaginated).

practice by people who allow themselves to be called artists is another question.

Just as the meanings of the words "art," "science," and "technology" have varied greatly throughout history, so has the role in society of the various practitioners. Perhaps technology has been the most constant in its aims. Science has encompassed many different approaches to the collection and analysis of data, just as art, in different places and periods, has combined in vastly different degrees the functions of decoration, symbolism, illustration for didactic purposes, the projection of feeling, and (by no means the least important) pure enjoyment. In what follows the "art" may sometimes be of a kind beneath the notice of an art historian, but it will always be concerned with a man's doing something that is not strictly necessary for the performance of a function, something extra done to give enjoyment to the producer himself and usually also to others who subsequently come in contact with his work.

Not all peoples have regarded "art" as a separable human activity, and the self-conscious production of paintings, sculpture, and *objets d'art,* like the organized commerce in them, has by no means always occupied the privileged place that it has had in Europe since the Renaissance. Most of what follows is concerned with the decorative arts—those arts relegated to the minor category in most museums today—although it might be remarked that the best of today's nonobjective paintings have more in common with sensitively wrought useful objects of ceramic and metal than they have with many of the "fine" arts displayed on museum walls.

There is some analogy between the exploration and exploitation of the materials of nature in chalcolithic times and earlier, the detailed exploration of the forms of nature that followed increased representational skill in the 13th and 14th centuries, and the experiments with perspective, light, and shadow in the Italian Renaissance. The driving force in all three was an essentially scientific curiosity directed to the discovery of some fairly practical means of achieving an aesthetic end.

The relation between art and the artist's materials was well discussed by Henri Focillon.[4] Remarking that art is bound to weight, density, light, and color, he says that it is borne along by the very matter it has sworn to repudiate: matter in its raw state "evokes, suggests and propagates other forms according to its own laws." The ceramics of the Far East appear to be "less the work of a potter than a marvellous conglomerate created by subterranean fire or accident." The raw stuff of

[4] Henri Focillon, *Vie des formes* (Paris, 1947); English translation, *The Life of Forms in Art* (New York, 1948), passim, esp. pp. 31–41, 76. The illustrations in the English edition are poorly selected to reinforce the author's points.

[Chinese ink brush painting] partakes of both water and smoke . . . yet . . . such a painting possesses the extraordinary secret of being able to stabilize these elements and at the same time leave them fluid and imponderable." Though overemphasis on technique is clearly dangerous, Focillon believed that "the observation of technical phenomena not only guarantees a certain objectivity to [the studies of] a historian but affords an entrance in the very heart of the problem by presenting it in the same terms and from the same point of view as it is presented to the artist." In discussing the artist's various techniques to get different qualities of line, shadings, and graduations, "such alchemy does not, as is commonly supposed, merely develop the stereotyped form of an inner vision: it constructs the vision itself; gives it body and enlarges its perspectives."

Technique, of course, mainly gives details of form, not the gross outlines and balances. Nevertheless, much of the refinement of an artist's vision as he works toward its realization comes from his interaction with his materials. The whole quality of a line and surface depends upon both the material and the tool as well as upon the artist's hand, whose movements they subtly control. Compare the same pictorial concept as it is realized in different media—with a brush in oils, watercolor, or tempera on canvas, wood, or paper; by printing from a metal plate with intaglio lines made by etching or engraving or from surfaces left in relief on a chiseled wood block; or by repoussé work, tracing or otherwise working directly on the final metal surface. It is understandable that those students who must work from reproductions of works of art are usually more interested in iconography than in the more subtle questions of technique and quality, but it is regrettable that technical ignorance should so frequently prevent art historians from considering the whole experience of the artist. In much the same way, science historians have tended to overlook the less logical side of science.

The Discovery of the Properties of Matter

In studying man's earliest history, when the evocative qualities of certain forms and the power of symbolism in nonrandom shapes and sounds was being discovered, it is difficult to separate things done for "pure" aesthetic enjoyment from those done for some real or imagined "practical" purpose. The man who selected for admiration a beautifully shaped and textured stone was yielding to a purely aesthetic motivation, but the man who molded clay into a fertility figurine was simultaneously an artist, a scientist learning to understand the properties of matter, and a technologist using these properties to achieve a definite purpose. Supposedly most of the innumerable fertility figures recovered by the archaeologist's spade from periods even before 20,000 B.C. were made as a kind of industry, acquired for reasons of fashion, and employed practically to make more probable some desired result. This does not, however, destroy their fundamental aesthetic quality.

More important is the fact that in the earlier stage of discovery, first of form and later of materials that, once shaped, would retain desirable form, the motive can hardly have been other than simply curiosity, a desire to discover some of the properties of matter for the purpose of internal satisfaction. Paradoxically man's capacity for aesthetic enjoyment may have been his most practical characteristic, for it is at the root of his discovery of the world about him, and it makes him want to live. It may even have made man himself, for, to elaborate a remark by the poet Nabakov, it seems likely that verbal language (to which anthropologists now assign vast evolutionary advantage) was simply a refined use of the form-appreciating capabilities first made manifest in singing and dancing.

A natural step after the collection and admiration of unusual natural stones and animal or vegetable debris would have been the use of the properties of some natural materials to produce unnatural shapes and textures in others. This supposedly began by matching the hard cutting edge of stone to softer wood, hide, sinew, and bone, and was followed by the discovery and exploitation of the special properties of a host of substances. The last were mainly minerals that could be ground and used as pigments, undoubtedly far more for the decoration of the body and other long-perished surfaces than for the incredibly preserved cave paintings that we admire so greatly today. It is not only the nature of the record that makes one feel the joy that early man took in the discovery of the properties of materials. The cracking propensity of different stones, the plasticity of moist clay, the fine granular color of pigments were all used for what they are and appreciated directly by the senses in shaping or in use.

Aesthetically satisfactory forms have repeatedly developed from interaction between cultural requirements and the real properties of a new material or technique: the forms are not just superimposed. A returning sensitivity to this is at least partially behind the present passion for primitive art, for a simpler technology makes the properties of materials more evident.

Over and over again scientifically important properties of matter and technologically important ways of making and using them have been discovered or developed in an environment which suggests the dominance of aesthetic motivation. The presence of flowers in Neanderthal graves[5] suggests that the transplanting of flowers for enjoyment preceded the development of agricultural technology for food supply. The first use of both ceramics and metals occurs in decorative objects. Fire-hardened figurines of clay precede fired pots in many Middle Eastern archaeological sites. The 7th millennium B.C. copper dress ornaments and beads at Chatal Huyuk in Anatolia and at Ali Kosh in Iran considerably precede the use of copper for weapons, though the useful needle appears early. Although there is some evidence for earlier pyrotechnological experiments with ores, the replacement of simply hammered native copper by smelted metal did not occur until about the time that copper oxide was being used in blue glazes on ceramics, though probably only after high temperatures had become available for firing useful ceramic sickle blades.[6]

The modern metallurgist uses alloying elements to strengthen metals and to lower their melting point; he cold-works to harden and anneals to soften. He uses their differing chemical reactivities, immiscibilities, and surface energies in refining and joining processes. The discovery of all these effects is very old. To take a single point in history, an examination of the jewelry and other metal objects from the famous Royal graves at Ur,[7] dated about 2600 B.C. (figs. 1 and 2), reveals knowledge of virtually every type of metallurgical phenomenon except the hardening of steel that was exploited by technologists in the entire period up to the end of the 19th century A.D. One must not, of course, overlook the fact that royal burial objects are far from being representative samples of contemporary use of any material. The court would appropriate the best work to its own ends, but just for this reason it provides the best index of both the most novel and the most sophisticated techniques.

[5] Arlette Leroi-Gourhan, "Le Neanderthalien IV de Shanidar," *Bulletin de la Société Préhistorique Française* (Comptes rendus séances mensuel) 65 (1968):79–83. See also Walter Sullivan, *New York Times,* June 13, 1968.

[6] Though blue frit is characteristically an Egyptian product, the earliest examples of it are two frit vessels and some seals and amulets from Mesopotamia, in the Tall Halaf levels at Tall Arpachiyah near Ninevah, a period which lasted from roughly 4900 to 4300 B.C. (M. E. L. Mallowan and J. C. Rose, "Excavations at Tall Arpachiyah 1933," *Iraq* 2 [1933–35]:1–178. For a later appearance of it see Hans Wulff et al., "Egyptian Faience: A Possible Survival in Iran," *Archaeology* 21 (1968):98–107.

[7] H. J. Plenderleith, "Metals and Metal Technique," in *Ur Excavations,* ed. C. L. Wooley, vol. 2, *The Royal Cemetery* (London, 1934), pp. 284–310, also plates 138 and 162. For a discussion of early metallurgy, T. A. Wertime, "Man's First Encounters with Metallurgy," *Science* 146 (1964):1257– C. S. Smith, "Materials in Civilization and Science," *Science* 148 (1965):908–17. The best comprehensive history of metallurgy is that by Leslie Aitchison, *A History of Metals,* 2 vols. (London, 1961).

The transition from copper ornaments to axes and swords of bronze in the 4th millennium B.C. was paralleled in the 15th and 16th centuries A.D. by the transition from the casting of monumental bronze doors, statuary, and especially bells, to the casting of cannon. If the objects themselves are not sufficient evidence, a comparison of the vivid circumstantial account of bell founding given by the early 12th-century

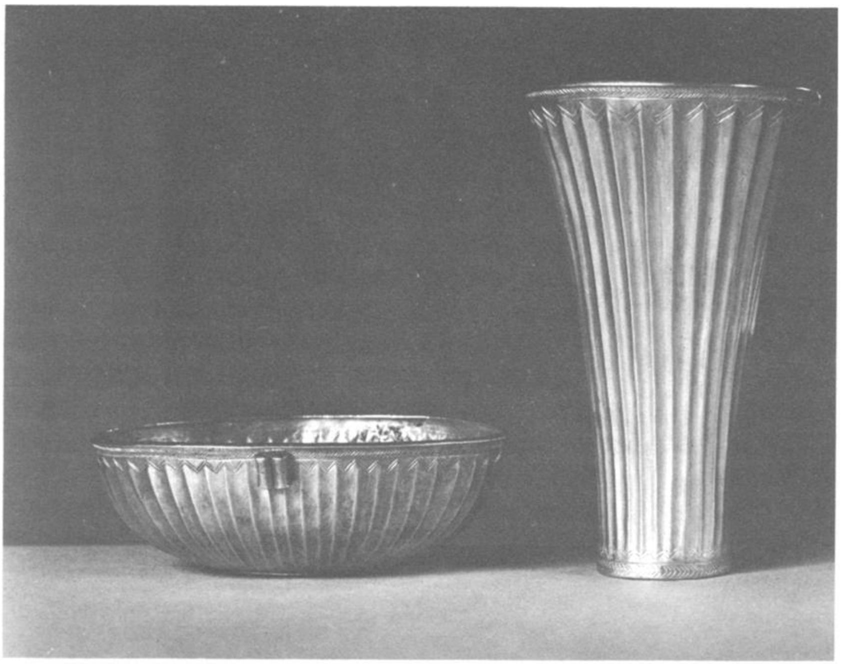


FIG. 1.—Gold beaker and cup made by raising from sheet metal, decorated by repoussé work and tracing. Cup height, 15.5 cm. From the Royal Graves at Ur, ca. 2600 B.C. (Photo courtesy University Museum, Philadelphia.)

artist-craftsman Theophilus[8] with the discussion of the casting of cannon by the eminently practical Biringuccio[9] some four centuries later will show how much the warrior depended upon the churchmen's technique. To be sure, existing technology is applied to whatever need may

[8] Theophilus, "De diversis artibus," manuscript treatise, ca. A.D. 1123; Latin text and translation by C. R. Dodwell (London, 1961); translation with technical notes by J. G. Hawthorne and C. S. Smith (Chicago, 1963). Chapters 85–87 deal with bell casting.

[9] Vannoccio Biringuccio, *De la pirotechnia* (Venice, 1540); trans. C. S. Smith and M. T. Gnudi (New York, 1942), pp. 2550–60.

be seen by a government or people: my point is only that the *invention* of a technique has, until recently, been more likely to occur in an aesthetically sensitive environment than in a practical one. We will see later that even the development of efficient quantity-production methods owed much to the art industries, if not directly to the artist.

Technology's debt to the artist is inseparable from the converse. Though both the most exquisite and the most ugly objects can be made

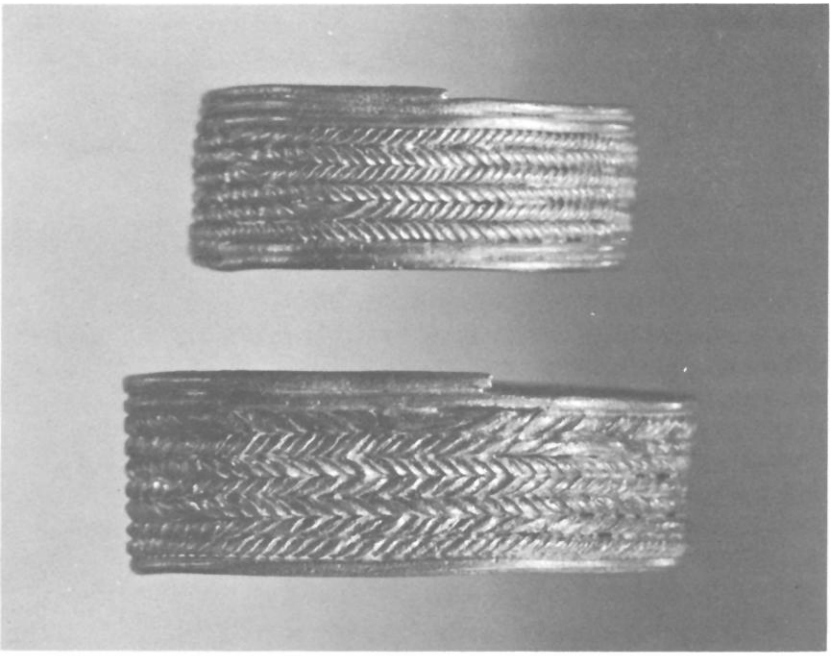


FIG. 2.—Gold rings made from square and round wire almost invisibly soldered, 2.1 and 1.7 cm diameter. From the Royal Graves at Ur, ca. 2600 B.C. (Photo courtesy University Museum, Philadelphia.)

with the same technique, technique is essential to beauty. The technique of the artist merges by invisible stages into the technology of his materials. Among the best examples of this are superb cast bronzes of Shang Dynasty China.[10] The earliest ones reflect an advanced ceramic technology, so necessary in making the molds, and have linear decora-

[10] Noel Barnard, *Bronze Casting and Bronze Alloys in Ancient China* (Canberra and Nagoya, 1961); R. J. Gettens, *The Freer Chinese Bronzes*, vol. 2, *Technical Studies* (Washington, D.C., 1970). A superb collection of bronzes illustrating stylistic development was assembled for an Asia House exhibition in 1968. Every item is described and illustrated in the catalogue (Max Loehr, *Ritual Vessels of Bronze Age China* [New York, 1968]).

tion obviously cut into the mold (fig. 3). The almost brutal strength that characterizes the later bronzes of the Shang and early Chou periods, the flanges, and almost every aspect of their form arise in a direct interplay between design and the practical details of the foundry. The molds were divided into a number of sections that would have produced unpleasant breaks in the surface decor had these not been designed for division, and leaky mold joints would have produced ugly fins if these had not been exaggerated into flanges the edges of which could easily

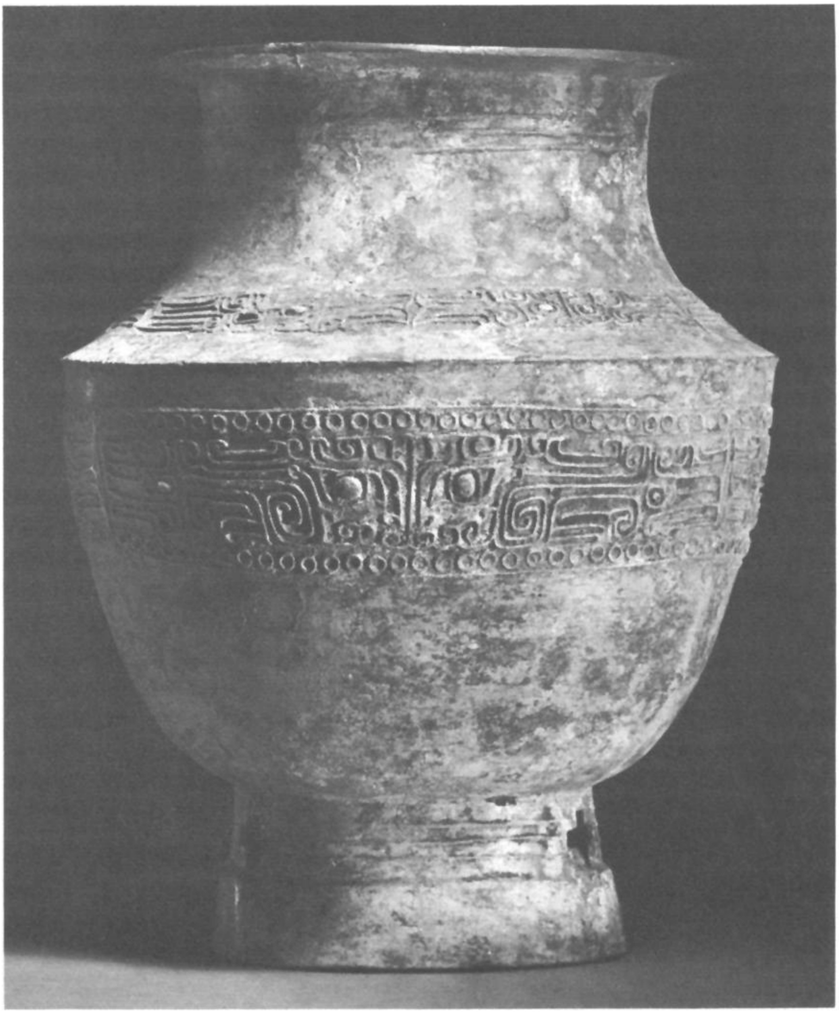


FIG. 3.—Chinese cast bronze ceremonial vessel, type *tsun*. Early Shang Dynasty. Height, 27 cm. (Photo courtesy Arthur M. Sackler Collection, New York.)

be dressed (fig. 4). The attractive difference in quality of the fine and the bold intaglio lines probably arises in the technical difference between carving the former directly into the mold surface and making the latter by applying convex lines of clay to a molded concavity. Still later comes the building up of designs from a few units by the use of some method of three-dimensional replication (fig. 5)—a clear forerunner of the printing process.

FIG. 4.—Chinese cast bronze ceremonial vessel, type *ting*. Late Shang Dynasty or early Chou (11th–10th century B.C.) (Courtesy Fogg Art Museum, Harvard University.)

In view of the centrally important role that welding plays in today's space-age structures, it is interesting to note the facility with which Greek and Roman founders welded together the parts of their statuary. Almost any classical bronzes, when closely studied, reveal some patching of foundry defects, but recent studies[11] have uncovered the widespread use of some process (not yet fully understood but clearly involving the running-in of superheated molten metal) for making joins

FIG. 5.—Chinese cast bronze bell (detail). Chou Dynasty. The design is built up in three successive stages of replication. (Courtesy Freer Gallery.)

[11] S. Delbourgo, "L'étude au laboratoire d'une statue découverte à Agde," *Bulletin du Laboratoire du Musée du Louvre* (1966), pp. 7–12; H. Lechtman and A. Steinberg, "Bronze Joining: A Study in Ancient Technology," in *Art and Technology: A Symposium on Classical Bronzes*, ed. S. Doerringer (Cambridge, Mass., 1971); A. Steinberg, "Joining Methods in Large Bronze Statues, in *The Application of Science in the Examination of Works of Art* [Seminar III], ed. W. J. Young (Boston, in press).

between precast parts, with an accuracy and permanence that would challenge a modern welder (fig. 6).

The technique of casting-on preceded this joining of long seams and was widely exploited in the Middle East, in Europe,[12] and in the Far East[13] because of the freedom that it gave to the designer (fig. 7). It also enabled different metals to be combined, as in the application of an elaborately detailed cast bronze handle to a serviceable steel blade—a

Fig. 6.—Welded joint in Roman cast bronze statue, natural size. (Courtesy Arthur Steinberg and the Museum of Fine Arts, Boston.)

literal welding of beauty and utility—and in the combination of cold-worked and cast bronze.

In later times the little mouth blowpipe used by the goldsmith in his soldering operations was adapted to laboratory use, first for the examination of ores and later as the basis of the first comprehensive scheme

[12] Hans Drescher, *Der Überfangguss* (Mainz, 1958).

[13] R. F. Gettens, "Joining Methods in . . . Ancient Chinese Bronze Ceremonial Vessels," in W. J. Young, ed., pp. 205–17. Also see n. 10 above.

of qualitative chemical analysis.[14] The larger blast lamps worked with foot bellows that were used for making glass beads and for decorating them with colored enamels (fig. 8) led directly to the oxy-gas blowpipe. This was first used in high-temperature research about 1782, became fully commercial with the melting of platinum in the 1850s, and finally became the modern welding torch.

The decoration of pottery with colored pigments and later with

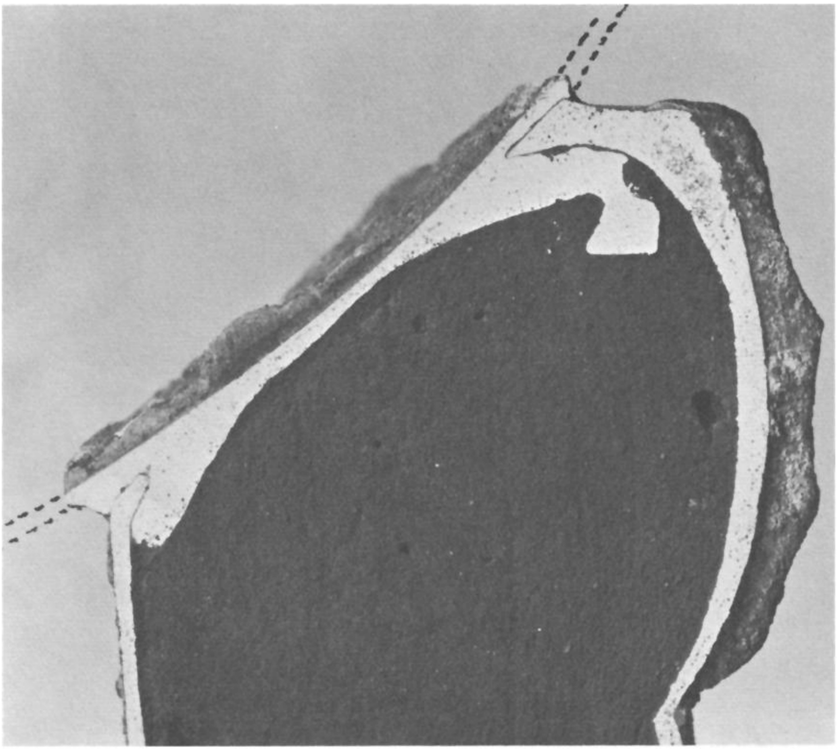


FIG. 7.—Section of cast-on leg in Late Chou bronze vessel. (Photo by John Gettens, courtesy Freer Gallery.)

[14] The history of the blowpipe has yet to be written. Blowing through pipes to urge charcoal (?) fires for smelting and melting doubtless preceded the use of bellows. The small mouth blowpipe with a lamp or candle was used by jewelers for local soldering operations and in the 17th century was suggested for testing ores. Comprehensive schemes of chemical analysis based on it were developed in Sweden in the last half of the 18th century but were slowly displaced, except in the field, by wet methods of analysis. Blowpipe analysis was regarded as an essential part of the training of a young chemist until very recently, and I still trace my feel for the nature of most chemical substances and reactions to my work with the blowpipe as a schoolboy.

glazes repeatedly brought to man's attention the chemical diversity of natural minerals and led to new techniques. The cementation process that was probably used in the 4th millennium B.C. to make Egyptian blue frit (faience) involves very subtle behavior of alkalies and silicates in differential contact with lime and silica surfaces.[15] It is highly probable that it gave rise directly to the manufacture of the first "sand-cored" glass vessels (which probably had a calcareous, not a siliceous,

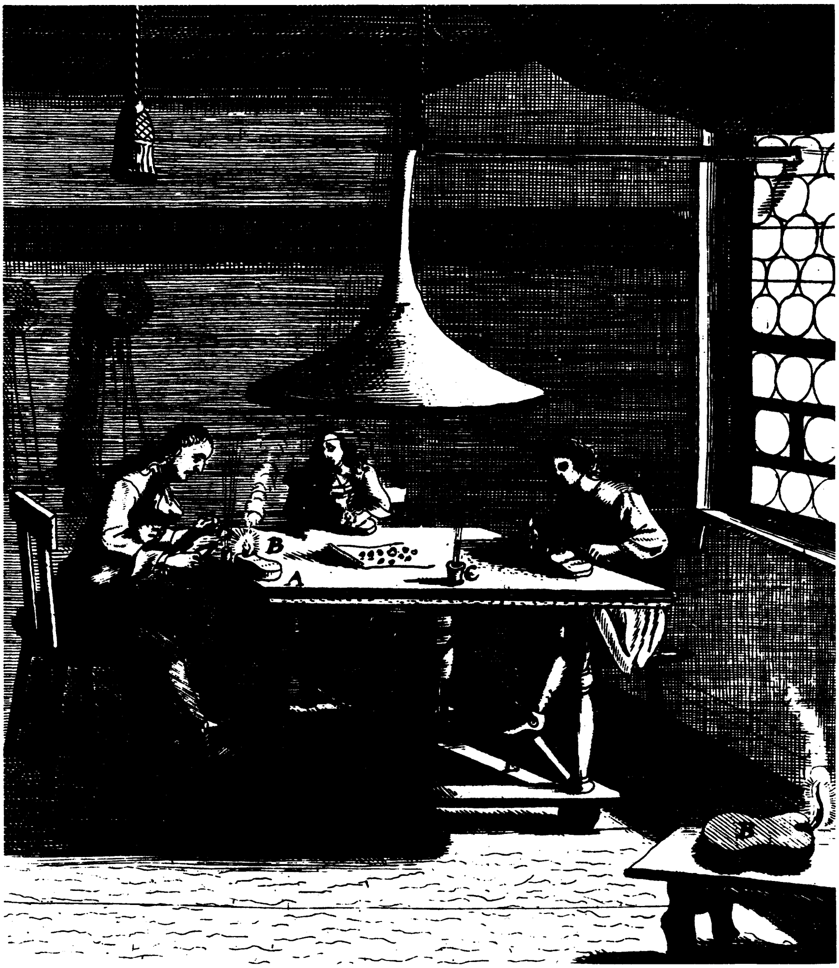


FIG. 8.—Blast lamps being used in making colored enameled glass beads. Johan Kunckel, *Ars vitraria experimentalis* (1679). (Photo courtesy Corning Museum of Glass.)

[15] Wulff et al. (see n. 6).

core). Though its relationship to early metallurgy has not yet been explored, this cementation process gives hints of the way in which the first alloys may have been made. It may relate to the smelting of complex sulphide ores by the use of highly alkaline fluxes, to say nothing of its later use in the soldering, parting, and coloring of gold and eventually in the making of brass and steel.

In the 18th century, European desire to duplicate beautiful porcelain from the Orient inspired not only geological search but also experiments in high-temperature chemistry and the development of the first realistic methods of chemical analysis for anything but the precious metals. Reports of the large-scale operations at Ching-te-Chen may have inspired the integration of mass-production operations at Wedgwood's factory in Staffordshire, and, at the other end of the spectrum of knowledge, it was an interest in porcelain that led to Réaumur's studies of the devitrification of glass, which later played a role in the understanding of lava and the development of Hutton's plutonic theory of the earth.[16]

Chinese fireworks for pleasurable celebration inspired more diverse chemical experimentation than did military explosives. Today's rocket ships and missiles are an outgrowth of fun-fireworks, and their guidance systems depend on knowledge first acquired from that ubiquitous toy, the top. All optical devices have their roots in the polishing of ancient mirrors and the cutting of accurate facets on gems for a more decorative glitter. The chemist's borax-bead test, now alas passé, arose from the use of metal oxides in making stained glass windows and colored enamels (as well as fake gem stones).

Colors and chemistry are inseparable. The earliest pigments were naturally occurring minerals, but the preparation of artificial ones, such as red and white lead, verdigris, and marvelous sublimed vermilion, mark a chemical industry in classical times. The subtleties of surface tension on which the modern flotation process for the beneficiation of ores depends were first used in the purification of lapis lazuli to give fine ultramarine.[17] The important metal powder industry of today began

[16] C. S. Smith, "Porcelain and Plutonism," in *Toward a History of Geology: Proceedings of the New Hampshire Conference on the History of Geology, 1967*, ed. C. J. Schneer (Cambridge, Mass., 1969), pp. 317–638.

[17] Cennino d'A. Cennini, "On the Character of Ultramarine Blue and How to Make It," *Il libro dell'arte* [ca. A.D. 1400], ed. and trans. D. V. Thompson, Jr. (New Haven, Conn.: 1933), chap. 62. A more complete account of the flotation process was recorded slightly later in the Bologna manuscript reported and translated by Mary P. Merrifield, *Original Treatises Dating from the XIIth to XVIIIth Centuries on the Arts of Painting in Oil*, 2 vols. (London, 1849).

with gold ink.[18] Art historians rarely go behind the blue and gold splendor and the iconography of a medieval illuminated manuscript to see the ingenious technology that made it possible and that reflects men's lives on another, no less necessary, level. It is the same with organic dyes: think of the chemical knowledge behind an oriental rug or an emperor's robes! The chemist's indicators and his eventual awareness of *p*H came directly from the chameleon colors of the miniature-painter's turnsole. . . . The list is endless.

The Development of Mechanical Technology

The relation between design, structural engineering, and knowledge of materials in architecture is a well-known example of the inseparability of aesthetic and technological factors. Here it must suffice to make only the passing comment that it has usually been nonutilitarian structures such as temples and monuments that have stretched the limits of existing techniques and led to the development of new ones.[19]

The popular belief that technology is recent is partly based on the fact that intricate machines were, in fact, slow to develop. The advanced knowledge of materials in the ancient world was not paralleled by mechanical devices of seemingly comparable ease of discovery. The ancient military devices (which have usually followed not far behind aesthetic needs in promoting discovery) and hoisting machines of importance to the builder are all relatively simple. Mechanical devices of any intricacy appear only as toys, as aids to priestly deceptions, or as theatrical machinery. It was not utility in the usual sense—though it may have been a search for the public's money—that prompted the mildly ingenious devices described by Hero of Alexandria.[20] It may be, as has often been suggested, that the availability of cheap labor rendered the Persians, Greeks, and Romans unable to appreciate the advantages of mechanical power; but their failure to develop other types of intricate mechanisms is, I believe, attributable to the fact that the aesthetic

[18] C. S. Smith, "The Early Development of Powder Metallurgy," in *Powder Metallurgy*, ed. John Wulff (Cleveland, 1942), pp. 4–17; Shirley Alexander, "Medieval Recipes Describing the Use of Metals in Manuscripts," *Marsyas* 12 (1966):34–51; Alexander, "Base and Noble Metals in Illumination," *Natural History* 74, no. 10 (1965):31–39.

[19] S. Giedion, *Space, Time and Architecture* (Cambridge, Mass., 1953); Norman Davey, *History of Building Materials* (London, 1961); L. F. Salzman, *Building in England Down to 1540* (Oxford, 1952); Marion E. Blake, *Ancient Roman Construction in Italy* (Washington, D.C., 1947).

[20] A. G. Drachmann, *The Mechanical Technology of Greek and Roman Antiquity* (Copenhagen, 1963); B. S. Brumbaugh, *Ancient Greek Gadgets and Machines* (New York, 1966).

rewards to beginning experimentation by the curious in this area are not large. Indeed, for simple mechanical experiments to be intriguing, they require a kind of overlay of intellectual analysis: they are too easily reproducible to provide a rich and varied sensual experience of the kind that comes directly from play with minerals, fire, and colors. Not until the mid-20th century have artists shown much desire to experiment with machinery, and their efforts sometimes seem to be more directed toward catching up with and exploiting the technologists' world than toward leading it.

The association of the earliest clocks with mechanical automata was a natural one, for, with the possible exception of organ makers, only the makers of automata had the necessary skill and sense of mechanism.[21]

Machine tools, like materials and mechanisms, had a period of prehistory within the decorative arts. The earliest is probably the rotary drill, which, though it was perhaps developed for hafting axes, found wider use in making beads, seals, stone pots, and sculpture.[22] The inverse geometric motion of material against a fixed tool begins with the potter's wheel and progresses to the simple lathes that supposedly produced the soft-stone products of Glastonbury and the Roman bronze objects such as mirrors and pots having decorative bottoms with deep, heat-catching circular grooves (fig. 9). Then followed Theophilus's 12th-century description of lathes for turning bell molds as well as for the molds of pewter pots and the metal pots themselves. By this time rotary motion was commonly used in the grindstone. The first machine with intermittent motion after the Oriental rice-pounding mill is Theophilus's little device for cutting the criss-cross ground for decorative overlay of precious metal on iron. The cam- and template-guided lathes of Jacques Besson (1578) not only cut screws but also turned decorative woodwork of great variety. They were followed by the ornamental turning lathes of the 17th–19th centuries, used mainly by gentlemen hobbyists and for decorating gold snuff boxes. These were devices of great mechanical ingenuity applied to a mechanically trivial purpose (fig. 10); nevertheless, they provided the experimental environment in which definable motions were generated not only as a basis for instrument making and later industrial machinery but also to disseminate a feeling for the composition of mathematical curves.[23] The toy-

[21] Alfred Chapuis and Edmond Droz, *Automata: A Historical Account and Technical Study* (Neuchatel, 1958). See also *Les automates dans les oeuvres d'imagination* (Neuchatel, 1947) and several other works by Chapuis.

[22] V. Gordon Childe, "Rotary Motion," in *A History of Technology*, ed. Charles Singer et al. (London, 1954), 1:187–215.

[23] The apogee of ornamental turning and its gadgetry is recorded in John Jacob Holtzapffel, *Turning and Mechanical Manipulation*, vol. 5, *The Principles and Prac-*

like nature of these lathes resulted in their being rather briefly dismissed in the standard machine-tool histories,[24] but it is easy to see how the desire to produce a decorative effect was once more the motivation for the discovery of phenomena that would later be applied to more serious purposes.

FIG. 9.—Cast bronze funerary bucket with lathe-tool marks on bottom. Roman, ca. A.D. 200. 25 cm diameter. (Courtesy W. J. Young, Museum of Fine Arts, Boston.)

---

tices of Ornamental or Complex Turnings (London, 1884). The book by A. K. Snowman, *Eighteenth Century Gold Boxes of Europe* (London, 1966), illustrates innumerable surfaces, both enamelled and plain, whose decorative charm derives directly from engine turning.

[24] Robert S. Woodbury, *A History of the Lathe to 1850* (Cleveland, 1961); L. T. C. Rolt, *A Short History of Machine Tools* (London and Cambridge, Mass., 1965).

Decorative fountains—for example, Versailles with its magnificent pumps and pipes—stretched the capacities of hydraulic engineers more than did plebeian water supply. Savery's fire engine was pumping water for a garden in Kensington in 1712 not long after its use in mine drainage.

The Graphic Arts

The introduction of printing illustrates the same point, though here the art is even less separable from the technology. The obvious advantage of transmitting information in written form kept thousands of

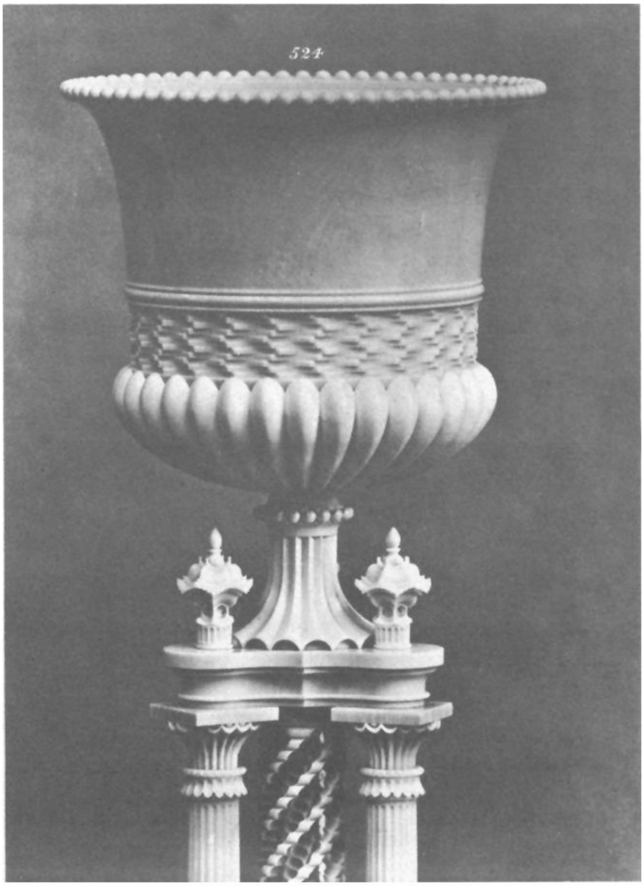


FIG. 10.—Ivory cup. One of the products of the 19th-century ornamental turning lathe, showing some of the complicated shapes made possible by the mechanical combination of simple motions. J. J. Holtzapffel, *Turning and Mechanical Manipulation* (London, 1884), vol. 5.

scribes busy for millennia, but the functional business of recording the commands of the government or the information needed by merchants did not lead to printing—this came from the desire to reproduce images and patterns. The ceramic decorative stamps at Chatal Huyuk,[25] the cylinder seals made in such profusion throughout the Middle East, the tools for the impression of decorative details in ceramic vessels and tiles as well as in molds for casting, the punches for repetitive stamping of metal, the dies for striking metal coins, and the block printing of textiles—all these precede "useful" typographic printing and lay the groundwork for it. The sequence from rubbing to woodblock print to movable type in the Far East is a direct one.[26] The first true printing was for the dissemination of a Buddhist sutra—utility and aesthetics united in the service of religion. In Europe, although the precise stages of the invention are hard to trace, the sequence is similar. The reproduction of pictures with text from woodblocks was a popular art early in the 15th century, though for the step to reusable type Gutenberg's solution involved the transfer of technique from a humbler craft, that of the pewterer, whose permanent molds with replaceable parts for decorative detail and whose alloy needed little change to make type.[27] The earliest type seems to have been cast from a tin-base alloy perhaps containing bismuth, but cheaper, harder lead alloys were common in the 16th century and thereafter.

A strong aesthetic motivation is visible in the works of the early typographers. Much of it obviously derived from the desire, or perhaps the necessity, of duplicating the quality of the manuscripts with which they were initially in competition. But art and technology are even more inextricably interwoven in the reproduction of pictures, which began before typography but received an enormous impulse from their use to illustrate printed books. Though to some extent the mere possibility of making multiple copies is the enemy of art, limited reproduction brings an artist's works to a greater audience, and the techniques themselves give rise to aesthetic qualities not otherwise obtainable. Woodcuts, etchings, lithographs—especially if the artist's hand prepares the printing surface—are often preferable to unique works executed in the traditional media of the painter.

Print-making from intaglio lines in metal plates was late in appearing,

[25] James Mellaart, *Çatal Huyuk: A Neolithic Town in Anatolia* (London and New York, 1967).

[26] T. F. Carter and L. C. Goodrich, *The Invention of Printing in China and Its Spread Westward*, 2d ed. (New York, 1955).

[27] The earliest description of typecasting is in V. Biringuccio, *De la pirotechnia* (Venice, 1540), where it appears appropriately in the chapter on the pewterer's art.

but its roots are deep. Decorative engraving on the surfaces of bone or soft stone objects, of course, precedes the use of metal, and it was widely used pictorially on three-dimensional objects of bronze, gold, and silver. The earliest date on a print from an engraved plate is 1446. Some playing cards printed about four years later have attractive animal designs that are similar to some of the marginalia in the great manuscript Bible of Mainz (dated 1452–53, now in the Library of Congress), and Lehmann-Haupt[28] has suggested that the plates may have originated in abortive experiments by an engraver working in collaboration with Gutenberg, who at the very time and in the same city was at work on his famous Bible and would naturally have liked marginal embellishment matching the best contemporary manuscripts to appear alongside his typographic text. Plausible and attractive though this hypothesis is, there is no intaglio printing that can be definitely associated with Gutenberg. In any case, for hints as to possible technical steps behind the invention, we must move to Italy, where the first engraved prints—those of Maso di Finiguerra, 1452–55—were made slightly later than in northern Europe. Sulphur casts associated with the Italian prints are preserved in both the British Museum and the Louvre.[29] Goldsmiths were accustomed to make such replicas of engraved objects, both to check the designs before filling them with niello and to provide a record for future use. It was a simple matter to make a mold (perhaps of plaster) from the engraving and to obtain an exact replica of the original intaglio lines by casting sulphur in it; smearing this with soot and oil would make the design clear and produce a general effect of black lines on a yellow background much like the final niello on gold. Transfer to paper would follow naturally and soon render the cast copies obsolete. Northern engravers may have been more ingenious: the casting of the mold material on a dirty engraving might have suggested direct transfer to paper without the need for double molding or a sulphur intermediary. In any case, fine prints could not have been made in the 14th century had not centuries of earlier work with niello developed both the technique of using the graver as well as the sense of design appropriate to it, had not the caster of art bronze had experience in the

[28] Hellmut Lehmann-Haupt, *Gutenberg and the Master of the Playing Cards* (New Haven, Conn., and London, 1966). An excellent discussion of the beginning of the graphic arts is provided by the works of A. M. Hind, especially *A Short History of Engraving and Etching* (London, 1908), and *An Introduction to a History of Woodcut . . .*, 2 vols. (London, 1935).

[29] A. M. Hind, *Nielli, Chiefly Italian of the XV Century: Plates, Sulphur Casts and Prints Preserved in the British Museum* (London, 1939). For reproductions of other sulphurs and niello prints, see also John G. Phillips, *Early Florentine Engravers and Designers* (Cambridge, Mass., 1955).

replication of models with fine detail, and had not the new oil-based inks and presses become available for the printer. Once the process of transferring to paper had been invented, it spread rapidly, and artists throughout Europe produced prints which used to the full the possibilities of rendering fine detail and controlled shading that were implicit in the technique. A few years later, shortly after 1500, engraved plates began to meet competition from those in which some or all of the lines were bitten with acid, giving them a special quality that many artists prefer to engraved ones. Here, too, an old technique was ready to be adapted to a new purpose, for armor makers had been using etching in the decoration of the more elaborate of their products for at least a century, probably much longer.[30]

The beginnings of etching—the removal of metal by localized chemical attack—are obscure. Supposedly jewellers and coppersmiths had long used vegetable acids or minerals such as copperas to remove the oxide scale produced by annealing their ware, but there is no early record of this practice. Chemical attack was certainly used in the cementation process to remove silver from solid gold at least as early as the 6th century B.C.[31] The earliest etching done with a reserved design is actually that on calcareous shells at Snaketown Pueblo in Arizona,[32] dating from the 1st century B.C. A related effect of a slightly later date is seen in the reserved areas of silver in the depletion-gilded sheet ornaments of copper-silver-gold alloy found in northern Peru.[33]

In the Old World, swords made of different kinds of iron and steel welded together into a consciously decorative pattern appear in La Têne sites, and supposedly some chemical attack would have been used to

[30] James G. Mann, "The Etched Decoration of Armour," *Proceedings of the British Academy* 28 (1942):17–44.

[31] Sidney Goldstein, in a private communication, reports that fragments of hammered gold that had unmistakably been cemented were uncovered in a 550 B.C. workshop site at Sardis. They were found during the 1968 campaign of the Fogg Museum at Sardis. The earlier Lydian coins were of unrefined electrum.

[32] Emil W. Haury, "Etched Shells," in *Excavations at Snaketown,* ed. H. Gladwin et al., Medallion Papers, vol. 25 (Tucson: University of Arizona Press), pp. 148–53.

[33] H. Lechtman, "Ancient Methods of Gilding Silver—Examples from the Old and the New Worlds," chap. 1 in *Science and Archaeology,* ed. R. H. Brill (Cambridge, Mass., to be published in Spring 1971). The presence of reserved areas of silver in "gold" objects from Peru has not previously been noted, although there have been speculations on bimetallic construction. A paper to be published in *The Application of Science in the Examination of Works of Art,* ed. W. J. Young (see n. 11 above), shows that pre-Columbian gilding was done cold, not by hot cementation. Basic ferric sulphate, a natural mineral, will remove copper and silver from a dilute gold alloy, leaving a porous layer of pure gold which is easily consolidated by gentle heat or by burnishing.

reveal the pattern. Quite apart from its decorative function, the visibility of the pattern in the welded composite would serve simultaneously to control the work in the smithy and to provide a kind of index of quality to the customer. The patterns on the swords of the Franks and Vikings (fig. 11) are referred to in Viking sagas in terms that leave

Fig. 11.—Tip of a pattern-welded iron sword. Merovingian (6th century A.D.). Width, 3.9 cm.

no doubt as to their visibility.[34] The beautiful textures of Damascus swords were also acclaimed by poets long before the technique of forging and etching them was described.[35] These were certainly etched to bring out the pattern, and etching was probably done on the European blades, though polishing alone can leave a just-visible texture on the surface if there is enough slag mixed in the metal. Japanese swords owe both their effectiveness and their beauty to the distribution of intensely hard areas left by an intricate control of the forging and heat-treating operations. These, with finely dispersed slag particles, are subtly revealed in the final polishing operation. There is, in fact, no better symbiosis of the highest aesthetic and technical standards[36] than in these swords.

Some paragraphs in Pliny may refer to chemical attack on iron, but the first clear reference to etching in European literature is in the 8th-century chemical manuscript at Lucea, *Compositiones variae*, which contains a recipe for the treatment of an iron surface with a mixture of corrosive salts containing copper as a preliminary to gilding it. A similar technique appears in the ninth-century *Mappae clavicula*.[37] With the omission of copper and the use of a stop-off to localize the effect, decorative etching was born (fig. 12). Although the earliest extant etched decoration is on late 15th-century iron armor, there is earlier evidence for its use. Conrad Kyeser's 1405 manuscript, *Bellifortis*,[38]

[34] H. R. E. Davidson, *The Sword in Anglo-Saxon England* (Oxford, 1962). On the metallurgy of the pattern-welded blades, see C. S. Smith, "The Pattern Welded Blade," in *A History of Metallurgy* (Chicago, 1960), chap. 1, and the references cited therein.

[35] C. S. Smith, "The Damascus Blade," in *A History of Metallography* (Chicago, 1960), chap. 3; C. Panseri, "Damascus Steel in Legend and in Reality," *Gladius* 4 (1965):5–66.

[36] B. W. Robinson, *The Arts of the Japanese Sword* (London, 1967); C. S. Smith, "A Metallographic Examination of Some Japanese Sword Blades," *Doc e Contributi per la Storia della Metallurgia*, no. 2 (1957), pp. 42–68.

[37] H. Hedfors, ed. and trans., *Compositiones ad tingenda musiva . . .* (Uppsala, 1932); Thomas Phillipps, "A Manuscript treatise . . . entitled Mappae Clavicula," *Archaeologia* 32 (1847):183–244; Wilhelm Ganzenmuller, "Ein unbekanntes Bruchstück der Mappae Clavicula aus dem Anfang des 9. Jahrhunderts," *Mitteilungen zur Geschichte der Medezin der Naturwissenschaft der und der Technik* 40 (1941):1–15. A translation of the *Mappae clavicula* by J. G. Hawthorne and C. S. Smith is to be published.

[38] Conrad Kyeser, *Bellifortis*, facsimile, transcript, and German translation by G. Quarg, 2 vols. (Berlin, 1967). Versions of similar recipes, without, however, using distillation to make strong acid, appear in several 15th-century sources. The first printed account of etching is in the anonymous Dutch *T. Bouch vā Wondre* (Brus-

FIG. 12.—Etched design on Italian helmet, Milan, 16th century (detail.) (Courtesy John Woodman Higgins Armory, Worcester, Mass.)

describes the preparation of distilled nitric acid for this purpose, and he even calls it *aqua martis*, in clear reference to its use on iron. It seems highly probable that the discovery of this first mineral acid about a century earlier had come directly from the experimental distillation of an etching mixture containing saltpeter and acid sulphates. Parenthetically, hydrochloric acid, distilled from a mixture of chlorides and sulphates, also appears first in connection with decorative embellishment—in a work on dyeing, *Plictho*, published in 1548—and 1589 decorative etching with it is described, but on marble, not metal.[39]

The technique of etching passed directly from arms to the production of etched iron plates for printing, which was at first a part-time activity of armorers. But, having begun as art, etching eventually began also to influence science. As the Damascus and Merovingian swords showed, etching is a sensitive means of revealing heterogeneity in steel, but metallurgists did not begin to use it consciously for this purpose until 1762. In the period between 1773 and 1786, observations on the etching of Damascus gun barrels, which were then being made in Europe, led to the first identification of carbon as the material responsible for the differences between wrought iron, steel, and cast iron.[40] The investigation of an essentially decorative phenomenon, and an oriental one at that, thus led directly to the most important single scientific discovery in metallurgical history!

Soon thereafter etching gave rise to a new decorative technique known as *moiré métallique*.[41] This was invented in 1814 and aroused considerable excitement for a few decades. (Fig. 13 shows a fire lighter

---

sels, 1513, reprinted with commentary by H. G. T. Frencken, Roermund, 1934). Next is the important little pamphlet *Von Stahel und Eysen* (Nuremberg, 1532), which was reprinted many times both by itself and with other material in the series of *Kunstbüchlein* and other books of secrets. For a modern English translation, see C. S. Smith, *Sources for the History of the Science of Steel 1532–1786* (Cambridge, Mass., 1968), pp. 37–38.

[39] L. Reti, "How Old Is Hydrochloric Acid?" *Chymia* 10 (1965):11–23; Sidney Edelstein and H. C. Borghetty, eds., *The Plictho of Gioanventura Rosetti* (Cambridge, Mass., 1969); G. B. della Porta, *Magiae naturalis libri viginti* (Naples, 1589). The 1658 anonymous English translation of the section "How to Grave Porphyr Marble without an Iron Tool" is reproduced in C. S. Smith, *Sources for the History of the Science of Steel*, pp. 37–38. See also Haury (n. 32 above).

[40] C. S. Smith, "The Discovery of Carbon in Steel," *Technology and Culture* 5 (1964):149–75.

[41] C. S. Smith, "Moiré métallique," in *A History of Metallography* (Chicago, 1960), pp. 63–65. A later description of the process in an American hardware catalog is quoted by H. J. Kaufman, *Early American Ironware, Cast and Wrought* (Rutland, Vt., 1966).

made by this technique.) It was simply etched and lacquered tin plate, but the plate was sometimes treated by local heating and cooling to give very fancy crystallization patterns, even semblances of flowers and landscapes!

New methods of printing illustrated books repeatedly redounded to

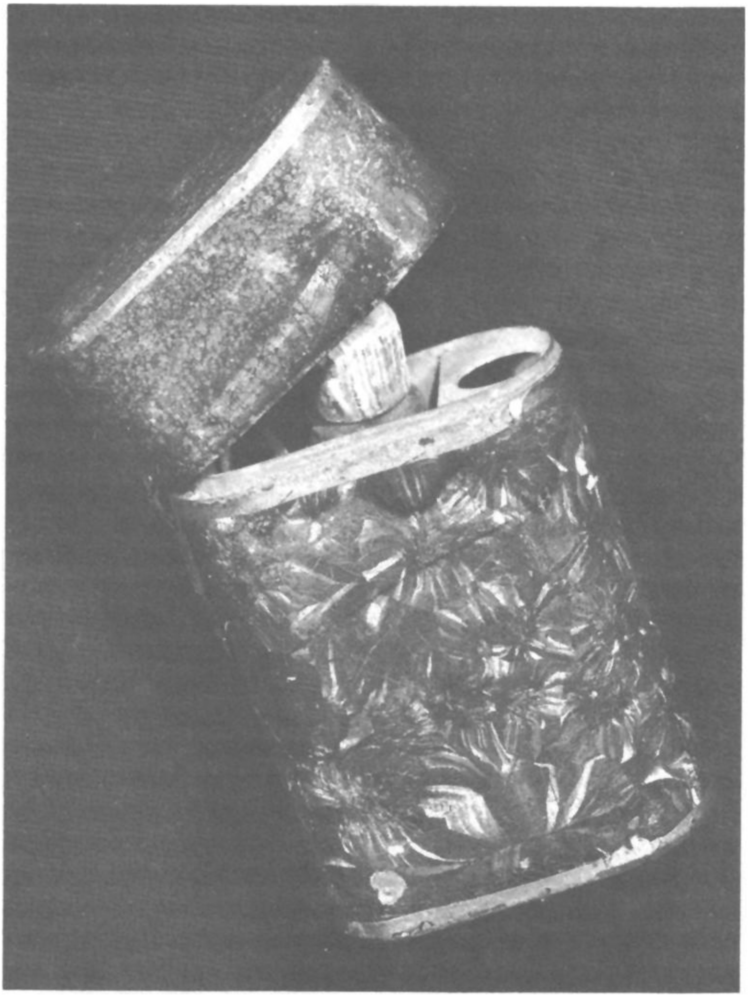


FIG. 13.—"Instantaneous light box" with case made of green *moiré métallique*. Made in London about 1820 by "J. Watts and Co., Chymists No. 478 Strand." Height, 8.0 cm. This device made fire by bringing a wooden match tipped with potassium chlorate and sulphur into contact with concentrated sulphuric acid. *Moiré métallique* was tin-plated iron that was given a special treatment to develop a fancy crystallization, subsequently etched and covered with colored lacquer. (Photo courtesy Bryant and May Ltd. and Science Museum, London.)

the advantage of both science and technology. An interesting printing technique—first described in the 16th century[42]—was to make direct impressions of objects such as leaves by coating them with printer's ink and impressing them directly on paper. The process (which is not unrelated to the much earlier and more versatile oriental method of producing rubbings on paper laid over objects with details in relief) was later called "nature self-printing." In the 18th-century a number of botanical books were published with illustrations printed this way, the first being J. H. Kniphof's *Botanica in originali*, published in 1733.[43] The same technique was used by von Schreibers and Widmanstätten in 1813 for recording the etched structure of a section of the Elbogen meteorite. Their print was a spectacular improvement in clarity and accuracy over the lithographs of other meteorites that accompanied it in their published book[44] or the engravings by Gillet de Laumont in the *Annales des Mines* of 1815. For a time thereafter many methods of obtaining relief or intaglio impressions of an object directly on a printing surface were experimented with for both scientific and other purposes.[45] Nature printing from a collage of textured surfaces is the basis of a flourishing school of printmakers today.

The early history of photography itself is a classic example of the symbiosis of art and invention. Della Porta in 1558 recommends the *camera obscura* as a device to lighten artists' labors and help them with perspective. Niepce's famed photochemical etchings on glass (1826) were done to reproduce art, not reality. The processes of Daguerre and Talbot were of both worlds, as photography has been ever since. When the invention of photomechanical methods displaced most other methods in the printer's shop, etching had become a common laboratory technique. The science of metallography—indeed, practically the whole structural side of modern materials science—stems from the work of Henry Clifton Sorby in 1863–64 in the famous steelmaking center of

[42] Alexis [pseud.], *Secreti . . . del Alessio Piemontese* (Venice, 1555). There were innumerable subsequent editions and translations of this book, which is the most complete of all the early books of recipes for artists, craftsmen, and housewives.

[43] For a history of nature printing, see Ernst Fischer, "Zweihundert Jahre Naturselbstdruck," *Gutenberg Jahrbuch* (1933), pp. 186–213.

[44] Carl von Schreibers, *Beyträge zur Geschichte und Kenntniss meteorischer Stein- und Metall-massen* (Vienna, 1820). See also Smith, *A History of Metallography,* pp. 150–56; and "Note on the History of the Widmanstätten Structure," *Geochimica et Cosmochimica Acta* 26 (1962):271–72.

[45] Alois von Auer, *Der polygraphische Apparat* (Vienna, 1853). This includes a portfolio of fine prints made by all methods of reproducing illustrations then known: several are of scientific subjects.

Sheffield, which was the world center of supply for engravers' steel plates. By applying to the preparation of laboratory specimens the methods used for giving these plates their fine finish, and by using etching, which he had heard discussed at a meeting of the local Literary and Philosophical Society, Sorby was able to reveal for the first time in history the true microstructure of steel without disfigurement by fracture or deformation.[46] In the present connection it is interesting to note that the next paper on metallography, by the German railroad engineer Adolf Martens in 1878, was directly inspired by some work on the quality control of metal for use in the exquisite art castings of iron for which Germany was rightly famous at the time.

Electroplating and Electrical Engineering

Electrochemistry is another area in which the interest of the artist or the art industry accelerated scientific knowledge and technological development. An old and pretty parlor trick was the *Arbor Dianae,* mentioned with other "metallic vegetations" in most chemical textbooks of the 17th and 18th centuries. Eighteenth-century assayers knew of the electrochemical series (though they did not call it that) in the form of sequential replacement of silver in solution by copper, copper by iron, and iron by zinc.

The medieval use of an acid cupriferous solution to give a coating of copper on iron was mentioned above. Such electrolytic replacement remained a common observation and was sometimes used for recovering copper from waste mine waters as well as to confuse people with the semblance of transmutation. It gave rise to a minor art in the 17th century in the form of a very pleasant ware made from cement copper in the town of Herrengrund in the Bohemian Erzgebirge.[47] These objects bear inscriptions reflecting their polymetallic origin (as in fig. 14) or cryptic jingles such as on a wine cup in my possession, which reveals a common source of scrap iron for the process:

> Ein Pferd mich vor mit füssen trat,
> da ich noch Eisen ware,
> durch ziment wassers baad
> bring ich gut freünd zu baare.

[46] Norman Higham, *A Very Scientific Gentleman: The Major Achievements of Henry Clifton Sorby* (Oxford, 1963). For Sorby's work on steel, see Smith, *A History of Metallography*, pp. 169–85; A. R. Entwisle, "An Account of Exhibits Relating to Henry Clifton Sorby," *Metallography 1963* (London, 1963); and papers by C. S. Smith, D. W. Humphries, and Norman Higham in *The Sorby Centennial Symposium on the History of Metallurgy* (New York, 1965).

[47] Gustav Alexander, *Herrengrunder Kupfergefässe* (Vienna, 1927).

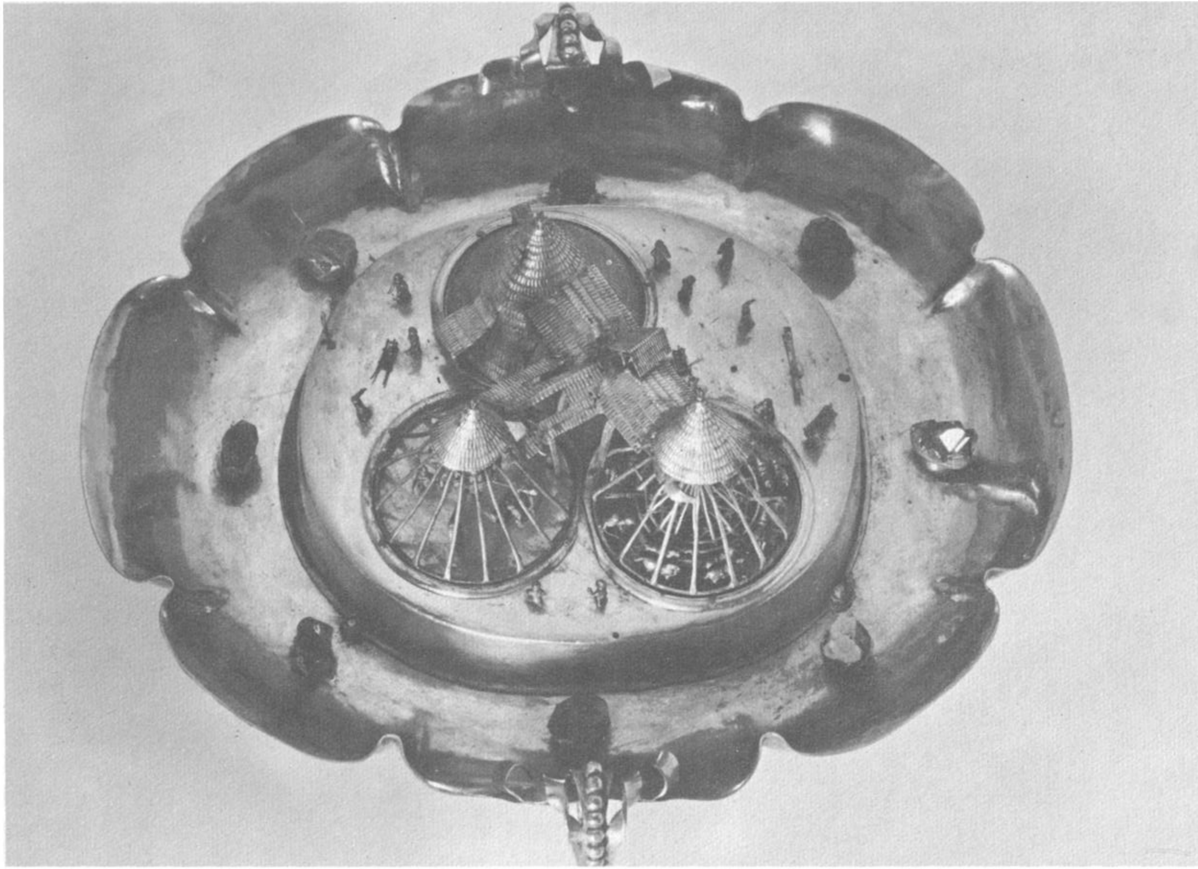


Fig. 14.–Copper dish (Herrengrund ware) containing model of minehead equipment with working miners and mineral specimens. Length, 33 cm. Heavily gilded. The inscription, "Eisen war ich, Kupfer bin ich, Silber trag ich, Gold bedecket mich," refers to the recovery of the copper from mine waters by displacement with scrap iron. Made in Herrengrund, Bohemia, early 18th century. (Photo courtesy Abegg-Stiftung, Bern.)

It has been reported that this ware was shaped in iron and then plated by immersion in cement water. Some folk objects were certainly so made, but the real Herrengrund ware is actually nonmagnetic and was probably made from cement copper powder that was melted, refined, cast, hammered into sheets, and shaped as any other copper would have been. The role of electricity in these operations was not, of course, suspected, anymore than it was in the mysterious decay of rudder irons on the English ship Phoenix that in 1670 had been sheathed with sheet lead, which had just then become available in wide sheets from the new rolling mill at Deptford.[48]

If any of these effects had been looked at by a sufficiently curious mind, Galvani's discovery could easily have been made a century or more earlier and without the intervention of a frog. However, even after Galvani and Volta, even after Wollaston's and Cruickshank's demonstrations of the cathodic deposition of copper and other metals, and even after Michael Faraday's elucidation of the laws of electrolysis, no use was made of the phenomenon until 1838. In that year Jacobi, Spencer, and Palmer, in somewhat confused priority, all began the art of electrodeposition for the duplication of coins and other small art objects[49] as well as for the reproduction of printing surfaces—at first for illustrations (fig. 15) and later for letterpress.

Henry Bessemer later claimed that in 1832, when a young man of nineteen, he had reproduced plaques by electrolysis, but he did not publish. After 1838 this quickly became a very popular hobby and resulted in widespread knowledge of electricity. Smee, writing in 1842, remarked that "there is not a town in England that I have happened to visit, and scarcely a street of this metropolis [London] where prepared plasters are not exposed to view for the purpose of alluring persons to follow the delightful recreation afforded by the practice of electro-metallurgy." The new metallurgy quickly spread from copper to other metals. Gold was naturally one of the first, but the most commercially significant was the electrodeposition of silver upon the beautiful white copper alloy now known as nickel silver which rapidly displaced the more expensive Sheffield plate. The base alloy itself had been imported from China for about two centuries and was used for fireplace equip-

[48] Thomas Hale, *An Account of Several New Inventions of Improvements Now Necessary for England* . . . (London, 1691).

[49] The early books on electroplating devote much space to the advocacy of one or another view of its history. See particularly Alfred Smee, *Elements of Electro-metallurgy* (London, 1842; 2d ed., 1843); George Gore, *The Art of Electrometallurgy* (London, 1877). The process of electrotyping, commercially used in 1840, spread rapidly aross the Atlantic (see fig. 12).

ment, candlesticks, and other domestic objects of some elegance,[50] but it was not analyzed until 1776 (twenty-five years after the discovery of nickel itself), and it was not available commercially until 1833, just in time for its wedding with the new plating method.

Electroplating soon became an important art industry. It provided economic support for the beginning of the nickel industry, it provided a large market for electromechanical generators before the development of electric light, and, with telegraphy, it provided training and

FIG. 15.—Print from one of the earliest American electrotype plates. This formed the frontispiece of Daniel Davis, Jr., *Manual of Magnetism* (Boston, 1842), alongside a print from the original engraved plate, indistinguishable from the copy.

experience for innumerable men who were soon to combine their empirical knowledge with a growing science to give birth to electrical engineering.

Art for the Masses

The story of electroplating is only one example among many in which a desire to simulate a precious material in a cheaper form has stimulated technical advance. All students of the historical literature on iron and steel know Réaumur's classic *L'art de convertir le fer forgé en acier et l'art d'adoucir le fer fondu* (Paris, 1722), with its curious subtitle *faire des ouvrages de fer fondu aussi finis que de fer forgé.*

[50] Alfred Bonnin, *Tutenag and Paktong* (Oxford, 1924).

Réaumur specifically states that his incentive was to provide cheaply for the masses decorative objectives of the kind expensively made and previously available only to the rich. It was not engineering devices but elaborately chiselled door-knockers (fig. 16) at which he first aimed—but he ended by extolling the virtues of mass production of interchangeable parts in industry generally.[51] The motive of cheap art was also behind Réaumur's development of his "porcelain," a devitrified glass of the type that has recently been revived in superior form.[52]

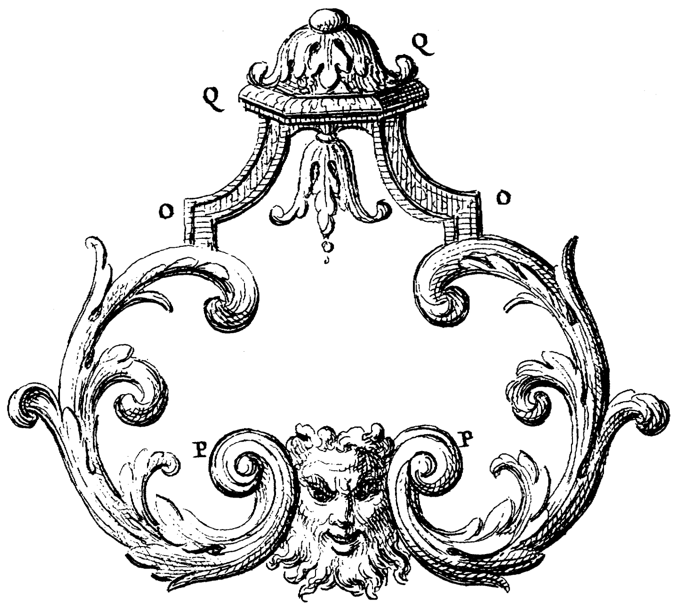


FIG. 16.—A door knocker—the first use of malleable cast iron. From R. A. F. de Réaumur, *L'art de convertir le fer forgé en acier* . . . (Paris, 1722), plate 16. This figure is the cover design of this issue of *Technology and Culture.*

He was also the first to suggest the use of wood pulp in the making of paper.

Today's steel rails, I-beams, and other structural shapes also have their origin in decorative needs—the rolling of H-shaped lead cames for stained glass windows. Around 1750, fancily profiled sections of iron for use in

[51] R. A. F. de Réaumur, *L'art de convertir le fer forgé en acier* . . . (Paris, 1722); trans. A. G. Sisco (Chicago, 1956), pp. 340–59.

[52] R. A. F. de Réaumur, "L'art de faire une nouvelle sorte de porcelaine, . . . où de transformer le verre en porcelaine," *Mémoires de l'Académie des Sciences,* 1739 (published 1741), pp. 370–88. Réaumur's porcelain is extremely rare, but a box that seems to be made of it has recently been described by R. Strasser, *Journal of Glass Studies* 9 (1967):118. See also n. 16 above.

balcony railings, window moldings, and the like were being made in grooved rolls, three decades before Henry Cort applied the process to the large-scale consolidation of wrought iron bars.[53] Another even more portentous mechanical invention was Jacquard's loom (1801) with its punched-card control: this was not needed for plain fustian but for the fanciest of lace.

Parcel gilding can be justified on purely aesthetic grounds as producing an agreeable contrast in color, as in inlay, but most gilding operations have been done simply to save money and to make expensive-looking objects available to others than the rich. The preparation of thin gold leaf, the most extreme utilization of the malleability of any metal, is similarly inspired. The fact that composition gradients could be produced in solid metals was made quite clear, long before diffusion became a subject of scientific inquiry, by the common use of gilding via gold amalgam, as in Europe, or by chemical methods of surface enrichment, as in pre-Columbian South America and in Japan.

The many changes of properties and surface coloration of metals produced by goldsmiths could hardly have failed to support the belief that transmutation is possible—as indeed it is, if "transmutation" is not limited to nuclear changes but is applied to major changes of physical properties.[54] Hopkins,[55] in particular, has argued that alchemy was an outgrowth of the joining of Greek philosophy with a knowledge of workshop practices. Yet the value of empirical knowledge naturally fades as a field advances, and the replacement of alchemy by modern chemical theory is attributable more to the logical than to the practical approach.

The above examples show that the art industries have contributed greatly to the development of techniques and to the knowledge of reactions on which today's science and technology are based. Perhaps, indeed, the mixture of aesthetic and commercial motivation involved in such developments was quantitatively the most powerful stimulus of all, for basic discovery of new effects inspired only by curiosity is

[53] The lead rolling mill is illustrated in Jost Amman's woodcut of the glazier (Jost Amman, illustrator, *Eigentliche Beschreibung aller Stände auff Erden . . . durch . . . Hans Sachs* [Frankfurt, 1568]). Plates showing the profiled iron shapes and some new window designs that they made feasible appear in [Bullot] *Mémoire sur les ouvrages en fer et en acier qui se fabriquent dans la manufacture Royale d'Essonne par le moyen du laminage* (Paris, 1753). For more detail, see Cyril Stanley Smith, "Rolled Architectural Iron, 1753," to be published in a forthcoming issue of *Technology and Culture.*

C. S. Smith, "Matter versus Materials: A Historical View," *Science* 162 (1968): 637–44.

[55] A. J. Hopkins, *Alchemy: Child of Greek Philosophy* (New York, 1934).

by its very nature rare, as rare, perhaps, as any great individual work of art.

Sources in Art for the History of Technology

Little reliance can be placed on any of the written sources relating to technological history prior to about A.D. 1500 unless they are confirmed by contemporary nonverbal evidence. Even today technologists are not noted for literacy, and men like the Benedictine monk Theophilus (early 12th century), whose hands were accustomed to both the hammer and the pen, have always been rare. In books ideas naturally fare better than technology. Moreover, the chances of survival of written technological information in medieval libraries were not high. For all this, there are many records that must be studied in the absence of anything better, and in these a strong bias toward the decorative arts is evident. Following the Roman Vitruvius's *De architectura*, the best pre-Renaissance technological sources are the Leyden papyri, the Lucca manuscript entitled *Compositiones variae*, the *Mappae clavicula*, Theophilus's *De diversis artibus*, and Eraclius's *De coloribus et artibus romanorum*. Every one of these deals with the artist's materials and techniques, to the exclusion of almost everything else. Manuscripts presenting primary information on machinery for warfare, mining, and other industrial occupations do not exist before the 14th century. The famed Theophilus's manuscript of about A.D. 1123 is an outstanding source of pure technology, though he confines himself to giving intimate details on painting, stained glass, and metal work for the embellishment of the church.

Although it was far from the artist's conscious intent, many paintings of religious subjects, especially those of the 13th–16th centuries, convey information on agricultural and building techniques, and they are particularly important in reflecting current attitudes toward labor and machinery. Lynn White, jr.,[56] has studied from this viewpoint the changing depiction of the seven Virtues. As late as the 12th century Temperance was in little esteem, but during the 13th century she became identified with measure and subtly associated with internal and external control. By 1350 she is depicted with the newly invented sandglass; by 1450 (in a manuscript now in Rouen) all the seven Virtues are depicted with technological appurtenances, but Temperance displays eyeglasses, rowell spurs, a mechanical clock, and a tower windmill. The showing in such a scene of these objects—all recent inventions—expresses "a reverence for advancing technology, a sense of its spiritual

[56] Lynn White, jr., "The Iconography of Temperantia and the Virtuousness of Technology," in *Action and Conviction in Early Modern Europe*, ed. T. K. Rabb and J. E. Seigel (Princeton, N.J., 1969), pp. 197–219.

value, which is peculiar to the West and which has been essential for the building of industrial society."

There are innumerable illustrations known to historians of art but almost untapped by their technical confreres in which an artist interested in human activity (either for its own sake or to satisfy an ecclesiastical or princely command) used the decorative aspects of tools and mechanical devices, quite commonly with disregard of mechanical details but nevertheless providing useful information to the historian.[57] And, of course, the very materials of the artist are themselves a superb record of the technology that produced them, a record that can be read in intimate detail by modern laboratory techniques. The output of the spectroscope, microscope, and X-ray spectrometer will soon become as important to the technological historian as his older verbal sources, to which we now return.

The two earliest printed works on the prosaic subject of steel both have artistic overtones. The first was a little pamphlet on etching, the *Stahel und Eysen* (1532) mentioned above (n. 38), and the second was on ornamental ironwork and locks—Mathurin Jousse, *La fidelle ouverture de l'art de serrurier*, published at La Flèche in 1627 (fig. 17). By this time the artist was aiding the technologist in substantial ways, for the new techniques of producing more accurate representation of visual appearances served increasingly to convey precise technical information. The accurate, detailed drawings of the liquation process for the desilverization of copper, and those of lathe details, pile drivers, etc., contained in the 15th-century "Hausbuch" of the Mendel brothers[58] and in the Nuremberg "Hausbuch,"[59] are a far cry from earlier illustrations in which technology is only incidentally reflected. One of the leading German illustrators of the 16th century, Jost Amman, sought inspiration in the technical crafts for eighty-six of the 118 woodcuts in his popular Book of Trades.[60]

[57] Three useful collections of paintings and other works illustrating technological scenes and devices are Heinrich Winkelmann, *Dar Bergbau in der Kunst* (Essen, 1958); Vaclav Husa et al., *Traditional Crafts and Skills* (Prague & London, 1967); F. D. Klingender, *Art and the Industrial Revolution* (London, 1947). Many works on the history of technology reproduce artists' works as illustrative material. Emil E. Ploss, *Ein Buch von alten Farben* (Heidelberg and Berlin, 1962), as befits its subject, is an unusually fine mixture of historical and artistic material.

[58] Wilhelm Treue et al., eds., *Das Hausbuch der mendelschen Zwölfbrüderstiftung zu Nurnberg . . .*, 2 vols. (Munich, 1965).

[59] Helmuth T. Bossert and Willy F. Storck, eds., *Das mittelalterliche Hausbuch* (Leipzig, 1912).

[60] Amman (n. 53 above). The same blocks were used in printing a Latin edition in the same year.

By the mid-16th century many carefully written books on both science and technology were being printed with woodcut illustrations (fig. 18). Both the biological sciences and technology required and inspired some of the best efforts of the artist in rendering realistic details without confusion. The woodcuts in the well-known treatises of Vesalius, Agricola, and Ercker are about as attractive as any book

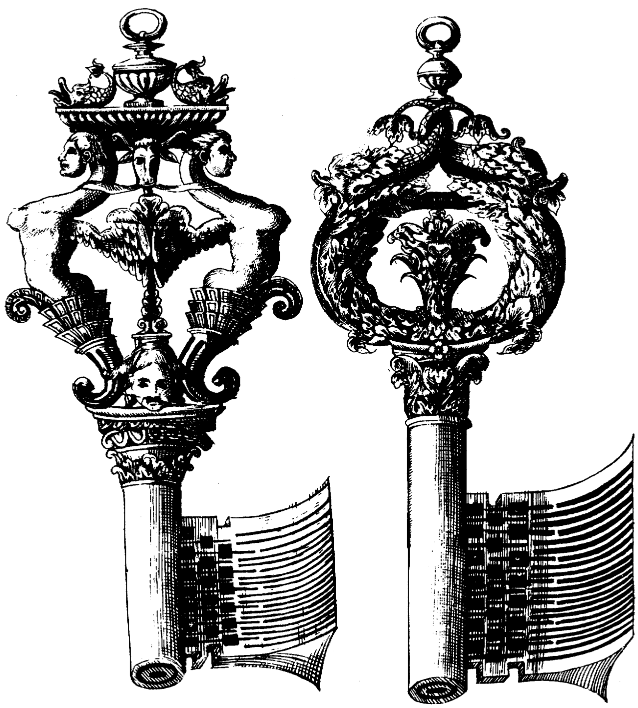


FIG. 17.—Design for a key of forged and chiseled iron. From M. Jousse, *La fidelle ouverture de l'art de serrurier* (1627), plate 1.

illustrations of the 16th century, and their instructional value was correspondingly high. The mystical side of alchemy, though scientifically sterile, appealed to the artist's imaginative approach and gave rise to some attractive books.[61]

In general, physics, with its abstract concepts and simple diagrams, neither attracted nor needed the artist. If a physicist used illustrations at all, they were likely to be in the form of colorless linear diagrams making visible the geometry implied by his equations. Galileo in 1638

[61] John Read, *Prelude to Chemistry* (London, 1936); Read, *Humour and Humanism in Chemistry* (London, 1947).

depicts a weed-encrusted stone wall supporting his elastically-deflected beam (fig. 19), but later elasticians eschewed such realism. The terrellae in William Gilbert's *De magnete* (London 1600) have a pleasant look, perhaps contributed by the man who cut the block, and to illustrate his observations on the magnetization of cooling iron, Gilbert allowed himself the luxury of including a woodcut view of a blacksmith's shop that is in the direct tradition of the series of such views

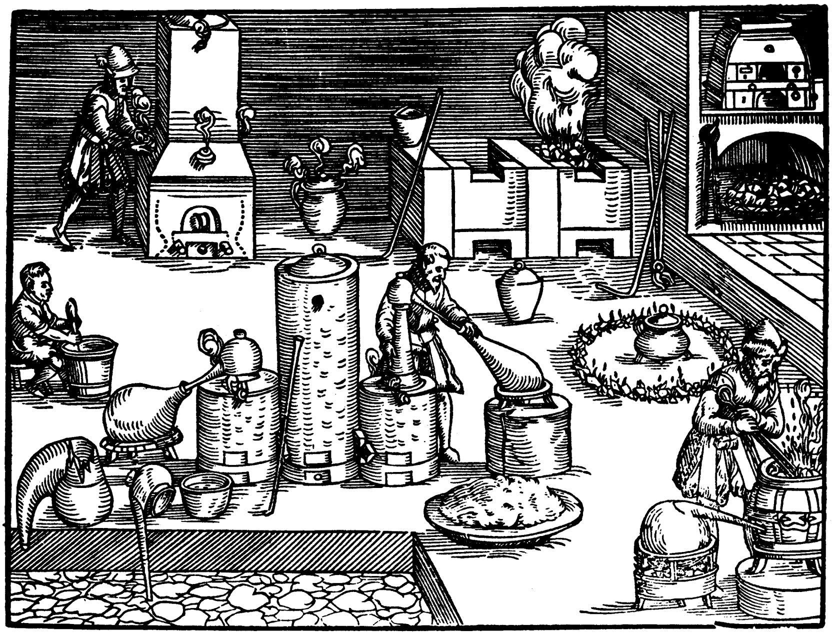


FIG. 18.—Woodcut view of an assay laboratory. From Lazarus Ercker, *Beschreibung allerfurnemisten mineralischen Ertzt und Berckwercksarten* (Prague, 1574).

in the medieval *Speculum humanae salvationis*, where they illustrate (amid changing hearth and anvil design and with occasional detachability of the horse's leg to simplify the smith's work) metallurgy's first contribution to the fine arts—Tubal Cain's rhythmic clangor giving rise to the idea of melody in a listener's mind. Gilbert's other illustrations, however, are purely linear diagrams.

The engravings in Robert Hooke's *Micrographia* (1665) reflect both approaches. Most of these are well-shaded sensitive representations of exciting vistas in the New Landscape that his microscope was exposing for the first time (fig. 20), but Hooke's diagrams of the paths of rays of light (fig. 21) have a sharp austerity which matches the abstraction of the idea and which came to characterize most scientific diagrams thereafter.

The engraved copper plates preferred by 17th- and 18th-century book publishers over wood blocks permitted accurate delineation of apparatus and were excellent for showing machinery (fig. 22). The enormous growth of the graphic arts in 18th-century France coincided

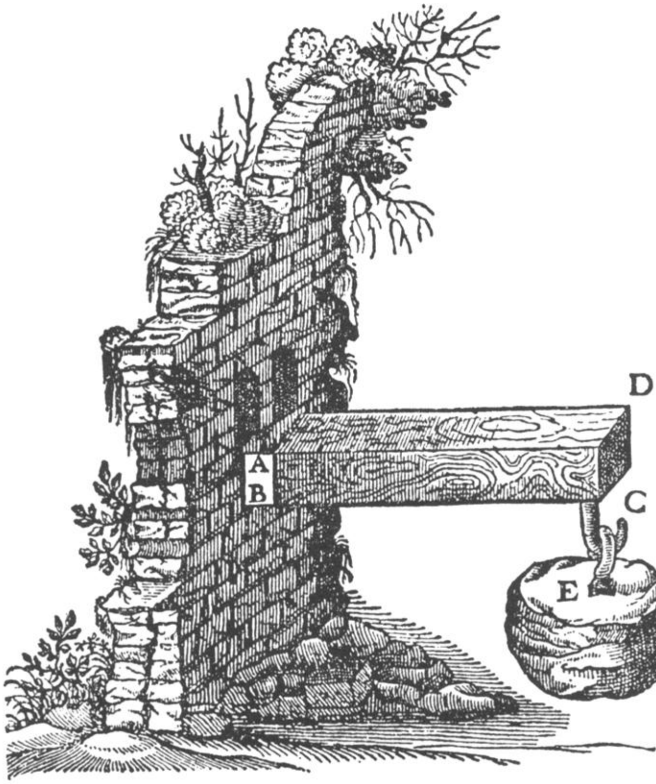


FIG. 19.—Woodcut diagram to illustrate the bending stresses in a beam. Galileo, *Discorci e dimostrazioni matematichia intorno a due nuove scienze* (Leiden, 1638).

with a rationalist viewpoint to result in the publication of massive collections of engravings of technical subjects. The hundreds of folio-size plates in the series of *Descriptions des arts et métiers* published by the French Academy of Sciences (fig. 23) and the seven volumes of plates accompanying Diderot's famed *Encyclopédie* provide a profuse record of technical crafts and industry. Our knowledge of the technology of that time is probably more complete than that of any other period in history, for before this there was scant interest in making records and after it the profusion of technology both outran the possibility of fully recording it and stifled an interest in the details of its minor variations.

On the Segregation of Disciplines

The conscious separation and classification of an activity or viewpoint as science, technology, or art is recent and came about rather slowly. It is misleading to apply modern classifications to earlier periods in which distinguishable professions did not exist and a desired end result dominated over conscious particularities of method. Nevertheless, it is obvious from the above that I regard the somewhat less fully intellec-

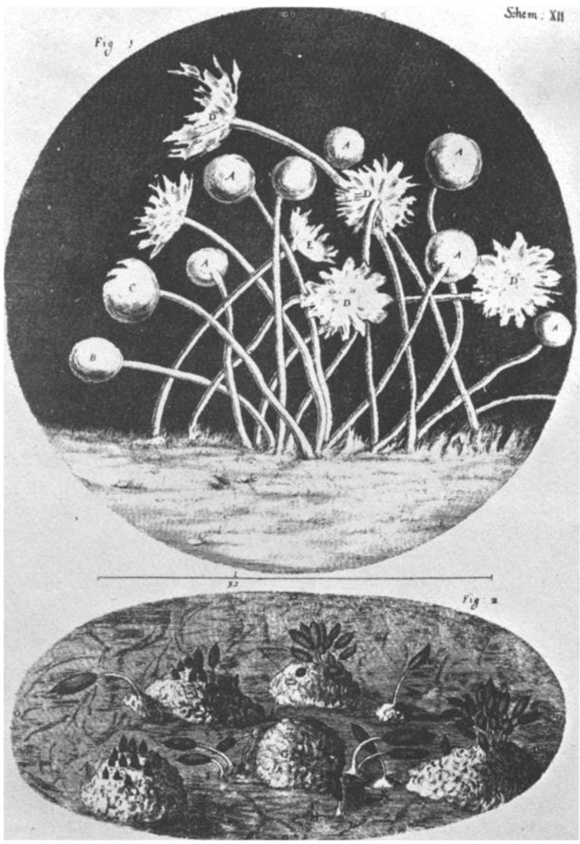


FIG. 20.—Engraving showing "nothing else but the appearance of a small white spot of hairy mould, multitudes of which I found to bespeck and whiten over the red covers of a small book" (Robert Hooke, *Micrographia* [1665], plate 12). The scale line is 1/32 inch, corresponding to an original magnification of about fifty.

tualized activities of the technologist as having much in common with those of the artist and, until recently, interacting rather less with those of the scientist.

The Renaissance marks a natural interaction between a rejuvenated art and a beginning science.[62] In the 14th century many artists delighted in using their newly awakened powers of observation and their increased skill in representation to embellish the margins of manuscripts

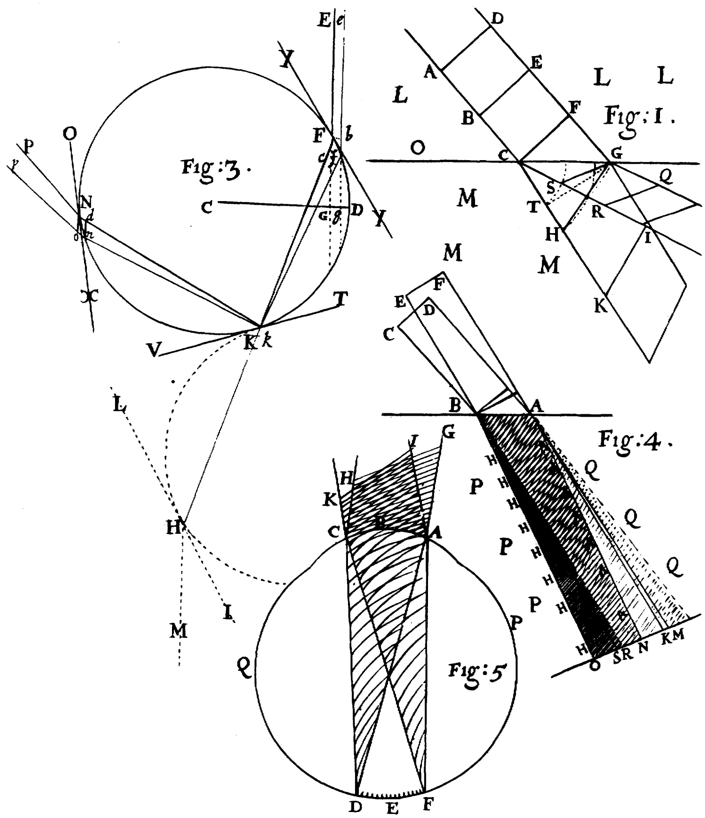


FIG. 21.—Engraved diagrams showing paths of rays of light in the eye and in other media. R. Hooke, *Micrographia* (1665), plate 6.

[62] Giorgio de Santillana, "The Role of Art in the Scientific Renaissance," in *Critical Problems in the History of Science*, ed. Marshall Claggett (Madison, Wis., 1959); reprinted with other essays in *Reflections on Men and Ideas* (Cambridge, Mass., 1968). See also the essays in H. H. Rhys, ed., *Seventeenth Century Science and Arts* (Princeton, N.J., 1961).

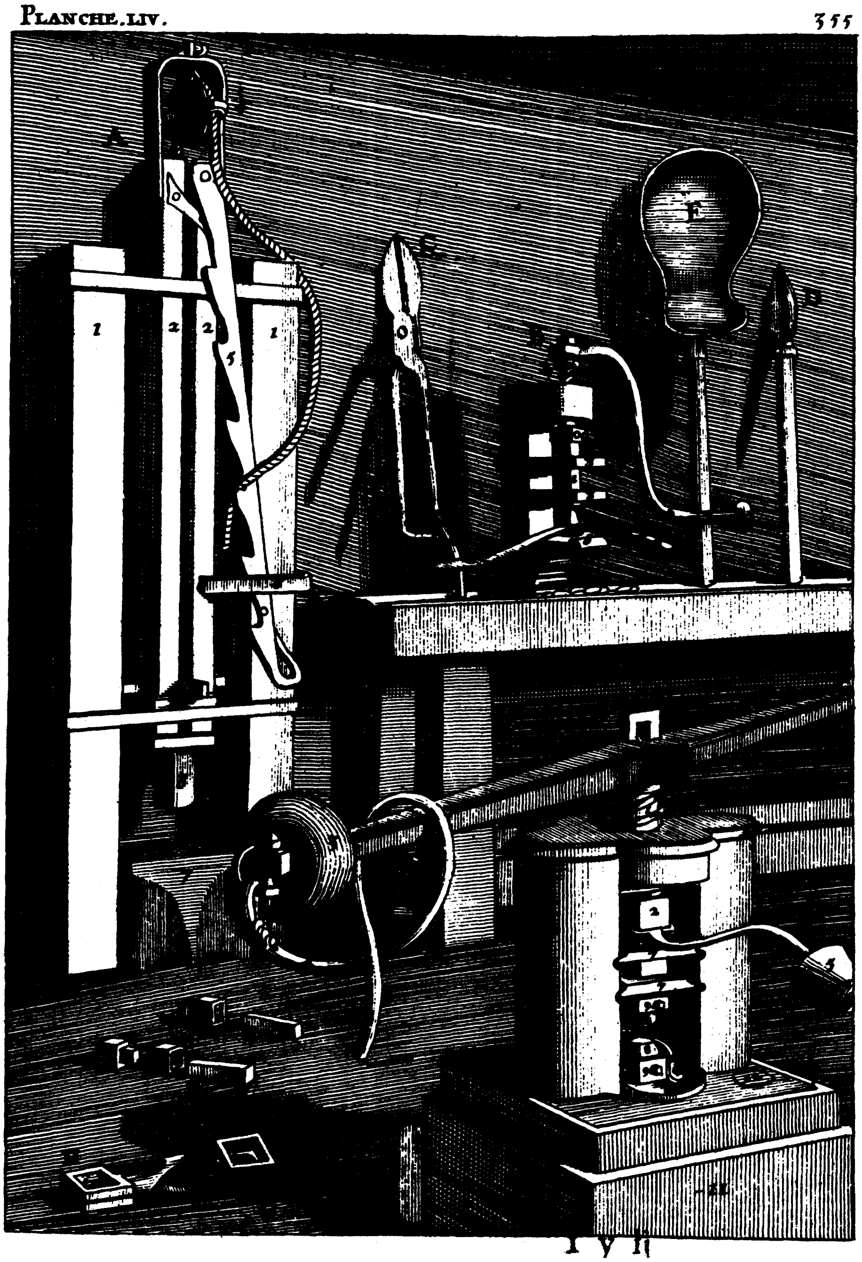


FIG. 22.—Engraving showing machinery for blanking and striking coins. André Félibien, *Principes de l'Architecture* (Paris, 1676).

with precisely limned naturalistic living forms,[63] while at the same time they came to observe and emphasize the essential aspects of their subjects in a manner that later became appropriate for scientific illustration. Conscious studies of the interaction of light with matter and almost mathematical considerations of perspective are reflected in the mid-16th-century paintings and sculpture by Ghiberti, Brunelleschi, and others, who both set the tone of the new times and absorbed its spirit. Fifteenth-century writing about art is very different in tone and

FIG. 23.—Interior view of workshop producing hammered copper vessels. From Duhamel du Monceau, "Déscription de la manufacture du cuivre de M. Raffaneau établie près d'Essone," appended to —. Galon, *L'art de convertir le cuivre rouge . . . en lation* . . . ([Paris], 1764).

intent from the earlier collections of pigment recipes or the practical how-to-do-it treatise of Theophilus. Stillman Drake[64] shows how the conflict of theory and experiment in 16th-century music contributed directly to the development of the style of Galileo and other great 17th-century scientists. Yet, as the different viewpoints that had been combined in the artists' activities came to be consciously realized, an inevitable result was that each of them should become a separate field of specialization.

[63] G. Evelyn Hutchinson, "Psychological and Aesthetic Factors in the Progress toward Realism, A.D. 1280–1480" (paper presented at Symposium on Art and Science, AAAS meeting, December 1968).

[64] Stillman Drake, "Renaissance Music and Experimental Science," *Journal of the History of Ideas* (October 1970).

The artist became an important individual, highly visible in society, while craftsmanship took on a markedly lower status. Logical thought had always aided the artist in making his materials conform to his vision, but when the critical interplay between logic and experiment was consciously separated as a method of learning about the world, it became the new science, growing and changing beyond all recognition of its origins. For four centuries now it has outrun the other aspects of the artists' approach and has done so by exploiting the power of partial isolation. If mathematics could deal with music and perspective, it could also deal with falling bodies—but it could handle the planets better than a terrestrial feather, for it only applies to ideal isolated systems of simplified forces and bodies, one or two or at most a very few things at a time. Science in its very essence is simple. The new physics could deal with ideally elastic bodies, but it could do nothing with plasticity or with the host of other structure-sensitive properties on which the arts depend.

The geometry of perspective could be well handled by mathematicians, but the perspective of color could not be. The artist's intuitive knowledge of the psychology of perception has interacted strongly with science in the 20th century, but for the most part science developed without art, and art was affected by science only through the changing world view that science promulgated or indirectly through the effect of science on technology.

The experience at moments of insight must be much the same among creative men of all kinds. However, the communication of new ideas, and especially their validation in terms that others will accept, is vastly different in different fields. As science and technology have become simultaneously broader in scope and more precise in individual purpose, their connection with art has become less and less apparent. Despite the austere and magnificent beauty of the order that is being uncovered by science, art has remained closer to technology than it has to science. As science has discovered the strength of simplicity, technology has become more complex. There is even a kind of aesthetic quality displayed by the interdependent relationships between the parts of an intricate machine, a complex process, or a large organization. Order per se is not art, and neither is complexity, but the finding of order in complexity is.

Looking back from the 20th century, it is obvious that engineers, if not exactly aesthetes, have always had a rich and valid aesthetic experience in building their structures and devising their machines. A Newcomen engine at work with its massive rocking beam of oak mounted in a simple stone structure, with clanking chains and resonant iron bars, with its fiery furnace and jets of steam and its slow and irregular oscillation, *was* a work of art even if it was not consciously built as one or so appreciated in its own time. A modern artist, Garry Rieveschl, has recently proposed building a full-scale working Newcomen engine as a public monument, providing at once a reconstruction of a forgotten experience, a glance at a critical moment in technological history, and a reminder of the beauty and portentous quality of new contrivances.

In the theater, documentary drama is a similar art form based on a selective reconstruction of the past. Less conventional is the work of another Boston artist, Harris Barron, who has devised a performance which by effectively and unforgettably evoking the emotional experience of early aviators exemplifies the way in which art can extend human experience. Perhaps a poet on the first lunar landing would have done more for technology than an astronaut; certainly it will be poetic interpretations of space travel that will remain most in men's minds.

Symbolism in Art and Science

Both art and science are basically symbol-making activities, and both have the quality of yielding metaphors that match far more than their creators intended. The scientist's equations and the conceptual models on which they are based often relate to other parts of nature which are mathematically similar but physically unrelated. This relationship is matched by that fundamental evocative quality of art, in which relationships developed by the artist with one aspect of form in mind turn out to suggest many other things to the eye of a viewer who has had different experiences. The artist consciously exploits the similarities in shape, color, texture, orientation, or other qualities of things of quite different natures; in fact, if there were not some such resonance, the viewer of a picture would find little to hold his attention. The scientist finds that a few basic patterns reappear at different levels and in different systems, but this is mainly because the types of interaction between the few units with which alone he can deal are, after all, quite limited: simplicity and symmetry do not allow many alternative arrangements.

Historically, it is interesting to note occasions on which the decorative artist has developed designs that later were reinvented to represent important scientific concepts. One of the best examples of this is the use of circular mosaic tiles to build up two-dimensional polygonal patterns having all the characteristics of order, symmetry, and angular relationship between planes that are the basis of crystallography. This occurs in the Sumerian palace at Uruk, built in the middle of the 4th millennium B.C. (fig. 24). Mosaicists ever since have been displaying examples of the combinatorial possibilities of simple geometric forms, none more magnificently than the Islamic tile workers of the 15th–17th centuries A.D.[65] The crystal lattice dislocation, which was conceived in 1926 and has become extremely important in solid-state physics, was modeled much earlier in the fitting of medieval suits of mail armor, in the studded decoration of Japanese cast-iron tea kettles, and with slight distortion in innumerable other repetitive designs.

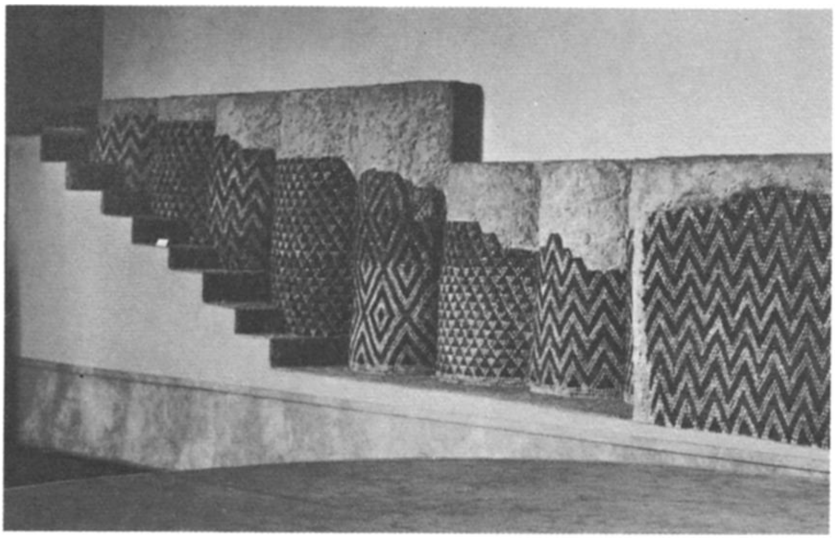


FIG. 24.—Inlaid mosaic decoration on columns at the palace at Uruk, ca. 3500 B.C. (National Museum, Berlin. Photo courtesy Bildarchiv Foto-Marburg.)

The best three-dimensional models of the close-packed-sphere arrangement of atoms on a crystal lattice occur in the famed granulation work of Etruscan goldsmiths in the 6th century B.C., though the technique was already 2,000 years old at that time. Figure 25 shows an octahedral ear ornament composed of tiny gold spheres, made in Persia in the 9th or 10th century B.C. Curiously, none of the Greek atomists hit on the basic principle that these things illustrate, namely, that the mere stacking of equal isotropic spheres would give rise to the directional anisotropy of crystalline matter; and it was left to Johannes Kepler in 1611 first to publish this principle in a scientific treatise—if that is the proper term for his playful essay inspired by the hexagonality

[65] Edith Muller, *Gruppentheoretische und strukturanalytische Untersuchungen der maurischen Ornamente aus der Alhambra in Granada* (Ruschlikon, 1944).

of the snowflake.[66] Though the symmetry (not always sixfold!) of the snowflake appears commonly enough on today's Christmas cards, its decorative qualities do not seem to appear in art until after its depiction in scientific works. In the Far East, window lattice patterns representing interfering ice crystals on a frozen pond are common,[67] but the

FIG. 25.—Gold earring in polyhedral form composed of gold granules accurately soldered together and unconsciously illustrating the concept of the crystal lattice. From Marlik, ca. 1000 B.C. The granules in the top and bottom polyhedra are of different sizes, and their junction in the central plane illustrates an intercrystalline boundary. (Photo courtesy Iran Bastan, Tehran.)

[66] Johannes Kepler, *Strena seu de Nive Sexangula* (Frankfurt, 1611); trans. Colin Hardie (Oxford, 1966).

[67] Daniel S. Dye, *A Grammar of Chinese Lattices*, 2 vols. (Cambridge, Mass., 1937).

earliest oriental use of true snowflake symmetry appears to be that on a Japanese sword guard made by Harukiro Hirata in 1828, obviously related to the drawings (fig. 26) that were made by Toshitsura Doi under the influence of Dutch science and published five years later.[68]

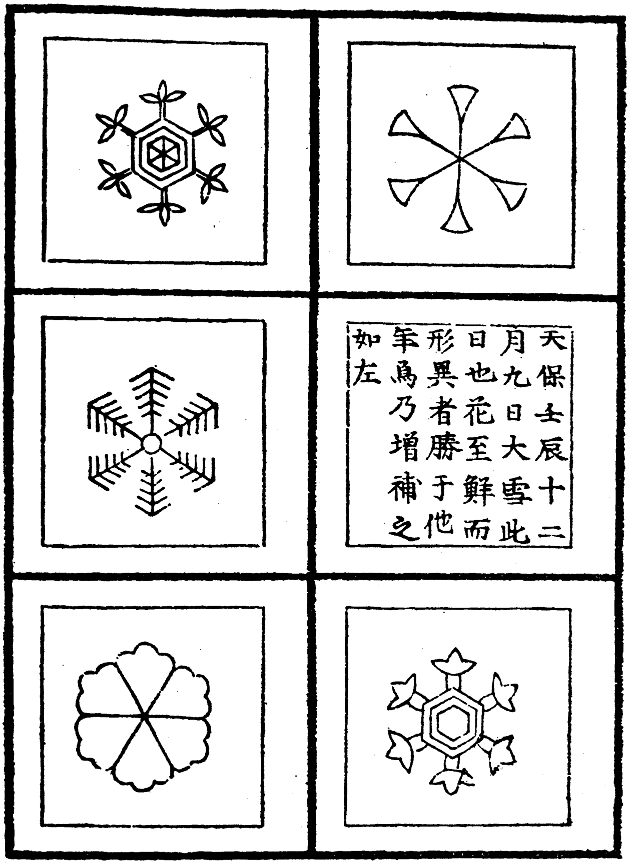


FIG. 26.—Drawings of snowflakes made with the "Dutch glass." Toshitsura Doi, *Sekka zusetsu* (Tokyo, 1833).

Figure 27 shows an elegant iron guard made somewhat later by a member of the famed Goto family.

Doi's drawings have a symmetry that is quite un-Japanese. Virtually every drawing of a snowflake that has been published, whether in a work of art or of science, depicts almost exact symmetry, reflecting

[68] Toshitsura Doi, *Sekka zusetsu* [Illustrations of snow crystals] (Tokyo, 1833; suppl., Tokyo, 1840). Both sections were reproduced with extended commentary and a summary in English by Teisaku Kobayashi (Tokyo, 1968).

the unwarranted but firm belief in the basic order of nature and the inability of the eye to see the unexpected. A glance at any photograph, or, better, the flakes themselves, will show many small differences between the six dendritic branches of even the best flake. And, of course, most snow falls as irregular aggregates displaying no symmetry whatever.

FIG. 27.—Japanese sword guard with snowflake design. Iron with inlay. Goto School, ca. 1850. (Photo courtesy Toledo Museum of Art.)

The principle of the crystal lattice is in every elementary textbook today; yet it proved difficult to accept and, despite Hooke's elaboration of the idea and Huygens's very effective use of it in explaining the properties of calcite crystals, it virtually disappeared for two and a half centuries as scientists preferred the concept of elementary polyhedra and, later, more elegant mathematical abstractions.[69]

Among the innumerable geometric patterns painted on early pottery in most cultures, there are many reminiscent of the magnetic-domain patterns of today's solid-state physicist. A more recent example of an artist's prescience lies in the work of the Dutch artist Maurits Escher, whose experiments with space-filling and repetitive patterns later pro-

[69] John G. Burke, *Origins of the Science of Crystals* (Berkeley, Calif., 1966).

vided the illustrations for an introductory book on symmetry for students of crystallography[70] and who in 1942 illustrated color-group symmetries "well before official crystallography even thought about them" and quite independently of their mathematical treatment by Schubnikov. Islamic tile workers had used them earlier, however, notably in the Alhambra (fig. 28).

In recent years there have been many exhibitions and books relating

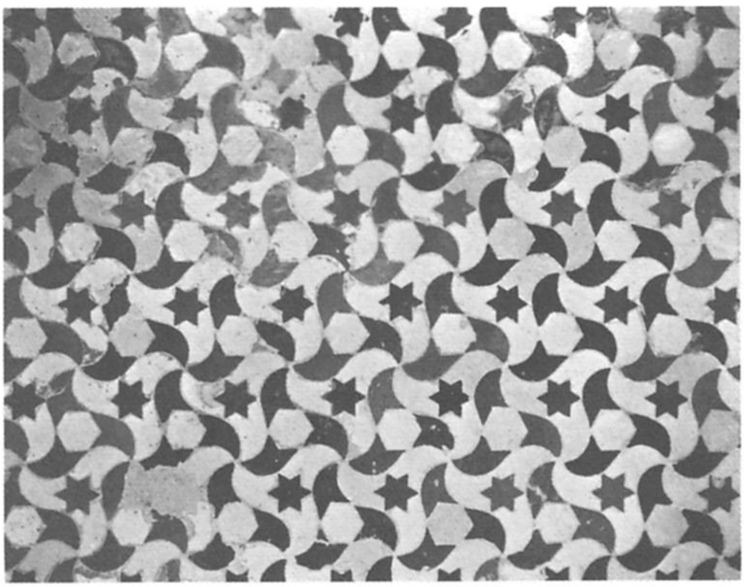


FIG. 28.—Mosaic tile work in the Alhambra at Granada, ca. A.D. 1325. (Photo by Phylis Morrison.)

scientific photographs to abstract art.[71] One of my photomicrographs of a copper silicon alloy once hung in the Museum of Modern Art! A particularly interesting coincidence is in some of the paintings of

[70] Caroline H. Macgillavry, *Symmetry Aspects of M. C. Escher's Periodic Drawings* (Utrecht, 1965); M. C. Escher, *Grafiek en tekengen* (Zwolle, 1960); translation, *The Graphic Work of M. C. Escher* (London, 1961, 2d ed., enl., 1967).

[71] See, for example, Gyorgy Kepes, ed., *Structure in Art and Science* (London and New York, 1965); Philip C. Ritterbush, *The Art of Organic Form* (Washington, D.C., 1968); Georg Schmidt, ed., *Kunst und Naturform: Form in Art and Nature* (Basel, 196[illegible] Paul Weiss, "Beauty and the Beast: Life and the Rule of Order," *Scientific Monthly* 81 (1955):286–99; Lancelot L. Whyte, ed., *Aspects of Form: A Symposium on Form in Nature and Art* (London, 1951); "The Artist Looks at the Scientist's World," exhibition organized by the Renaissance Society, University of Chicago, 1952.

Piet Mondrian which were later found to have almost exactly matched the microstructures of some cubic crystals containing randomly nucleated plates of a precipitated phase growing at right angles to each other until interference.[72] Such correspondence, of course, is only possible in a period in which artists are unconcerned with representation of the human world and are, for whatever reason, seeking a simplicity commensurate with that of the physicist. Perhaps, however, sculpture and paintings with human symbolism will some day be found to have played a similar role in connection with psychological science. Op art certainly belongs in laboratories studying the simpler aspects of the physiology of perception.

Let us return to history. If the crystal lattice was slow of conception, the idea of the atom, of course, was not. Is there perhaps a connection between the use of pebble mosaics to depict human and animal figures in Greece in the 4th century B.C. Greece and Leucippus's and Democritus's theories of matter? The concept that the distinguishing characteristics of matter arose in the shape, order, and orientation of parts in aggregation was certainly illustrated by the new mosaic forms, even if it was not suggested by them. At the present day, the printer's halftone is useful to illustrate information theory and discussions on structural hierarchy in matter.[73] And, again, in the same vein, is it absurd to suspect some connection between the revived receptivity to atomism at the end of the 16th century and the concurrent interest in the fine structure of a work of art that accompanied the new graphic methods? A rapid improvement in the quality of metal engraving accompanied the making of niello prints. Shading in both a woodcut and an engraving depends upon the control of discrete, nearly invisible lines which build up to a recognizable body. The painter, with continuous gradation of darkening and lightening even within a single brush stroke, does not need this kind of analysis; neither does the goldsmith with his repoussé bas-relief. Woodcuts in the nonatomistic Orient, exploiting mainly a variable quality of line and texture, are basically different from Western ones.

Today and Tomorrow

After this excursion into some of the past interaction between art and technology, it is tempting to speculate on their joint futures. It is fashionable today to note the similarity between the artist's creative insight

[72] R. W. Cahn, "Art in Science, Science in Art," *Museum* (UNESCO) 21 (1968): 16–21.

[73] Paul A. Weiss, "1 + 1 $\neq$ 2," in *Neurosciences: A Study Program,* ed. Gardner C. Quarton, Theodore Melnechuk, and Francis O. Schmidt (New York, 1967), pp. 801–21.

and that of the scientist, but for some reason the technical side of art has been downgraded as "mere" technique. Yet the handling of matter will always be necessary to give reality to the artist's all-important vision. Without it he cannot influence the minds and feelings of other people. Moreover, since technique relates more closely to the everyday experience of most men and women, especially when they are young, an interest in it can provide a path to the deeper meaning of art and lead to an understanding of things that the intellectual has never been able to communicate. The artist, if not every art historian, has always known that technology is a basically important human activity.

The recent trend away from representational art in the Western world has, however, been accompanied by a perceptibly increased interest in techniques and materials and consequently by more widespread appreciation of the "minor" arts that make up such a major part of the archaeological record if not of art-historical writings.[74] This has inevitably been accompanied by an increased interest in Oriental art, for, while Western art tells mainly about individuals, ideas, and institutions, Oriental art (in accord with Oriental philosophy) tells more about nature—and it sensitively uses the properties of matter to do this. The subtle representations of the Chinese landscape painter arise from the properties of colloidal carbon and water interacting with the capillarity of various surfaces. The ceramist of the Far East can sometimes reveal the essence of things even better than can a painter because the ceramist's product constitutes in itself a direct example of the balance of natural principles, not merely a representation of them, selected and controlled by the potter just enough to invite appreciation by the human eye and hand.

The natural forces operating on matter are at last being consciously utilized in Western art: For example, Jēkabs Zvilna of Toronto has been producing two-dimensional patterns (visible only with optical enlargement) that result from the interaction of surface tension, shrinkage stresses, volatility, viscosity, and other physical forces working on selected substances under boundary conditions that he deliberately sets up. Though both the method and the results bear some relation to what ceramists have long done, especially in Japan, the nonnatural scale and the freer choice of materials gives his patterns a highly contemporary look. The recent "Kalliroscope" of Paul Matisse is a simple but perpetually fascinating device revealing shear gradients in a liquid

[74] It is interesting to note that the published catalogs of three recent art exhibitions have long technical introductions: Herbert Hoffmann and Patricia Davidson, *Greek Gold* (Boston, 1965); Dominique de Menil, *Made of Iron* (Houston, Tex., 1966); David G. Mitten and Suzannah H. Doeringer, *Master Bronzes of the Classical World* (Cambridge, Mass., 1967).

containing floating micron-size reflecting platelets: patterns of endless variety and subtlety can be produced by controlling the conditions of turbulence or thermal convection, and close inspection even reveals to the naked eye the effects of molecular agitation (Brownian movement). Many other artists are experimenting with the aesthetic possibility inherent in systems the details of which are fully determined by physical forces, but the boundary conditions are set by human intervention.

Many artists are currently exploring the properties of polarized light, of kinetic and balanced motion and flow, of simple magnetic interaction (fig. 29), and of other phenomena which in the 19th century were used as rudimentary lecture demonstrations and laboratory experiments to evoke the interest of students in science.[75] It is high time that scientists admit that their experience in the laboratory is an aesthetic one, at times acutely so: the arid form of presenting their results has disguised this, and their respectable logical front often makes it invisible even to a student. The artist's interest in this aspect of science is very valuable. The introduction of scientific toys, under whatever name, to the general public and the opportunity to experience natural phenomena can only be applauded. The modern sculptor's skill in invoking viewer participation can aid enormously in the teaching of science, but his devices are usually of such simplicity that the initial feeling of pleasure cannot deepen by repetition into a rich aesthetic experience.

The visual excitement of the structures revealed by the microscope and electron microscope, of ion tracks in cloud chambers and interference patterns, has given rise to many fine exhibitions which have enriched the artist's vocabulary at the same time that they have heightened the scientist's sensibility. Yet it seems to me that in most of this the artist is just following others and is not fulfilling his particular role of revealing new significances in large, complex, perhaps social, patterns. Science is proliferating into more and more precise studies of more and more details. Higher energies beckon always away from the understanding of things on a human level to the smaller and simpler units of matter. So much knowledge has been acquired in this way that some scientists have claimed that no valid meaning can be established except by physical science. The most exciting frontier of biology has been on the molecular level, not life itself, which requires higher organization. After decades of neglect, however, something like old-fashioned natural history seems to be coming back into its own: the cell, biological form, and especially that comprehensive subject known as ecology, which is almost the art of science. Can the same thing

[75] Jack Burnham, *Beyond Modern Sculpture* (London and New York, 1968).

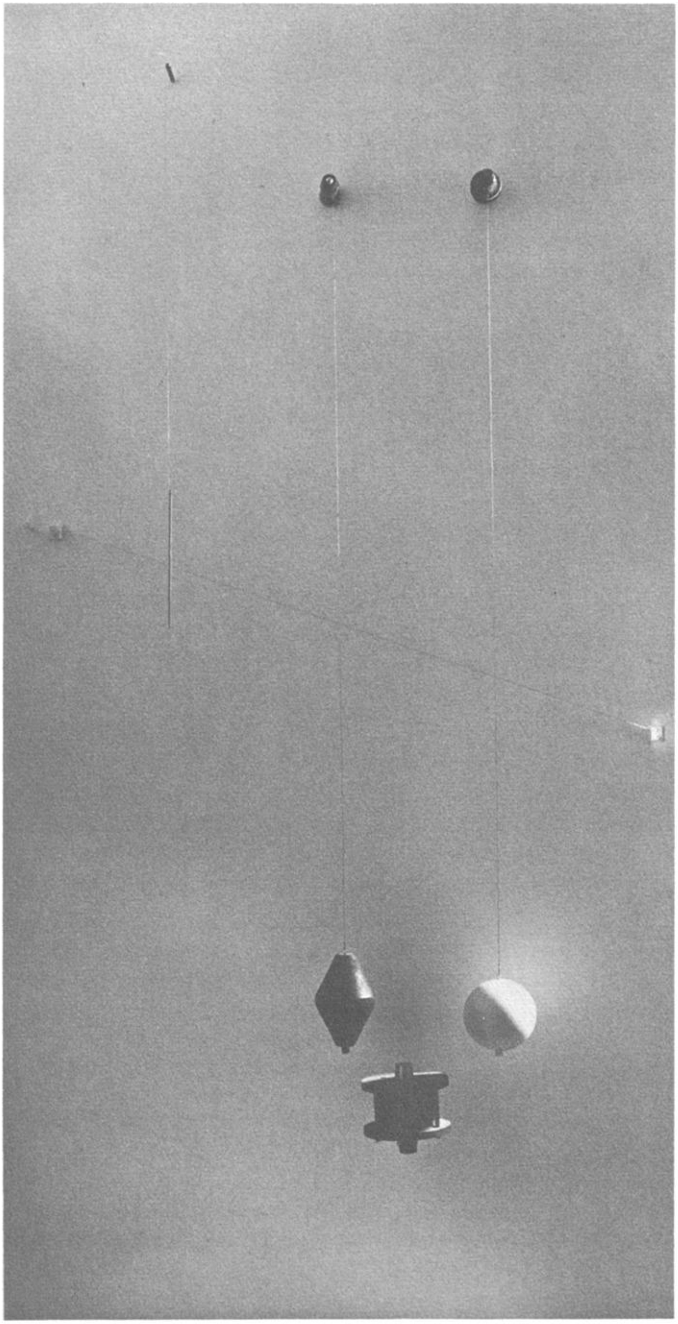


FIG. 29.—*Ballet magnétique* by Greek-American sculptor Takis (1968). (The solenoid, intermittently energized, causes irregular movements of the two pendulum bobs, which are suspended on steel wires that strike the taut transverse wire, producing musical notes and dancing of the suspended rod at the left.)

happen in other areas? And can the artist, when he has learned some of the rudiments of science and technology, help?

Throughout history there has been a slow separation of art from the arts, and of science from both. As science became more definite, it became increasingly useful to technology, and it has given precision to both the design and the control of processes. With art, however, the very utility of its contributions to industrial design and advertising seems rather to have forced out the one component of it that is most needed, and today we are faced with the curious phenomenon of art being mainly a comment and a much-needed protest rather than a constructive suggestion of a way toward deeper understanding. Artists have found much to interest them in both the scientific and technological world, and they have shown that there is much beauty even in things such as galvanized iron roofing and the intricacies of stairs and piping in a chemical processing plant, to say nothing of the elegant patterns of electronic gadgetry. The strength of steel and concrete and the beauty of a streamlined surface are proper aesthetic experiences in today's world, and they become more so as artists explore their meaning. Many sculptors have learned to enjoy the properties of steel and to exploit the cutting and welding torch in producing sculpture. The role of the artist in pointing to common things and making one pause to look at them has always been important. He now plays a similar role in relation to science, not only in finding the visual delights of the New Landscape[76] but also in calling attention to experiences of the other senses that are possible in a scientific or technological environment.

Technology is by its very nature complex and thus is incapable of being completely understood. There are two kinds of simplification that can make this complexity handleable. The first is the scientists' recognition of the units and their interaction on a small precise scale and the other is the recognition of the connectivity of units—which sometimes is systems analysis but more constructively is art. As technology has passed from the individual work of craftsmen to an aggregate of integrated systems, the significance of individual processes has been lost precisely at the moment that they become most efficient. The discovery of new techniques owes less to artistic curiosity but now occurs in well-financed research laboratories and is increasingly dependent on science. Yet does not the transition from craftsman to technologist itself suggest a new area in which the artist should play a role? The new level of complexity in technology requires a new level

[76] Gyorgy Kepes, *The New Landscape in Art and Science* (Chicago, 1956); Kepes, ed., *Structure in Art and Science* (New York, 1965).

of art, perhaps almost a social one. Indeed, the artist is needed now as never before, and only by an introduction of the artist's general sense of relationships will it be possible to restore the balance between social and individual needs. At least some of the artist's work will be devising schemes in which the pleasure of an intensely individual experience can interact with that of others to produce a more viable society than at present. The artist can highlight discrepancies and point up problems that should be solved before they become generally obvious, while it is the job of the technologist to say how to solve them, and to do so.

Here it should be noted that there are more possibilities of diversity above human scale than below it, and the dangers of oversimplification in social matters are correspondingly greater than in the realm of physics and chemistry. The more that individuals are able to enhance their differences without loss of contact, the richer their lives will be. Technology at last makes real diversity possible, but democratic egalitarianism is in danger of eliminating it. Part of the artist's job will be to oppose oversimplification in this world of immensely diverse possibilities. But needed beyond all else is an aesthetic feeling in the hearts and minds of technologists, who are so rapidly, at other peoples' behest, despoiling the Old Landscape.

Technology and Culture

The International Quarterly of the Society for the History of Technology

OCTOBER 1970, VOLUME 11, NUMBER 4
THE UNIVERSITY OF CHICAGO PRESS

Nastic *et al. Journal of Internet Services and Applications* (2015) 6:21
DOI 10.1186/s13174-015-0037-1

Journal of Internet Services and Applications
a SpringerOpen Journal

RESEARCH

Open Access

SDG-Pro: a programming framework for software-defined IoT cloud gateways

Stefan Nastic*, Hong-Linh Truong and Schahram Dustdar

Abstract

Recently, emerging IoT cloud systems create numerous opportunities for a variety of stakeholders in terms of optimizing their existing business processes, as well as developing novel cross-organization and cross-domain applications. However, developers of such systems face a plethora of challenges, mainly due to complex dependencies between the application business logic and the underlying IoT cloud infrastructure, as well as difficulties to provision and govern vast, geographically distributed IoT cloud resources. In this paper, we introduce SDG-Pro – a novel programming framework for software-defined IoT cloud systems. The main features of our framework include programming abstractions: Software-Defined Gateways, Intents, Intent Scopes, and Data and Control Points, as well as provisioning and governance APIs that allow for programmatic management of software-defined gateways throughout their entire lifecycle. The SDG-Pro framework enables easier, efficient and more intuitive development of IoT cloud applications. It promotes the everything-as-code paradigm for IoT cloud applications in order to provide a uniform, programmatic view on the entire development process. To illustrate the feasibility of our framework to support development of IoT cloud applications, we evaluate it using a real-world case study on managing fleets of electric vehicles.

Keywords: IoT cloud applications programming; Software-defined gateways; IoT cloud systems

1 Introduction

Emerging IoT cloud systems extend the traditional cloud computing systems beyond the data centers and cloud services to include a variety of edge IoT devices such as sensors and sensory gateways. Such systems utilize the IoT infrastructure resources to deliver novel value-added services, which leverage data from different sensor devices or enable timely propagation of decisions, crucial for business operation, to the edge of the infrastructure. On the other side, IoT cloud systems utilize cloud's theoretically unlimited resources, e.g., compute and storage, to enhance traditionally resource constrained IoT devices.

In order to facilitate development of IoT cloud systems, existing research and industry have produced numerous infrastructure, platform and software services as well as frameworks and tools [1–6]. These advances set a cornerstone for proliferation of (unified) IoT cloud platforms and infrastructures, which offer a myriad of IoT cloud capabilities and resources. Lately, we have been exploring software defined approaches and introduced a design methodology and a set of software defined principles for IoT cloud [7] in order to facilitate utility-oriented delivery of the IoT cloud resources, provide elasticity support for the IoT cloud systems and enable automated and logically centralized provisioning of the geographically distributed IoT cloud infrastructure. Generally, the software-defined IoT cloud systems abstract from low-level resources (e.g., hardware) and enable their programmatic management through well-defined APIs. They allow for refactoring the underlying IoT cloud infrastructure into finer-grained resource components whose functionality can be (re)defined in software, e.g., applications, thus enabling more efficient resource utilization and simplifying management of the IoT cloud systems. However, most of the contemporary approaches dealing with IoT cloud are intended for platform/infrastructure providers and operations managers. Therefore, from the developer's perspective there is a lack of structured, holistic

*Correspondence: snastic@dsg.tuwien.ac.at
Distributed Systems Group, TU Wien, Argentinierstrasse 8/184-1, 1040 Vienna, Austria

approach to support development of the IoT cloud systems and applications. Concrete abstractions and mechanisms, which enable efficient, more intuitive and scalable application development still remain underdeveloped.

1.1 Contributions

In this paper we introduce SDG-Pro – a novel programing framework for software-defined IoT cloud systems. The main contribution of the paper is SDG-Pro's programming model. It provides a unified, programmatic view for the entire development process (*everything as code*) of IoT cloud applications, thus making it easily traceable and auditable. We demonstrate the main advantages of our programming model in terms of easier, efficient and more intuitive application development, by using a real-world case study on managing fleets of vehicles.

This paper substantially extends and refines our previous work presented in [7, 8]. In [8], we introduced a programming model for developing cloud-based IoT services. The SDG-Pro framework extends that approach, by introducing programming support for developing the (edge) device services, i.e., monitoring and control tasks (Section 4.2). In [7], we introduced a conceptual model and main design principles for software-defined IoT cloud systems. The SDG-Pro framework builds on these concepts and extends our previous work by introducing comprehensive programming support for unified development, provisioning (Section 4.3) and governance of (Section 4.4) IoT cloud applications.

1.2 Paper organization

The remainder of the paper is structured as follows. In Section 2, we describe a motivating scenario and main research challenges; Section 3 outlines the design of the SDG-Pro framework; Section 4 presents the framework's programming model; In Section 5, we outline SDP-Pro's runtime support; Section 6 presents our experiments; Section 7 discusses the related work; Finally, Section 8 concludes the paper and gives an outlook of our future research.

2 Motivation and research challenges

2.1 Scenario

Let us consider a realistic application scenario in the domain of vehicle management that we will refer to throughout the paper.

Fleet Management System (FMS) is a real-world IoT cloud system responsible for managing fleets of electric vehicles deployed worldwide, e.g., on different golf courses. The vehicles are equipped with an *on-board device*, acting as a gateway to vehicles' sensors and actuators, as well as offering resources to execute device services in the vehicles. The vehicles communicate with the cloud via 3G or Wi-Fi networks to exchange telematic and diagnostic data. On the cloud, FMS provides different applications and services to manage this data. Relevant services include realtime vehicle status, remote diagnostics, and remote control. In general, different stakeholders rely on the FMS applications to manage their portion of the fleet and optimize tasks specific to their business model.

The cloud plays a crucial role for the FMS due to several reasons. Besides utilizing the edge infrastructure, e.g., vehicle sensors and the on-board devices, the FMS heavily relies on cloud infrastructure to be able to process and reliably store vast amounts of sensory data as well as to connect large number of vehicles and provide centralized, simultaneous access to the geographically distributed fleet (needed by the services such as emergency remote fleet control). Further, since different stakeholders manage different portions of the fleet, i.e., are allowed to access specific vehicles and their data, FMS has to be able to support multiple tenants. In addition, since many of FMS's services can be elastically scaled down in off peak times, e.g., during the night, the elastic nature of cloud plays a significant role, especially in terms of costs control, since systems of such scale as our FMS incur very high costs in practice (e.g., of computation or networking).

The FMS runs atop a complex IoT cloud infrastructure, which includes a variety of IoT cloud resources. Figure 1 gives a high-level overview of the common elements in FMS's architecture and deployment. FMS deployment topologies span across the entire IoT cloud infrastructure, i.e., from large data centers to the edge of the network, resulting in complex dependencies among the business logic services, but also between such services and the underlying infrastructure. Therefore, developers need to consider numerous infrastructure resources and their properties such as availability of sensors, devices ownership and their location.

The FMS applications perform a variety of analytics and are mostly characterized by a reactive behavior. They receive (monitoring) data, e.g., a change in vehicles' operation and, as a response, perform (control) actions. Such monitoring and control tasks are executing in heterogeneous, dynamic FMS environment and interact with many geographically distributed vehicles and their low-level capabilities, e.g., engine control points. Further, FMS applications have different requirements regarding communication protocols. For example, the fault alarms need to be pushed to the services, e.g, via MQ Telemetry Transport (MQTT) and vehicle's diagnostics should be synchronously accessed via RESTful protocols such as Constrained Application Protocol (CoAP) or Simple Measurement and Actuation Profile (sMAP). Therefore, from the developers perspective such tasks and capabilities need to be decoupled from the underlying physical

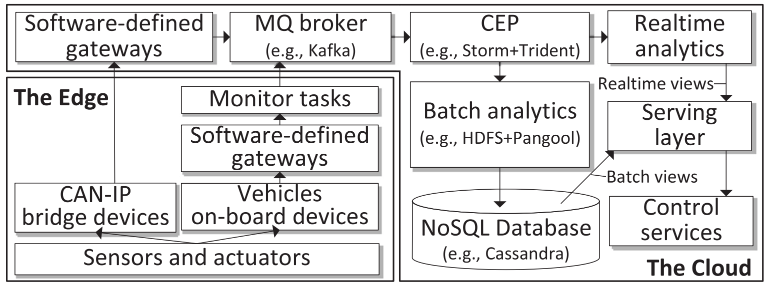


Fig. 1 Overview of FMS architecture and deployment

infrastructure, but also easily specified, provisioned and managed programmatically (as code), without worrying about the complexity of low-level device services, communication channels and raw sensory data streams.

The currently limited development support regarding the FMS requirements and features (as discussed in Section 7), renders the development of its applications a complex task. Consequently, system designers and application developers face numerous challenges to design and develop IoT cloud applications.

2.2 Research challenges

RC-1 – The development context of IoT cloud applications has grown beyond writing custom business logic (e.g., services) components to also considering the involved IoT devices (e.g., their capabilities) as well as the deployment and provisioning of such services across the IoT cloud infrastructure. The main reasons for this are complex and strong dependence of the business logic on the underlying devices (and their specific capabilities), novel (resource) features that need to be considered, such as device location and the heterogeneity of the utilized IoT cloud resources. Unfortunately, developers currently lack suitable programming abstractions to deal with such concerns in a unified manner, from early stages of development.

RC-2 – IoT cloud applications execute in very dynamic, heterogeneous environments and interact with hundreds or thousands of physical entities. Therefore, monitoring and controlling these entities in a scalable manner is another challenge for developers of IoT applications, mainly because they need to dynamically identify the scope of application's actions, depending on the task at hand, but also express its business logic independently of the low-level device capabilities.

RC-3 – The IoT cloud applications mostly rely on common physical infrastructure. However, IoT cloud infrastructure resources are mostly provided as coarse-grained, rigid packages. The infrastructure components and software libraries are specifically tailored for the problem at hand and do not allow for flexible customization and provisioning of individual resource components or runtime topologies. This inherently hinders self-service, utility-oriented delivery and consumption of IoT cloud resources at fine granularity levels.

RC-4 – Due to dynamicity, heterogeneity and geographical distribution of IoT cloud, traditional provisioning and governance approaches are hardly feasible in practice. This is mostly because they implicitly make assumptions, such as physical on-site presence, manual logging into devices, understanding device's specificities, etc., which are difficult, if not impossible, to achieve in IoT cloud systems. In spite of this, techniques and abstractions, which provide a programmatic, conceptually centralized view on system provisioning and runtime governance are largely missing.

In the rest of the paper, we introduce our SDG-Pro framework and focus on describing and evaluating its programming model for IoT cloud applications.

3 The SDG-Pro framework

The main aim of our SDG-Pro (*Software-Defined Gateways Programing framework*) is to provide programming support for IoT cloud application developers, which offers a set of adequate programming abstractions to facilitate overcoming the aforementioned challenges. To this end, SDG-Pro, enables expressing application's provisioning, governance and business logic programmatically, in a uniform manner. By raising the level of programming abstraction, SDG-Pro reduces the complexity of application development, makes the development process traceable and auditable and improves efficacy and scalability of application development.

SDG-Pro does not propose a novel software-defined approach for IoT cloud systems. It builds on the design principles that were elicited in our previous research [7, 9, 10] and adopts development methodology we proposed in [11], extending them to provide programming abstractions that facilitate development of the essential application artifacts. SDG-Pro's programming model is designed to enforce the main design principles of software-defined IoT cloud systems at application level, from the early development stages.

3.1 Main design principles and development methodology for software-defined IoT cloud systems

As we have shown in [7], software-defined IoT cloud systems comprise a set of resource components, provided by IoT cloud infrastructure, which can be provisioned and governed at runtime. Such resources (e.g., sensory data streams), their runtime environments (e.g., gateways) and capabilities (e.g., communication protocols and data point controllers) are described as software-defined IoT units.

The software-defined gateways (cf. Fig. 1) are a special type of such units and they are the main building blocks of IoT cloud infrastructure, e.g., similar to VMs in cloud computing. In our conceptual design of software-defined IoT cloud systems, such gateways abstract resource provisioning and governance through well-defined APIs and they can be composed at different levels, creating virtual runtime topologies for IoT cloud applications. This enables opening up the traditional infrastructure silos and moving one step higher in the abstraction, i.e., effectively making applications independent of the underlying rigid infrastructure. As we have extensively discussed in our previous work [7, 9, 10], the main design principles of software-defined IoT systems include:

Everything as code – All the concerns, i.e., application business logic, but also IoT cloud resources provisioning and runtime governance, should be expressed programmatically in a unified manner, as a part of the application's logic (code).
Scalable development – The programming abstractions exposed to the developers need to support scalable application development, i.e., shield the developers from dealing with the concerns such as manually referencing individual devices or managing the low-level data and control channels.
API Encapsulation – IoT cloud resources and capabilities are encapsulated in well-defined APIs, to provide a uniform view on accessing functionality and configurations of IoT cloud infrastructure.
Declarative provisioning – The units are specified declaratively and their functionality is defined programmatically in software, using well-defined API and available, familiar software libraries.
Central point of operation – Enable conceptually centralized (API) interaction with the software-defined IoT cloud system to allow for a unified view on the system's operations and governance capabilities (available at runtime), without worrying about low-level infrastructure details.
Automation – Main provisioning and governance processes need to be automated in order to enable dynamic, on-demand configuring and operating the software-defined IoT cloud systems, without manually interacting with IoT devices (e.g., logging in the gateways).

As proposed in [11], building IoT cloud systems includes creating and/or selecting suitable software-defined IoT units, provisioning and composing more complex units and building custom business logic components. This (iterative) development process is structured along four main phases (cf. Fig. 2): i) Initially, developers need to design and implement the software-defined IoT units or obtain them form a third-party, e.g., in a market-like fashion. Among other things the IoT units support execution of the light-weight device services (monitor and control tasks), i.e., from the software engineering perspective they encapsulate such tasks, comprising domain libraries; ii) Next, the developers need to design and provision the required application topologies. This process includes implementing the dependencies among the business logic services, but also between such services and the underlying infrastructure; iii) Building custom business logic components mainly involves developing device services and implementing reactive business logic (e.g., cloud services) around the device services; iv) The developers implement operational governance logic for managing the IoT units during application runtime.

3.2 SDG-Pro architecture overview

Generally, the SDG-Pro framework is distributed across the clouds, i.e., large data centers, and "small" IoT devices,

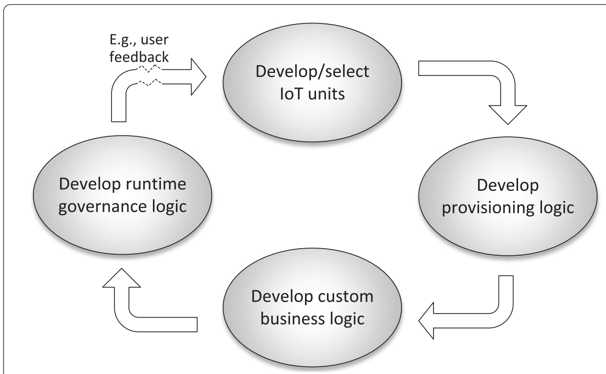


Fig. 2 Most important steps in development methodology for software-defined IoT cloud systems (partial view)

e.g., physical gateways or cloudlets. It is designed based on the microservices architecture, which among other things enables evolvable and fault-tolerant system, while allowing for flexible management and scaling of individual components. Figure 3 gives a high-level overview of SDG-Pro's architecture and main IoT cloud application artifacts. These artifacts can be seen as executables produced by the aforementioned development methodology. To support the development of such artifacts our framework provides a set of programming abstractions (depicted as gray components in Fig. 3 and described later in Section 4) and runtime support mechanisms (Section 5).

The runtime mechanisms are part of the SDG-Pro's Runtime support (cf. Fig. 3), which underpins the programming abstractions exposed to the developers, i.e., provides an execution environment for IoT cloud applications. It takes over a set of responsibilities such as placement of the software-defined gateways, their runtime migration and elasticity control, infrastructure topology management and application scope coordination. By doing so, it does most of "heavy lifting" on behalf of the applications, thus supporting the developers to easier cope with the diversity and geographical distribution of the IoT cloud and enabling better utilization of the numerous edge devices.

Internally, our framework's runtime support comprises several microservices, which can be grouped into: API-Manager, IoT units management layer, Intents runtime container and Operational governance layer. The *API-Manager* exposes governance capabilities and the low-level data and control channels from the IoT cloud infrastructure to the applications via well-defined APIs and handles all API calls from such applications. The *IoT units management layer* provides mechanisms and agents to support instantiating, provisioning and deploying software-defined gateways programmatically and on-demand. The *Intents runtime container* is responsible to handle incoming application requests (Intents) and select, instantiate and invoke device services (tasks), based on the information provided in the intents. It also enables applications to dynamically delimit the scopes of their actions, by providing support for IntentScopes resolution. Finally, the *Operational governance layer* supports execution of the governance processes by enabling remote invocation of governance capabilities and mapping of API calls on underlying devices via governance agents.

4 SDG-Pro's programming model

4.1 Structure of IoT cloud applications

The main purpose of our programming model is to provide a programmatic view on the whole application ecosystem, i.e., the full stack from the infrastructure to software components and services. The main principle behind our programming model is *everything as code*. This includes providing support for writing IoT cloud applications' business logic, as well as representing the underlying infrastructure components (e.g., gateways) at the application level and enabling developers to programmatically determine their deployment and provisioning. Figure 4 shows a component diagram with the logical structure of IoT cloud applications. The main components of such application include: custom business

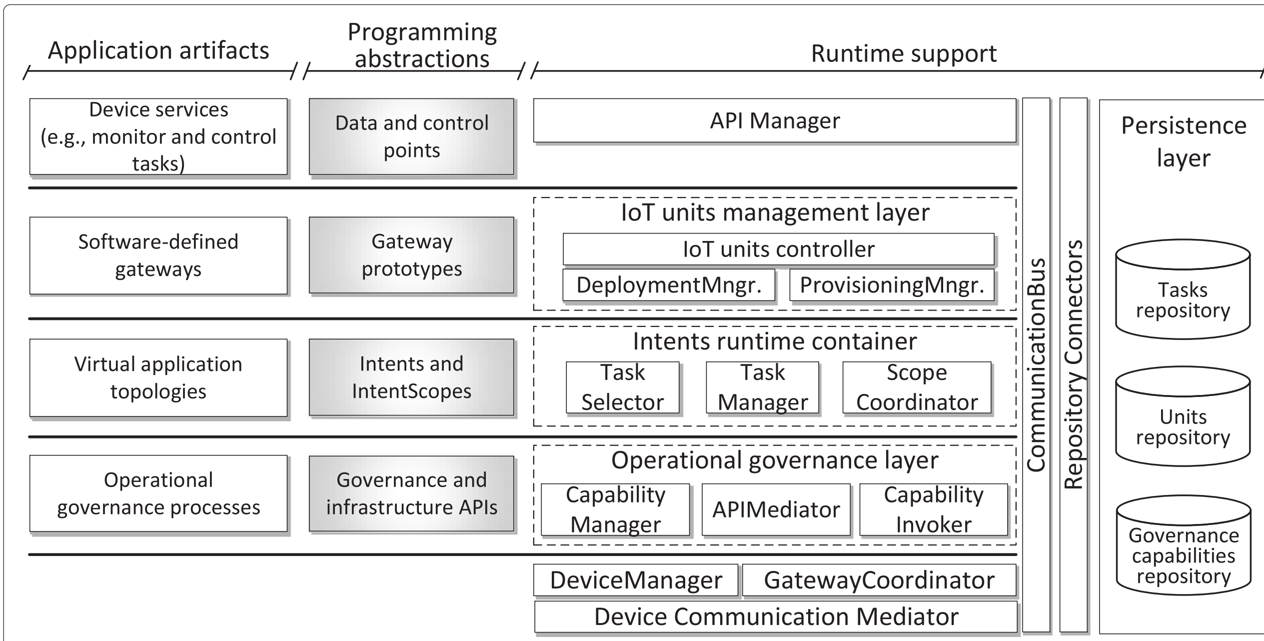


Fig. 3 Overview of SDG-Pro's architecture and main artifacts of IoT cloud applications

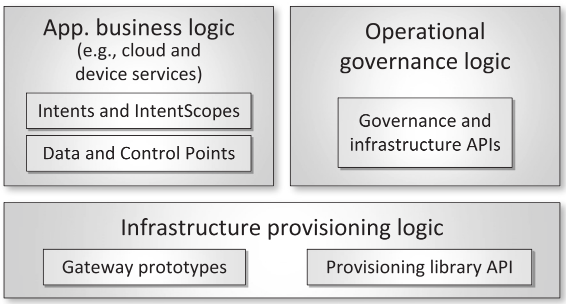


Fig. 4 Overview of IoT cloud application structure

logic components (e.g., cloud services and device services); resource provisioning and deployment logic (custom or stock component provisioning); and operational governance logic.

4.2 Programming support for business logic services

In SDG-Pro we distinguish between two types of business logic services: device-level services and cloud services. Device-level services are executed in IoT devices and implement control and monitor tasks. For example, a monitoring task includes processing, correlation and analysis of sensory data streams. To support task development, the SDG-Pro framework provides *data and control points*, which are described later in this section.

The cloud services usually define virtual service topologies by referencing the tasks. At the application level, we provide explicit representation of these tasks via *Intents*, i.e., developers write *Intents* to dynamically configure and invoke the tasks. Further, developers use *IntentScope*s to delimit the range of an *Intent*. For example, a developer might want to code the expression: "stop all vehicles on golf course X". In this case, "stop" is the desired *Intent*, which needs to be applied on an *IntentScope* that encompasses all vehicles with the location property "golf course X".

Intents Intent is a data structure describing a specific task which can be performed in a physical environment. In reality, *Intent*s are processed and executed on the cloud platform, but enable monitoring and controlling of the physical environments. Based on the information contained in an *Intent*, a suitable task is dynamically selected, instantiated and executed.

Depending on the task's nature, we distinguish between two different types of *Intent*s: *ControlIntent* and *MonitorIntent*. *ControlIntent*s enable applications to operate and invoke the low-level components, i.e., provide a high-level representation of their functionality. *MonitorIntent*s are used by applications to subscribe for events from the sensors and to obtain devices' context.

Figure 5 shows the *Intent* structure and its most relevant parts. Each *Intent* contains an ID, used to correlate invocation response with it or apply additional actions on it. Additionally, it contains a set of headers, which specify meta information needed to process the *Intent* and bind it with a suitable task during the runtime. Among other things, headers carry *Intent's* name and a reference to an *IntentScope*. Further, an *Intent* can contain a set of attributes, which are used by the runtime to select the best matching task instance in case there are multiple Intent implementations available. Finally, *Intent* can contain data, which is used to configure the tasks or supply additional payload. Generally, *Intents* allow developers to communicate to the system what needs to be done instead of worrying how the underlying devices will perform the specific task.

IntentScopes IntentScope is an abstraction, which represents a group of physical entities (e.g., vehicles) that share some common properties, i.e., a set of software entities in the cloud, which represent corresponding physical entities. The IntentScopes enable developers to dynamically delimit physical entities on which an *Intent* will have an effect. The SDG-Pro framework provides mechanisms to dynamically define and work with *IntentScope*s on the application level.

To define an *IntentScope* developers specify properties, which need to be satisfied by the physical entities to be included in the scope. To enable *IntentScope* bootstrapping, we provide a special type of *IntentScope*, which is called *GlobalScope*. It defines the maximal scope for an application and usually contains all physical entities administered by a stakeholder at the given time. Therefore, it is reasonable to assume that the *GlobalScope* is slow-changing over time and it can be configured by a user, e.g., a golf course manager. Our programming model allows *IntentScopes* to be defined explicitly and implicitly, i.e., developers can explicitelly add entities to the scope by specifying their IDs or recursively prune the *GobalScope*.

Formally, we use the well-known set theory to define *IntentScope* as a finite, countable set of entities (set elements). The *GlobalScope* represents the universal set, denoted as S^{max}, therefore, $\forall S(S \subseteq S^{max})$, where S is an *IntentScope*, must hold. Further, for each entity E in the system general membership relation $\forall E(E \in S | S \subseteq S^{max})$, must hold. Therefore, an entity is the unit set, denoted as S_{min}. Empty set $\emptyset$ is not defined, thus, applying an *Intent* on it results with an error. Finally, a necessary condition for an *IntentScope* to be valid is: *IntentScope* is valid iff it is a set S, such that $S \subseteq S^{max} \wedge S \not\equiv \emptyset$ holds. Equation 1 shows operations used to define or refine an *IntentScope*. The most interesting operation is $\subseteq_{cond} S$. It is used to find

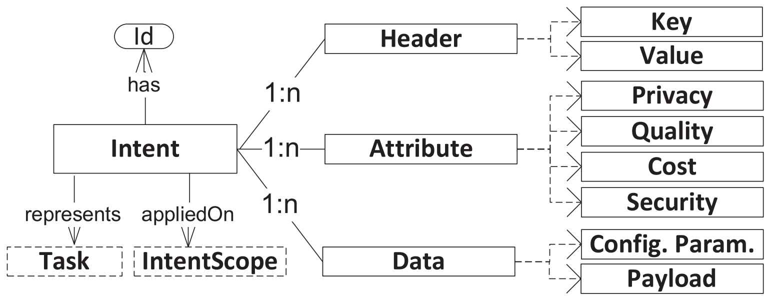


Fig. 5 Intent structure

a subset ($\hat{S}$) of a set S, which satisfies some condition, i.e., $E \in \hat{S} \mid E \in S \wedge cond(E) = True$.

$$S = S_{min}|S^{max}| \subseteq_{cond} S|S \cup S|S \cap S|S \setminus S \qquad (1)$$

Listings 1 and 2 show example usage of Intents and IntentScopes.

Listing 1 Example usage of MonitorIntent and GlobalScope

```
Intent eFault = Intent.newMIntent("EnergyFault");
//monitor whole fleet
eFault.setScope(IntentScope.getGlobal());
notify(eFault,this);//invoke task
//callback function called on event arrival
public void onEvent(Event e){//perform some action}
```

Listing 2 Example usage of ControlIntent and IntentScope

```
IntentScope cs = delimit(IntentScope.getGlobal(),
 Cond.isTrue(eFault)); //eFault defined in Listing1
Intent eCons = Intent.newCIntent("ReduceEnergy");
eCons.setScope(cs);//set intent scope
eCons.set("speed").value("5");
eCons.set("RPM").value("1100");
send(eCons); //invoke task
```

Data and Control points The main purpose of the data and control points is to support development of the light-weight device services. Generally, they represent and enable management of data and control channels (e.g., device drivers) to the low-level sensors/actuators in an abstract manner. Data and control points mediate the communication with the connected devices (e.g., digital, serial or IP-based) and also implement communication protocols for the connected devices, e.g., Modbus, CAN or SOX/DASP.

The data and control points enable developers to interact with sensory data streams and actuating functionality in a unified manner, independent of communication type, e.g., protocol. The most important concept behind data and control points are the *virtual buffers*, which are provided and managed by our framework. In general, such buffers enable virtualized access to and custom configurations of underlying sensors and actuators. They act as multiplexers of the data and control channels, thus enabling the device services to have their own view of and define custom configurations for such channels. For example, an application can configure sensor poll rates, activate a low-pass filter for an analog sensory input or configure unit and type of data instances in the stream.

Listing 3 Example usage of data point

```
DataPoint dataPoint = new DataPoint();
// Query available buffers
Collection<BufferDescription> availableBuffers
 = dataPoint.queryBuffers(new SensorProps(...));
// Assign the buffers to the data point
dataPoint.assign(availableBuffers.get(0));
dataPoint.setPollRate(5);
```

Listing 3, gives a general example of how to define a data point. It shows a data point with one stream of simple data instances that represent, e.g., vehicle's tire speed, based on the required sensor properties. By default the data points are configured to asynchronously push the data to the applications at a specific rate, which can be configured as shown in the example. The application defines a callback handler, which contains some data processing logic, e.g., based of complex event processing techniques. Additionally, the data and control points offer a `read` operator that can be used to sequentially (or in batch) read a set of instances from a stream, e.g., in order to perform more complex stream processing operations.

4.3 Programmatic infrastructure provisioning with software-defined gateways

The most important abstraction for provisioning IoT cloud infrastructure is the software-defined gateway. In our programming model software-defined gateways are treated as first-class citizens. This allows the developers to specify, manipulate and manage the IoT cloud infrastructure resources programmatically from within the application logic.

Generally, provisioning part of the application logic is used to programmatically specify the infrastructure dependencies, i.e., the state of the infrastructure required by the business logic services to execute correctly. To this end, our framework supports the developers to perform two main tasks. First, the developers can programmatically define the software-defined gateways and specify their internal structure. Second, our framework supports the developers to deploy such gateways atop IoT cloud (e.g., data centers or physical IoT gateways) form within the application logic. Therefore, provisioning logic is specified in software enabling the infrastructure dependencies and requirements to be defined dynamically and on-demand.

Figure 6 shows the typical structure of a software-defined gateway. We notice two important properties of software-defined gateways. First, to technically realize software-defined gateways SDG-Pro offers gateway prototypes. These are resource containers, used to bootstrap more complex, higher-level gateway functionality. Generally, they are hosted atop IoT cloud and enriched with functional, provisioning and governance capabilities, which are exposed via well-defined APIs. Currently, our framework supports software-defined gateways based on kernel supported virtualization, but virtualization choices do not pose any limitations, because, by utilizing the well-defined API, our gateway prototypes can be dynamically configured, provisioned, interconnected, deployed, and controlled.

Second, developers use the software-defined gateways to programmatically provision and deploy required application services, but also to configure an execution environment for such services. Therefore, by utilizing the provisioning APIs, developers can customize the software-defined gateways to exactly meet the application's functional requirements. For example (Fig. 6), they can dynamically configure a specific cloud communication protocol, e.g., CoAP or MQTT, select services runtime, e.g., Sedona VM or configure data and control points, e.g., based on Modbus.

In order to provision a software-defined gateway, initially the developers need to specify the software defined gateway prototypes. Listing 4 illustrates how the gateway prototypes are programmatically defined with our framework. In this example, a software-defined gateway is created from a gateway prototype, based on BusyBox. In the background the framework creates a Linux container and installs the provisioning and governance agents on it (more details about this process are given in Section 5). In general, the agents expose the provisioning APIs, which are activated and available at that point. Afterwards, a developer provides a configuration model for the gateway. In this example the gateway is configured to be deployed on a specific host by setting the "host address". In case it is not set the framework uses the deployment class to determine the gateway placement. Finally, the developer specifies the class that contains the internal provisioning logic.

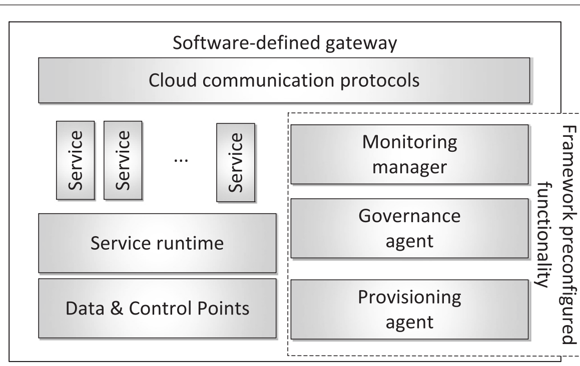


Fig. 6 Overview of software-defined gateway structure

Listing 4 Example of software-defined gateway prototype

```
//Define and configure a gateway
SDGateway gateway
= UnitsController.create(GType.BUSYBOX);
gateway.setId("gateway-X");
gateway.setHost("http://host_address");
gateway.setMetaData(Deployment.class);
gateway.addConfigClass(Provisioning.class);
```

Listing 5 illustrates our framework's support for dynamic provisioning of such gateways. The gateway provisioning logic contains the directives to internally provision the gateway, e.g., to install and configure device services, cloud communication libraries and data and control points. To this end, developers can use the framework's provisioning support, which contains the APIs, provided by the provisioning agent, and a provisioning library comprising a number of functions that facilitate provisioning of the software-defined gateways. In this example, we show how to provision a gateway with Sedona runtime.

Listing 5 Example of gateway provisioning API

```
String dest = ".../G2021/svm";
provisioner.CreateDirIfMissing(dest);
provisioner.CopyToDir("sedona-vm-1.2.28/svm", dest);
provisioner.setPermissions(dest, "a+x");
```

4.4 Programmatic governance and infrastructure APIs

4.4.1 Programmatic application governance

After an application is provisioned and deployed, a new set of runtime concerns emerges, e.g., dynamically reconfiguring sensor update rates or elastically scaling software-defined gateways. In order to address such concerns, application developers implement operational governance processes (cf. Fig. 4).

In our previous work [9, 10] we introduced a general approach for runtime operational governance in software-defined IoT cloud systems, as well as the concepts of operational governance processes that manipulate the states of software-defined gateways at runtime. Such processes can be seen as a sequence of operations, which perform runtime state transitions from a current state to some desired target state (e.g., that satisfies some non-functional properties, enforces compliance, or exactly meets custom requirements).

The core abstraction behind the operational governance business logic are *governance capabilities*. Generally, the governance capabilities represent the main building blocks of operational governance processes and they are usually executed in IoT devices. The governance capabilities encapsulate governance operations which can be applied on deployed software-defined gateways, e.g., to query the current version of a service, change a communication protocol, or spin up a virtual gateway. The framework enables such capabilities to be dynamically added to the system and supports managing their APIs. Generally, we do not make any assumptions about concrete capability implementations. However, the framework requires them to be packaged as shown in Fig. 7.

To enable programmatic operational governance our framework offers governance APIs that are used by application developers to install, deploy, manage and invoke the governance capabilities. Listing 6, shows examples of operational governance APIs exposed by our framework. In general, the operational governance processes are defined as a sequence of such API calls, performed by the IoT cloud applications.

Listing 6 Examples of operational governance APIs

```
/* General case of capability invocation. */
/deviceId/{capabilityId}/{methodName}/{arguments}?
arg1={first-argument}&arg2={second-argument}&...

/* Data points capability invocation example. */
deviceId/DPcapa/setPollRate/arguments&rate=5s
/deviceId/DPcapa/list

/* Capabilities manager examples. */
/deviceId/cManager/capabilities/list
/deviceId/cManager/{capabilityId}/stop
```

4.4.2 Intents API operators

Intent is a passive data structure. Therefore, we need to provide developers with operators to work with the *Intent*s. These operators encapsulate mechanisms to *select, instantiate and execute* tasks, based on the input *Intent*. Consequently, instead of dealing with the individual tasks, a developer is presented with a unified interface to communicate with the runtime systems. Listing 7 shows the core API operators that support working with Intents and IntentScopes.

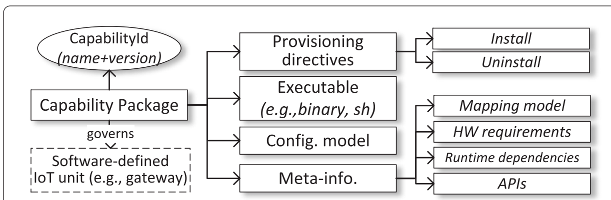


Fig. 7 Overview of capability package structure

Listing 7 Core Intent API operators

```
send(in ci:ControlIntent,out r:Result)
notify(in mi:MonitorIntent,in o:CallBackObj)
poll(in mi:MonitorIntent,out el:List<Event>)
delimit(in s:IntentScope,in c:Cond,out so:IntentScope)
```

The `send` primitive is used to communicate and execute a ControlIntent. When the `send` operator is invoked the container first selects suitable tasks to execute the ControlIntent by using Intent's headers. The task list is further filtered, based on Intent's attributes, e.g., quality requirements. Here, we use best-effort to find the best matching task implementation. Further, the selected task is configured with Intent's configuration parameters and a payload, and finally executed.

The core operators `notify` and `poll` are used to support working with the MonitorIntents. The operator `notify` is used by an application to subscribe for events, which are asynchronously delivered to the application. The `poll` operator is used to synchronously check the status of the environment, i.e., it will block application's main thread if the required event is currently unavailable.

The `delimit` operator is API equivalent of $\subseteq_{cond}$, defined in Section 4.2. It is used to define an IntentScope with entities, which satisfy a certain condition. Usually, when an application wants to determine the IntentScope, it will start by invoking `delimit` on the GlobalScope and further refine it by recursively applying this operator and/or using other scope operators.

5 SDG-Pro runtime mechanisms

In the current implementation of the SDG-Pro framework we provide a set of runtime mechanisms that underpin the programming abstractions (Section 4) and support application execution atop the IoT cloud. Generally, application execution includes instantiating, provisioning and deploying software-defined gateways; dynamically loading device services atop the gateways; instantiating virtual application topologies (with Intents and IntentScopes); and triggering execution of operational governance processes (on-demand, depending on the business logic). Next, we discuss the design and implementation of the most important SDG-Pro's runtime mechanisms in more detail.

5.1 Instantiating, provisioning and deploying software-defined gateways

Currently SDG-Pro supporta a version of software-defined gateways (prototypes), which is based on Linux Containers (LXC). When a developer instantiates a gateway prototype (e.g., as shown in Listing 4), the *IoT units controller* (cf. Fig. 3) performs three main tasks. First, it creates an instance of LXC and installs the *provisioning* *and governance agents* in the container. Second, the provisioning agent[1] executes the provisioning directives, supplied in a provisioning script (e.g., Listing 5). Finally, the *IoT units controller* deploys the gateway instance in IoT cloud.

Firstly, to instantiate a software-defined gateway our framework relies on Docker[2], i.e., more specifically on Docker deamon that offers a remote API for programmatic container creation. To bootstrap the instantiation, SDG-Pro provides a custom base image, which we developed atop a BusyBox user land on a stripped-down Linux distribution. In SDG-Pro, the *DeviceManager* is based on the Docker remote API, but it provides additional support for configuring and managing containers such as specifying the custom meta information (e.g., location) to provide more control over the software-defined gateways at the application level. As the last part of gateway instantiation, SDG-Pro installs its provisioning and governance agents that support execution of the subsequent phases.

Provisioning a software-defined gateway includes configuring, deploying and installing different artifacts such as device services, libraries (e.g., cloud communication protocols) and other binaries atop the newly created gateway instance. In the first step of the provisioning process, the *ProvisioningManager* creates artifacts image. In essence, it is a (compressed) set of component binaries and provisioning scripts. Next, the *DeploymentManager* places the image in the update queue. The provisioning agent periodically inspects the queue for new updates and when it is available the agent polls the image in the gateway (container) in a lazy manner. Additionally, SDG-Pro allows the components to be asynchronously pushed to the gateways, similarly to eager object initialization. Finally, the agent interprets provisioning scripts, i.e., performs a local installation of the binaries and executes any custom configuration directives.

Lastly, the SDG-Pro framework selects an IoT cloud node, i.e., an edge device or a cloud VM, and deploys the gateway instance on it. The main component responsible for gateways (containers) allocation and deployment is the *GatewayCoordinator*. In the current prototype, the GatewayCoordinator is built based on `fleet` and `etcd`. The `fleet` is a distributed init system that provides the functionality to manage a cluster of host devices, e.g., the IoT cloud nodes. The `etcd` is a distributed key/value store that supports managing shared configurations among such nodes. In order to allocate a gateway, i.e., select the best matching node in the IoT cloud, the GatewayCoordinator compares the available gateway attributes (e.g., location, ownership, node type, etc.) with the meta data of the available IoT cloud nodes. The gateway's meta data is obtained from a developer-specified configuration model. The nodes' meta data is provided by the *DeviceManager* and it is mostly maintained manually, e.g., by system administrators. At the moment, we only provide a rudimentary support for gateway allocation, i.e., SDG-Pro only considers static node properties. In the future, we plan to address this issue by including support for dynamic properties such as available bandwidth and providing support for runtime migration (reallocation) of software-defined gateways. Finally, after a node is selected, the GatewayCoordinator invokes the `fleet` to deploy the gateway on that node.

5.2 Intent-based invocation and IntentScope resolution

In the SDG-Pro framework, the communication among the main application components is performed via Intents. Generally, it follows a partial content-based publish/subscribe model and in the current prototype it is based on the Apache ActiveMQ JMS broker.

When an application submits a new Intent to SDG-Pro's *RuntimeContainer*, it first routes the Intent to the *TaskSelector*, which matches intent headers with device services (task) filters to find suitable services that match the Intent. Afterwards, the TaskSelector reads the Intent attributes and compares them with the task filters to find the best matching task. The attributes are represented as feature vectors and a multi-dimensional utility function, based on the Hamming distance, is used to perform the matching. Afterwards, the TaskSelector requests a service instance, by providing its description to the the TaskManager. It checks the validity of the mapping and, if it is valid, invokes the corresponding service. If no service is available the Intent is marked as failed and the invoker is notified.

In a more general case, when an Intent gets invoked on an IntentScope, the aforementioned invocation process remains the same, with the only difference that our framework performs all steps on a complete IntentScope, in parallel, instead on an individual gateway. To this end, the *ScopeCoordinator* provides dynamic resolution of the IntentScopes.

The IntentSope specifications are implemented as composite predicates which reference device meta information and profile attributes. The predicates are applied on the GlobalScope (Section 4.2), filtering out all resources that do not match the provided attribute conditions. The ScopeCoordinator uses the resulting set of resources to initiate the Intent mapping and invocation. The ScopeCoordinator is also responsible to provide support for gathering results delivered by the invoked device services. This is needed since the scopes are resolved in parallel and the results are asynchronously delivered by the software-defined gateways.

5.3 Invocation of runtime governance capabilities

As shown in Section 4.4, application developers define operational governance logic as a sequence of API calls to the governance capabilities. The *APIManager* is responsible to mediate (map) these invocations to the underlying infrastructure, i.e., the software-defined gateways. To this end it relies on the *CapabilityManager*, which is a cloud-based service and the governance agent, which is a lightweight HTTP deamon, preinstalled in software-defined gateway prototypes.

When an API request is submitted by an application, SDG-Pro performs following steps: it registers the capability, maps the API call, executes the capability, and returns the result. First, the APIManager registers the API call with the corresponding capability. This involves querying the capability repository to obtain its meta-information (such as expected arguments), as well as building a dynamic mapping model, which includes capability ID, a reference to a runtime environment (e.g., Linux shell), input parameters, the result type, and further configuration directives. The CapabilityManager forwards the model to the gateways (i.e. the governance agent) and caches this information for subsequent invocations. During future interactions, the framework acts as transparent proxy, since subsequent steps are handled by the underlying gateways. In the next step, the governance agent needs to perform a mapping between the API call and the underlying capability. By default, it assumes that capabilities follow the traditional Unix interaction model, i.e., that all arguments and configurations (e.g., flags) are provided via the standard input stream (stdin) and output is produced to standard output (stdout) or standard error (stderr) streams. This means, if not specified otherwise in the mapping model, the framework will try to invoke the capability by its ID and will forward the provided arguments to its stdin. For capabilities that require custom invocation, e.g., property files, policies, or specific environment settings, the framework requires a custom mapping model. This model is used in the subsequent steps to correctly perform the API call. Finally, the governance agent invokes the governance capability and as soon as the capability completes it collects and wraps the result. Currently, the framework provides means to wrap results as JSON objects for standard data types and it relies on the mapping model to determine the appropriate return type. However, this can be extended to support generic behavior, e.g., with Google Protocol Buffers.

6 Evaluation

6.1 Evaluation methodology

In this section we present a functional evaluation of the paper's main contribution – the SDG-Pro's programming model for IoT cloud applications. To validate SDG-Pro's programming model we follow evaluation design guidelines provided in [12]. The main objective of our qualitative analysis is twofold. First, to show that SDG-pro facilitates dealing with the challenges of designing and developing IoT cloud applications (RC1-RC4), we demonstrate how our programming model enforces the main design principles of IoT cloud systems, as justified in Section 3.1. Second, in order to show that SDG-Pro enables easier, efficient and more intuitive development of IoT cloud applications, we compare it against traditional programming model evaluation criteria that include: *readability, code simplicity, reusability, expressiveness* and *functional extensibility*.

Although analysis of non-functional properties of the runtime mechanisms (e.g., regarding system's scale) is not the main focus of this paper, we refer interested reader to our previous work (e.g., [10, 13]), where we partly show their performance.

6.2 Examples of FMS applications and services

To demonstrate the most important concepts and features of SDG-Pro's programming model, we present a set of real-life applications from our FMS system (Section 2). This example suite is designed to cover typical interactions and requirements of IoT cloud applications, such as realtime monitoring and data analytics, remote actuation and control, autonomous device tasks and offline data analytics, in order to show the *completeness* of SDG-Pro's programming model regarding its support w.r.t. the real-life requirements. The example applications are developed and deployed atop a virtualized IoT cloud testbed, based on CoreOS. In our testbed we simulate and mimic physical gateways in the cloud. The gateways are based on a snapshot of a real-world gateway, developed by our industry partners. The testbed is deployed on our local OpenStack cloud and it consists of 7 CoreOS 444.4.0 VMs, each running 150 LXCs, thus simulating approximately 1000 vehicles.

6.2.1 Example 1 – Energy consumption tracking

The FMS needs to monitor high-value vehicles' energy consumption in (near) real-time. In case any energy fault is detected, it must notify a golf course manager and put the vehicles in a reduced energy mode.

Listing 8 Remote monitoring of fleet's energy consumption

```
//select high-value vehicles
IntentScope s =
cont.delimit(IntentScope.getGlobal(),
Cond.greaterThan("price", "5000"));
Intent eFault = Intent.newMIntent("EnergyFault");
eFault.setScope(s);
cont.notify(eFault, this);//sub. to event
...
public void onEvent(Event e){
IntentScope ts = IntentScope.create(e.getEntityId();
Intent eCons = Intent.newCIntent("ReduceEnergy");
eCons.setScope(ts);//set task scope
cont.send(eCons); //send to all vehicles in ts
}
```

The most important part of this application in shown in Listing 8. To implement the monitoring behavior, developers only need to define an IntentScope (lines 2–4), in which they declare properties (e.g., metadata) that need to be satisfied by monitored vehicles, define a MonitorIntent and assign the desired scope to it (lines 5–7). Similarly, to implement a remote control behavior developers only need to define a ControlIntent (lines 12–14). In this example, it is natural to use asynchronous communication (of sensory data), thus a developer uses SDG-Pro's `notify` directive (line 8), to subscribe for the state changes in the environment.

This example demonstrates how easy it is to implement a real-time remote monitoring behavior. By introducing IntentSopes at the application level, SDG-Pro shields the developers from directly referencing the vast number of diverse physical entities and enables them to delimit the range of their actions on a higher abstraction level. Similarly, to perform an IoT control action or to subscribe for relevant events, developers only need to define and configure the corresponding Intents. This allows them to communicate to the system what needs to be done, instead of worrying how the underlying devices will perform the specific task.

6.2.2 *Example 2 – Scheduled maintenance check*

The FMS performs daily checks of the fleet's health. This is done mainly during the night, when the vehicles reside dormant in the club house, within the Wi-Fi range. The application reads the diagnostic data, gathered during the day, and analyzes them offline.

Listing 9 Scheduled maintenance check

```
Intent localCon = Intent.newMIntent("ConnType");
 localCon.setScope(IntentScope.getGlobal());
 IntentScope ds = container.
 delimit(Cond.eq("WLAN", localCon));
 ds.addObserver(this);

public void update(Observable obs, Object arg){
 Intent di = Intent.newMIntent("DiagnosticsLogger");
 di.setScope((IntentScope)obs);
 List<Event> data = container.poll(diagnostics);
 // send data to an analytics framework
}
```

To implement such behavior, application first needs to determine that a vehicle is connected to a local network. This is achieved by defining an active IntentScope, as shown in Listing 9, lines 1–4. Second, the application needs to gather vehicles' diagnostic data and store them, e.g., in a local database. To synchronously poll the vehicle data, a developer simply defines a MonotorIntent and uses the `poll` directive (lines 8–10).

This example, demonstrates several important points. First, since MonitorIntents can be used to define an IntentScope, SDG-Pro enables developers to dynamically (e.g., based on environment or context changes) determine application behavior. Second, since IntentScopes are observable, developers can specify complex conditions that will trigger an execution of the business logic, without having to write complicated queries and event processing schemes.

Finally, it is worth noticing that SDG-Pro does not provide support for the data analytics. However, we have shown that with little effort, by using intuitive concepts, an offline analytics application can obtain the required data, which can then be analyzed with data analytics frameworks, e.g., MapReduce.

6.2.3 *Example 3 – Diagnostics data logging*

This application periodically pools the data from the variety of vehicle's sensors e.g., engine status, battery status, transmission, etc. and stores them locally for later analysis (e.g., see Example 2).

Listing 10 Logging diagnostics data locally

```
//Create custom sensor from physical channel
BufferConf bc = new BufferConf("voltage_{i}n");
bc.setClass(BufferClass.SENSOR);
bc.getAdapterChain().add(
 new ScalingAdapter(0.0,100.0,10.0));
 bc.getAdapterChain().add(new LowpassFilter(0.30));
 BufferManager.create("lowpass-scaled", bc);
//Define diagnostics model
DataPoint diagnostics =
new ComplexDataPoint("lowpass-scaled","voltage_{i}n");
 DataInstance di = diagnostics.read();
//log the diagnostics data di
```

Listing 10 shows a partial diagnostics data model. The diagnostic data contains raw engine voltage readings and scaled voltage readings with low-pass filter, e.g., possibly indicating that something is taking the power away from the motor. To develop a custom sensor, developers only need to create a virtual buffer (referencing the base channel, e.g., raw voltage readings) and configure its adapter chain, as shown in lines 2–6. After creating a custom virtual sensor (line 7) application can treat this sensor as any other sensor. Consequently, a data model can then be easily defined with Data Points, as sown in lines 9–11. Storing the data is omitted for readability purposes.

Essentially, this example shows how our framework transparently virtualizes access to the same voltage sensor. This demonstrates two important features of the data and control points. First, since the SDG-Pro provides (virtually) exclusive access to the sensors (i.e., buffers act as multiplexers), developers can define custom configurations for the data streams, effectively creating an application-specific view of the sensors. An important consequence is that multiple applications can easily share the infrastructure, retaining a custom view of it. Second, since Data and

Control points support developers to interact with underlying devices in a unified manner, i.e., independent of the communication protocols or the input channel types, applications can define their (arbitrarily complex) data models by only specifying the required data points. These can be seen as volatile fields in traditional data model entities.

6.2.4 Example 4 – Energy fault detection

To detect vehicles over consuming battery an FMS service relies on powermeter, odometer and temperature sensors that are available in the vehicles and uses a custom algorithm to detect potential energy faults.

Listing 11 Device service for energy fault detection

```
DataPoint dp1 = DataPoint.
 create( "battery", "odometer" );
 DataPoint dp2 = new DataPoint();
//Since we have multiple temperature sensors
//we query them via the meta data
Collection<BufferDescription> tempBuffers =
 dataPoint.queryBuffers(
 new SensorProps("*temperature*"));
 dp2.assign(tempBuffers);
...
 //invoke energy fault detection algorithm
```

In Listing 11 we show a code snippet from the corresponding FMS service. Developers create two data points. The `dp1` combines the battery status and odometer readings and it asynchronously delivers the sensory readings to the service. The `dp2` queries the available temperature data channels, based on their meta data and aggregates the temperature readings from the available thermometers (lines 6–9). Among other things, the energy fault detection algorithm uses these data points and Complex Event Processing (CEP) techniques to determine potential energy faults, but its implementation is omitted in accordance with our nondisclosure agreement.

We notice that application obtains the temperature readings without directly referencing any physical sensor. Instead it generically queries the sensors' meta data. Further, since SDG-Pro takes care of synchronizing the sensors' readings, e.g., among the temperature sensors, developers can focus on custom data processing steps (algorithm). This is a crucial requirement to be able to develop portable applications, which do not directly depend on the physical infrastructure.

6.2.5 Example 5 – Provisioning and deploying application runtime environment

In order to execute an application/service (see Examples 1–4), developers need to provision a software-defined gateway and deploy it atop IoT cloud.

Listing 12 Creating a software-defined gateway

```
/* Snippet from Provisioning.java*/
//install JVM Compact Profile 1
String dest = ".../G2021/jvm";
provisioner.CreateDirIfMissing(dest);
provisioner.CopyToDir("jvm-profile1-1.8.0/*",dest);
provisioner.setPermissions(dest, "a+x");
...
/* Snippet from Gateways.java*/
SDGateway gateway
 = UnitsController.create(GType.BUSYBOX);
 gateway.addConfigClass(Provisioning.class);
 UnitsController.startParallel(gateway,
 IntentScope.getGlobal().asResource());
```

Listing 12, shows how to programmatically add Java Compact Profile runtime to a gateway and how to deploy instances of that gateway atop the vehicles' onboard devices. In lines 3–6 we show how developers can use the provisioning API to specify which custom resources are required in the gateway prototype. Further, this example show the most important parts related to gateways deployment, i.e., gateway instantiation from Docker-based Busybox prototype (lines 9,10), associating the configuration model with the prototype (line 11) and multiple deployment (lines 12,13).

This example shows a part of general SDG-Pro's provisioning API. We notice that our framework provides a generic API which can be used to declaratively configure different types of resources. This essentially enables developers to programmatically deal with complex IoT cloud infrastructure and its dependencies, i.e., the desired configuration baseline is specified locally and once for multiple application instances. SDG-Pro provides a unified view on defining and manipulating the infrastructure through software-defined gateways, but also offers a fine-grained access and control of the gateways configuration (e.g., container's base image).

6.2.6 Example 6 – Configuring application dependencies programmatically

The FMS applications have different dependencies and requirements e.g., regarding communication protocols. To guarantee correct application behavior, developers (or operations managers) need to correctly configure such infrastructure dependencies.

Listing 13 Configuring application dependencies

```
//install Modbus
provisioner.addDCPointResource("modbus/Modbus.sab");
provisioner.addDCPointResource("modbus/Modbus.sax");
provisioner.addDCPointResource("modbus/Kits.scode");
provisioner.addDCPointResource("modbus/Kits.xml");
//install MQTT client
RemoteLibrary mqttClient
 = provisioner.getFromURL(
 "http://..../mqtt-client-0.0.1.jar");
provisioner.installComProto(mqttClient.getBinary());
...
```

Listing 13 shows excerpt of typical FMS protocols configuration. Lines 2–6 show how developers can to configure Modbus device protocol (used by Data and Control Points) and MQTT cloud connectivity protocol (lines 7–10), e.g., used by MonitorIntents.

The most important thing to notice here is that SDG-Pro provides software-defined gateway specific provisioning APIs. This shows that our abstractions are designed in such manner to inherently support programmatic provisioning, by exposing well-defined API and providing runtime mechanisms which transparently enable inversion of control and late (re)binding of the dependencies. Also standard provisioning operations such as fetching a remote resource can be combined with specific provisioning APIs, as shown in lines 7–9. The most important consequence is that developers can design generic application business logic and transparently declare the desired infrastructure dependencies programmatically, e.g., in a separate application module.

6.2.7 Example 7 – Emergency governance process

In case of an emergency situation the FMS needs to increase the monitoring frequency of vehicles' sensors.

Listing 14 Example emergency operational governance process

```
Iterator<Vehicle> vehicles.iterator();
//for each vehicle on the golf course
List<DataPoint> dPoints = HTTPClient
 .invoke(".../APIManager/mapper/"
 +vehicles.next().getId()+"/DPcapa/list");
 for (DataPoint dp : dPoints) {
 HTTPClient.invoke(".../DPcapa/"
 +"setPollRate/args?rate=10s&id="+dp.getId()");
}
```

To satisfy this cross-cutting compliance requirement, developers need to develop an operational governance process [9, 10]. Listing 14 shows a code snippet form such emergency governance process. The most important part of the process is shown in lines 7–8, which show how a developer can use governance API to dynamically manipulate the edge of the infrastructure, in this case change the sensor update rate.

SDG-Pro takes over the responsibility of invoking individual governance capabilities (e.g., per vehicle), effectively shielding the developers from low-level infrastructure details. The most important consequence of having such governance API is that the governance logic can be specified programmatically and maintained locally. Also governance processes are completely separated from the business logic, thus the core business logic is not polluted with cross-cutting governance concerns.

In addition, since at the application level the infrastructure is perceived as a set of capabilities exposed through the governance API, the developers do not have to worry about geographical distribution, heterogeneity or scale of the IoT cloud infrastructure nor directly deal with individual devices.

6.3 Discussion

As shown on a set of real-life examples, our SDG-Pro framework enables addressing most of development concerns at application code level (*everything as code*). This provides advantages such as having a uniform view on the entire development process, which makes it easily traceable and auditable, but also enables exploiting proven and well-known technologies, e.g., source control or configuration management systems, during the entire application lifecycle. Moreover, it gives full control to developers and makes IoT cloud applications less infrastructure-dependent.

We have shown how SDG-Pro provides *API encapsulation* of the most important aspects related to gateway provisioning and governance. A key advantage of this approach is that developers do not need to explicitly worry about the underlying infrastructure. Rather, they perceive the complex and heterogeneous IoT cloud infrastructure as a set of uniform APIs that enable programmatic management of such infrastructure. Our SDG-Pro framework supports the developers to *declaratively provision* IoT cloud systems and to *automate* most of the provisioning process. This improves general readability and maintainability of the provisioning logic and simplifies the provisioning process. Additionally, by encoding the provisioning directives as part of application's source code, our framework makes the provisioning process easily repeatable. This reduces the potential errors, but more importantly enables continuous, automated enforcement of the configuration base line.

Regarding the governance processes, by providing a logically *centralized point of operation* of IoT cloud infrastructure, SDG-Pro supports developers to easily define desired states and runtime behavior of IoT cloud systems, but also enables automated enforcement of governance processes, which is crucial to realize (time) consistent governance strategies across the entire IoT cloud system.

We also notice a number of limitations of our approach. From the technical perspective, at the moment SDG-Pro offers a rudimentary mechanism for gateway allocation, which only considers static properties when deploying the software-defined gateways. Additionally, although IoT cloud systems include many mobile and unstable devices, the current prototype provides a limited support regarding the dependability concerns. However, optimization of gateway allocation and addressing the dependability issues related to device mobility are subject of our future work.

Furthermore, the set of proposed programming concepts is not exhaustive and especially the provisioning and governance APIs are in an active state of development and refinement. However, as we have shown on a set of real-life examples, SDG-Pro offers programming support sufficient to express many common behaviors of IoT cloud applications. Although our programming model has many important traits such as readability and simplicity, as well as facilitates writing reusable and portable application logic, in SDG-Pro's programming model, we trade flexibility and expressiveness for more intuitive and efficient programming of the IoT cloud applications. Finally, although developers utilize the well-known Java programming language, SDG-Pro introduces a number of new concepts that require an initial learning effort. However, by explicitly enforcing main design principles of software-defined IoT cloud systems, we believe that in the long-run our framework can reduce development time, potential errors and eventually the costs of application development.

7 Related work

Developing and managing IoT cloud systems and applications have been receiving a lot of attention lately. In [1, 14, 15] the authors mostly deal with IoT infrastructure virtualization and its management on cloud platforms. A number of different approaches (e.g., [2, 16]) employ semantics aspects to enable discovering, linking and orchestrating heterogeneous IoT devices. In [3, 17] the authors propose utilizing cloud for additional computation resources and approaches presented in [4, 18] focus on utilizing cloud's storage resources for sensory data. Approaches presented in [5, 19] deal with integrating IoT devices and services with enterprise applications based on SOA paradigm.

In [14] the authors focus on developing a virtualized infrastructure to enable sensing and actuating as a service on the cloud. They propose a software stack that includes support for management of device identification and device services aggregation. Although, this approach facilitates development of IoT cloud applications to a certain extent, contrary to SDG-Pro it does not define a structured programming model for developing such applications. In [1] the authors introduce sensor-cloud infrastructure that virtualizes physical sensors on the cloud and provides management and monitoring mechanisms for the virtual sensors. However, the support for sensor provisioning is based on static templates that, contrary to our approach, do not support dynamic provisioning of IoT cloud resources.

SenaaS [16] mostly focuses on providing a cloud semantic overlay atop physical infrastructure. It defines an IoT ontology to mediate interaction with heterogeneous devices and data formats, exposing them as event streams to the upper layer cloud services. Similarly, the OpenIoT framework [2] focuses on supporting IoT service composition by following cloud/utility based paradigm. It mainly relies on semantic web technologies and CoAP to enable web of things and linked sensory data. Such approaches address very important issues such as discovering, linking and orchestrating internet connected objects and IoT services, thus conceptually complementing our approach. Although, SDG-Pro relies on semantic concepts, e.g., it implements hierarchical namespaces and a proprietary taxonomy for sensor interoperability to support Data and Control Points, the semantic aspects are not the main focus this work. The aforementioned approaches mainly focus on providing different virtualization, management and (semantic-based) interoperability techniques for IoT devices. Therefore, such approaches can be seen as complementary to our own, as device virtualization sets the cornerstone for achieving IoT cloud systems. The SDG-Pro framework relies on the contemporary advances in IoT cloud and extends them with novel programming abstractions which enable everything-as-code paradigm, facilitating development of IoT cloud applications and making the entire development process traceable and auditable (e.g., with source control systems), thus improving maintainability and reducing development costs.

Putting more focus on the edge devices, i.e., IoT gateways, network devices, cloudlets and small clouds, different approaches have emerged recently. For example, in [20] the authors present a concept of fog computing and define its main characteristics, such as location awareness, reduced latency and general QoS improvements. They focus on defining a virtualized platform that includes the edge devices and enables running custom application logic atop different resources throughout the network. Similarly, Cloudlets [6] and small clouds [21] are introduced as intermediary infrastructure nodes (between the edge devices and data centers), which can be used to reduce network delays, processing time and costs. The SDG-Pro framework also aims at better utilization of the edge infrastructure, but we focus on providing a systematic approach, supporting application developers to address most of the infrastructure provisioning and governance issues programmatically, in a logically centralized fashion, by offering the software-defined gateways and well-defined provisioning and governance APIs.

Different approaches have exploited and extend software defined concepts to facilitate utilization and management of the pooled sets of shared IoT cloud resources, e.g., software-defined storage [22] and software-defined data center [23]. Also recent advances in more traditional software-defined networking (SDN) [24] have enabled easier management and programming of the intermediate network resources, e.g., routers. However, SDN mostly focuses on defining the networking logic, e.g., injecting routing rules into network elements. Conversely, our SDG-Pro addresses the more general problem of providing programming support for a general business logic of IoT cloud applications. It builds on our previous concepts to provide programming abstractions, which enforce earlier identified design principles of software-defined IoT cloud, in order to enable scalable, efficient and more intuitive application development.

Another related field is macroprogramming of sensor networks [25–28]. For example, in [25] the authors provide an SQL-like interface where the entire network is abstracted as a relational database (table). Contrary to their approach, we utilize more general set theory to define operations on our IntentScopes. This gives more flexibility to developers, since SDG-Pro also allows dynamic, custom properties to be included in scope definitions, but comes at the cost of additional performance overhead.

Similarly, in [27], the authors deal with enabling dynamic scopes in WSN, mainly addressing the important issues of task placement and data exchange (among the WSN nodes), in order to account for the heterogeneity of the nodes and enable logically localized interactions. Their approach can be seen as conceptually complementing the SDG-Pro, since task allocation and such interaction types are not the main focus of our framework. In [26], the authors propose the notion of logical neighborhood. Their approach is based on logical nodes (templates), which enable instantiating and grouping the nodes, based on their exported attributes. To facilitate communication within the neighborhoods, which is of greater importance in WSN, they also provide an efficient routing mechanism. In [28] the authors introduce an extensible programming framework that unifies the WSN programming abstractions in order to facilitate business processes orchestration with WSN. Despite the relevant efforts to integrate provisioning and business logic (e.g., template-based customizations [26]), the main focus of the aforementioned approaches is application business logic, while we address a more general problem of enabling everything-as-code paradigm, in order to also allow for capturing provisioning and governance logic for IoT cloud resources programmatically.

8 Conclusion and future work

In this paper we introduced the SDG-Pro framework for software-defined IoT cloud systems. We presented SDG-Pro's programming model for IoT cloud applications, which is designed to enforce the main principles of software-defined IoT cloud systems that were elicited in our previous research in this area [7, 9, 10]. By enforcing such principles on the application level, our framework enables easier, efficient and more intuitive application development. It provides a unified programmatic view on the entire development process (*everything as code*) making it easily traceable and auditable, thus reducing development time, errors and costs of application development.

In the future, we will continue the development of the SDG-Pro framework to address its current limitations, i.e., improve the gateways allocation mechanism to include support for dynamic infrastructure properties. We also plan to address the current limitations regarding the mobility aspects, especially the dependability issues related to the device mobility and mobility of software components, i.e., runtime migration of software-defined gateways. Finally, we plan to extend the current prototype to support elastic, on-demand scaling of the software-defined gateways.

Endnotes

[1] The provisioning agent is implemented as a light-weight service, based on Oracle Compact Profile1 JVM.

[2] https://docker.com/.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
SN is the main contributor of this work, which was undertaken as part of his Ph.D. studies. SN drafted most of the manuscript and carried out most of the framework design and development. HLT is a co-supervisor of SN, with contributions to conceptualizing software-defined gateways and programming in IoT cloud systems. SD is the supervisor of SN, with main contributions to conceptualization of software-defined IoT cloud systems. All authors read and approved the final manuscript.

Acknowledgements
This work is partially supported by the EU H2020 U-Test project, under grant No. 645463.

Received: 30 January 2015 Accepted: 1 October 2015
Published online: 30 October 2015

References
1. Yuriyama M, Kushida T (2010) Sensor-cloud infrastructure-physical sensor management with virtualized sensors on cloud computing. In: NBiS'10. pp 1–8. doi:10.1109/NBiS.2010.32
2. Soldatos J, Serrano M, Hauswirth M (2012) Convergence of utility computing with the internet-of-things. In: IMIS. pp 874–9. doi:10.1109/IMIS.2012.135
3. Chun BG, Ihm S, Maniatis P, Naik M, Patti A (2011) Clonecloud: elastic execution between mobile device and cloud. In: Proceedings of the Sixth Conference on Computer Systems. ACM, New York, NY, USA. pp 301–314. http://doi.acm.org/10.1145/1966445.1966473
4. Studei P, Mohomed I, Terry D (2010) Wherestore: Location-based data storage for mobile devices interacting with the cloud. In: Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond. ACM, New York, NY, USA. pp 1:1–1:8. http://doi.acm.org/10.1145/1810931.1810932
5. De Souza LMS, Spiess P, Guinard D, Köhler M, Karnouskos S, Savio D (2008) Socrades: A web service based shop floor integration infrastructure. In: The Internet of Things. pp 50–67. http://link.springer.com/chapter/10.1007%2F978-3-540-78731-0_4
6. Satyanarayanan M, Bahl P, Caceres R, Davies N (2009) The case for vm-based cloudlets in mobile computing. Pervasive Comput 8(4):14–23. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5280678

7. Nastic S, Sehic S, Le DH, Truong HL, Dustdar S (2014) Provisioning Software-defined IoT Cloud Systems. In: FiCloud'14. IEEE, Barcelona, Spain. pp 288–295. doi:10.1109/FiCloud.2014.52
8. Nastic S, Sehic S, Voegler M, Truong HL, Dustdar S (2013) PatRICIA - A novel programing model for IoT applications on cloud platforms. In: SOCA. IEEE Computer Society, Koloa, HI, USA. doi:10.1109/SOCA.2013.48
9. Nastic S, Voegler M, Inziger C, Truong HL, Dustdar S (2015) rtGovOps: A Runtime Framework for Governance in Large-scale Software-defined IoT Cloud Systems. In: Mobile Cloud 2015. IEEE, San Francisco, CA, USA. pp 24–33. doi:10.1109/MobileCloud.2015.38
10. Nastic S, Inziger C, Truong HL, Dustdar S (2014) GovOps: The Missing Link for Governance in Software-defined IoT Cloud Systems. In: WESOA14. Springer, Paris, France Vol. 8954. pp 20–31. doi:10.1007/978-3-319-22885-3_3
11. Inzinger C, Nastic S, Sehic S, Vögler M, Li F, Dustdar S (2014) MADCAT - A methodology for architecture and deployment of cloud application topologies. In: Service-Oriented System Engineering. IEEE Computer Society, Oxford, United Kingdom. pp 13–22. doi:10.1109/SOSE.2014.9
12. Mohagheghi P, Haugen Ø (2010) Evaluating domain-specific modelling solutions. In: Advances in Conceptual Modeling - Applications and Challenges. pp 212–21. http://link.springer.com/chapter/10.1007%2F978-3-642-16385-2_27
13. Voegler M, Schleicher JM, Inziger C, Nastic S, Sehic S, Dustdar S (2015) LEONORE – Large-scale provisioning of resource constrained IoT deployments. In: SOSE. IEEE, San Francisco Bay, CA, USA. pp 78–87. doi:10.1109/SOSE.2015.23
14. Distefano S, Merlino G, Puliafito A (2012) Sensing and actuation as a service: a new development for clouds. In: NCA. pp 272–275. doi:10.1109/NCA.2012.38
15. Hassan MM, Song B, Huh EN (2009) A framework of sensor-cloud integration opportunities and challenges. In: ICUIMC. ACM, New York, NY, USA. pp 618–626. doi:10.1145/1516241.1516350
16. Alam S, Chowdhury M, Noll J (2010) Senaas: An event-driven sensor virtualization approach for internet of things cloud. In: NESEA. pp 1–6. doi:10.1109/NESEA.2010.5678060
17. Kumar K, Lu YH (2010) Cloud computing for mobile users: Can offloading computation save energy? Computer 43(4):51–6
18. Zaslavsky A, Perera C, Georgakopoulos D (2013) Sensing as a service and big data. arXiv preprint arXiv:1301.0159. http://arxiv.org/abs/1301.0159
19. Kovatsch M, Lanter M, Duquennoy S (2012) Actinium: A restful runtime container for scriptable internet of things applications. In: Internet of Things. pp 135–142. doi:10.1109/IOT.2012.6402315
20. Bonomi F, Milito R, Zhu J, Addepalli S (2012) Fog computing and its role in the Internet of Things. In: MCC Workshop on Mobile Cloud Computing. ACM, New York, NY, USA. pp 13–16. http://doi.acm.org/10.1145/2342509.2342513
21. Sixsq NuvlaBox. http://sixsq.com/products/nuvlabox.html. [Online; accessed Jan-'15]
22. Thereska E, Ballani H, O'Shea G, Karagiannis T, Rowstron A, Talpey T, et al. (2013) IoTFlow: A software-defined storage architecture. In: SOSP. ACM, Farmington, PA, USA. pp 182–96. http://doi.acm.org/10.1145/2517349.2522723
23. Davidson, Emily A (Softchoice Advisor): The Software-Defined-Data-Center (SDDC): Concept Or Reality? http://tinyurl.com/omhmbfv. [Online; accessed Jan-'15]
24. Koldehofe B, Dürr F, Tariq MA, Rothermel K (2012) The power of software-defined networking: line-rate content-based routing using OpenFlow. In: Proceedings of the 7th Workshop on Middleware for Next Generation Internet Computing. ACM, New York, NY, USA. pp 3:1–3:6. http://doi.acm.org/10.1145/2405178.2405181
25. Madden SR, Franklin MJ, Hellerstein JM, Hong W (2005) TinyDB: an acquisitional query processing system for sensor networks. ACM Trans Database Syst (TODS) 30(1):122–73
26. Ciciriello P, Mottola L, Picco GP (2006) Building virtual sensors and actuators over logical neighborhoods. In: Proceedings of the International Workshop on Middleware for Sensor Networks. ACM, New York, NY, USA. pp 19–24. http://doi.acm.org/10.1145/1176866.1176870
27. Mottola L, Pathak A, Bakshi A, Prasanna VK, Picco GP (2007) Enabling scope-based interactions in sensor network macroprogramming. In: MASS 2007. IEEE Computer Society, Pisa, Italy. pp 1–9. http://dx.doi.org/10.1109/MOBHOC.2007.4428655
28. Casati F, Daniel F, Dantchev G, Eriksson J, Finne N, Karnouskos S, et al. (2012) Towards business processes orchestrating the physical enterprise with wireless sensor networks. In: ICSE'12. IEEE, Zurich, Switzerland. pp 1357–1360. doi:10.1109/ICSE.2012.6227080

Provisioning Software-defined IoT Cloud Systems

Stefan Nastic, Sanjin Sehic, Duc-Hung Le, Hong-Linh Truong, and Schahram Dustdar
Distributed Systems Group, Vienna University of Technology, Austria
Email: $\{lastname\}$@*dsg.tuwien.ac.at*

***Abstract*—Cloud computing is ever stronger converging with the Internet of Things (IoT) offering novel techniques for IoT infrastructure virtualization and its management on the cloud. However, system designers and operations managers face numerous challenges to realize IoT cloud systems in practice, mainly due to the complexity involved with provisioning large-scale IoT cloud systems and diversity of their requirements in terms of IoT resources consumption, customization of IoT capabilities and runtime governance. In this paper, we introduce the concept of *software-defined IoT units* – a novel approach to IoT cloud computing that encapsulates fine-grained IoT resources and IoT capabilities in well-defined APIs in order to provide a unified view on accessing, configuring and operating IoT cloud systems. Our software-defined IoT units are the fundamental building blocks of software-defined IoT cloud systems. We present our framework for dynamic, on-demand provisioning and deploying such software-defined IoT cloud systems. By *automating provisioning* processes and supporting *managed configuration models*, our framework simplifies provisioning and enables flexible runtime customizations of software-defined IoT cloud systems. We demonstrate its advantages on a real-world IoT cloud system for managing electric fleet vehicles.**

I. Introduction

Cloud computing technologies have been intensively exploited in development and management of the large-scale IoT systems, e.g., in [11], [16], [18], because theoretically, cloud offers unlimited storage, compute and network capabilities to integrate diverse types of IoT devices and provide an elastic runtime infrastructure for IoT systems. Self-service, utility-oriented model of cloud computing can potentially offer fine-grained IoT resources in a pay-as-you-go manner, reducing upfront costs and possibly creating cross-domain application opportunities and enabling new business and usage models of the IoT cloud systems.

However, most of the contemporary approaches dealing with IoT cloud systems largely focus on data and device integration by utilizing cloud computing techniques to virtualize physical sensors and actuators. Although, there are approaches providing support for provisioning and management of the virtual IoT infrastructure (e.g, [8], [16], [18]), the convergence of IoT and cloud computing is still at an early stage. System designers and operations managers face numerous challenges to realize large-scale IoT cloud systems in practice, mainly because these systems impose diverse requirements in terms of granularity and flexibility of IoT resources consumption, custom provisioning of IoT capabilities such as communication protocols, elasticity concerns, and runtime governance. For example, modern large-scale IoT cloud systems heavily rely on the cloud and virtualized IoT resources and capabilities (e.g., to support complex, computationally expensive analytics), thus these resources need to be accessed, configured and operated in a unified manner, with a central point of management. Further, the IoT systems are envisioned to run continuously, but they can be elastically scaled in/down in off-peek times, e.g., when a demand for certain data sources reduces. Due to the multiplicity of the involved stakeholders with diverse requirements and business models, the modern IoT cloud systems increasingly need to support different and customizable usage experiences. Therefore, to utilize the benefits of cloud computing, IoT cloud systems need to support virtualization of IoT resources and IoT capabilities (e.g., gateways, sensors, data streams and communication protocols), but also enable: i) encapsulating them in a well-defined API, at different levels of abstraction, ii) centrally managing configuration models and automatically propagating them to the edge of infrastructure, iii) automated provisioning of IoT resources and IoT capabilities.

In this paper, we introduce the concept of software-defined IoT units – a novel approach to IoT cloud computing that encapsulates fine-grained IoT resources and IoT capabilities in a well-defined API in order to provide a unified view on accessing, configuring and operating IoT cloud systems. Our software-defined IoT units are the fundamental building blocks of software-defined IoT cloud systems. They enable consuming IoT resources at a fine granularity, allow for policy-based configuration of IoT capabilities and runtime operation of software-defined IoT cloud systems. We present our framework for dynamic, on-demand provisioning of the software-defined IoT cloud systems. By automating main aspect of provisioning processes and supporting centrally managed configuration models, our framework simplifies provisioning of such systems and enables flexible runtime customizations.

The rest of the paper is structured as follows: Section II presents a motivating scenario and research challenges; Section III describes main principles and our conceptual model of software-defined IoT systems; Section IV outlines main provisioning techniques for software-defined IoT systems; Section V introduces design and implementation of our prototype, followed by its experimental evaluation; Section VI discusses the related work; Finally, Section VII concludes the paper and gives an outlook of the future research.

II. Motivation

A. Scenario

Consider a scenario about fleet management (FM) system for small-wheel electric vehicles deployed worldwide, on different golf courses. The FM is an IoT cloud system comprising golf cars' on-board gateways, network and the cloud infrastructure. The main features provided by the on-board device include: a) vehicle maintenance (fault history, battery health, crash history, and engine diagnostics), b) vehicle tracking (position, driving history, and geo-fencing), c) vehicle info (charging status, odometer, serial number, and service notification), d) set-up (club-specific information, maps, and fleet information). Vehicles communicate with the cloud via 3G, GPRS or Wi-Fi network to exchange telematic and diagnostic data. On the cloud we host different FM subsystems and services to manage the data. For example: a) *Realtime vehicle status*: location, driving direction, speed, vehicle fault alarms; b) *Remote diagnostics*: equipment status, battery health and timely maintenance reminders; c) *Remote control*: overriding on-board vehicle control system in case of emergency; d) *Fleet management*: service history and fleet usage patterns.

In the following we highlight some of the requirements and features of the FM system that we need to support:

- The FM subsystems and services are hosted in the cloud and heavily rely on the virtualized IoT resources, e.g., vehicle gateways and their capabilities. Therefore, we need to enable encapsulating and accessing IoT resources and IoT capabilities, via an uniform API.
- The FM system has different requirements regarding communication protocols. The fault alarms and events need to be pushed to the services (e.g, via MQ Telemetry Transport (MQTT) [1]), when needed vehicle's diagnostics should be synchronously accessed via RESTfull protocols such as CoAP [10] or sMAP [7]. The remote control system requires a dedicated, secure point-to-point connection. Configuring these capabilities should be decoupled from the underlying physical infrastructure, in order to allow dynamic, fine-grained customization.
- The FM system spans multiple, geographically distributed cloud instances and IoT devices that comprise FM's virtual runtime topologies. These topologies abstract a portion of the IoT cloud infrastructure, e.g., needed by specific subsystem, thus they should support flexible configuring to allow for on-demand provisioning.

The limited support for fine-grained provisioning at higher levels leads to tightly coupled, problem specific IoT infrastructure components, which require difficult and tedious configuration management on multiple levels. This inherently makes provisioning and runtime operation of IoT cloud systems a complex task. Consequentially, system designers and operations managers face numerous challenges to provision and operate large-scale IoT cloud systems such as our FM.

B. Challenges

RC1 – The IoT cloud services and subsystems provide different functionality or analytics, but they mostly rely on common physical IoT infrastructure. However, to date the IoT infrastructure resources have been mostly provided as coarse grained, rigid packages, in the sense that the IoT systems, e.g., the infrastructure components and software libraries are specifically tailored for the problem at hand and do not allow for flexible customization and provisioning of the individual resource components or the runtime topologies. This inherently hinders self-service, *utility-oriented delivery and consumption* of the IoT resources at a finer granularity level.

RC2 – *Elasticity*, although one of the fundamental traits of the traditional cloud computing, has not yet received enough attention in IoT cloud systems. Elasticity is a principle to provision the required resources dynamically and on demand, enabling applications to respond to varying load patterns by adjusting the amount of provisioned resources to exactly match their current needs, thus minimizing resources over-provisioning and allowing for better utilization of the available resources [9]. However, IoT cloud systems are usually not tailored to incorporate elasticity aspects. For example, new types of resources, e.g., data streams, delivered by IoT infrastructure are still not provided elastically in IoT cloud systems. Opportunistic exploitation of constrained resources, inherent to many IoT cloud systems further intensifies the need to provision the required resources on-demand or as they become available. These challenges prevent current IoT systems from fully utilizing the benefits cloud's elastic nature has to offer and call for new approaches to incorporate the elasticity capabilities in the IoT cloud systems.

RC3 – Dependability is a general measure of dynamic system properties, such as availability, reliability, fault resilience and maintainability. Cloud computing supports developing and operating dependable large-scale systems atop commodity infrastructure, by offering an abundance of virtualized resources, providing replicated storage, enabling distributed computation with different availability zones and diverse, redundant network links among the system components. However, the challenges to build and *operate dependable large-scale IoT cloud systems* are significantly aggravated because in such systems the cloud, network and embedded devices are converging, thus creating very large-scale hyper-distributed systems, which impose new concerns that are inherently elusive with traditional operations approaches.

RC4 – Due to dynamicity, heterogeneity, geographical distribution and the sheer scale of IoT cloud, traditional management and provisioning approaches are hardly feasible in practice. This is mostly because they implicitly make assumptions such as physical on-site presence, manually logging into devices, understanding device's specifics, etc., which are difficult, if not impossible, to achieve in IoT cloud systems. Thus, novel techniques, which will provide an unified and conceptually centralized view on system's configuration management are needed.

Therefore, we need novel models and techniques to provision and operate the IoT cloud systems, at runtime. Some of the obvious requirements to make this feasible in the very large-scale, geographically distributed setup are: (i) We need tools which will automate development, provisioning and operations (DevOps) processes; (ii) Supporting mechanisms need to be late-bound and dynamically configurable, e.g., via policies; (iii) Configuration models need to be centrally managed and automatically propagated to the edge of the infrastructure.

III. Principles and Building Blocks of Software-Defined IoT Systems

A. Principles of Software-Defined IoT

Generally, software-defined denotes a principle of abstracting the low-level components, e.g., hardware, and enabling their provisioning and management through a well-defined API [14]. This enables refactoring the underlying infrastructure into finer-grained resource components whose functionality can be defined in software after they have been deployed.

Software-defined IoT systems comprise a set of resource components, hosted in the cloud, which can be provisioned and controlled at runtime. The IoT resources (e.g., sensory data streams), their runtime environments (e.g., gateways) and capabilities (e.g., communication protocols, analytics and data

TABLE I. SUMMARY OF MAIN PRINCIPLES AND ENABLERS OF SOFTWARE-DEFINED IOT SYSTEMS

Research challenges	High-level principles	Enablers
• Flexible customization	• API encapsulation of IoT resources and capabilities	• Software-defined IoT units
• Utility-oriented delivery and consumption	• Fine-grained resources consumption	• Software-defined IoT topology (complex units)
• Self-service usage model	• Policy-based specification and configuration	• Centrally managed configuration models and policies
• Support for elasticity concerns	• Automated provisioning	• Automated units composition
• Operating dependable large-scale IoT cloud systems	• Cost awareness	• Runtime unit control and modification
• Central point of management	• Runtime elasticity governance	

point controllers) are described as *software-defined IoT units*. *Software-defined IoT units* are software-defined entities that are hosted in an IoT cloud platform and abstract accessing and operating underlying IoT resources and lower level functionality. Generally, *software-defined IoT units* are used to encapsulate the IoT resources and lower level functionality in the IoT cloud and abstract their provisioning and governance, at runtime. To this end, our *software-defined IoT units* expose well-defined API and they can be composed at different levels, creating virtual runtime topologies on which we can deploy and execute IoT cloud systems such as our FM system. Therefore, main principles of software-defined IoT systems include:

- API Encapsulation – IoT resources and IoT capabilities are encapsulated in well-defined APIs, to provide a unified view on accessing functionality and configurations of IoT cloud systems.
- Fine-grained consumption – The IoT resources and capabilities need to be accessible at different granularity levels to support agile utilization and self-service consumption.
- Policy-based specification and configuration – The units are specified declaratively and their functionality is defined programmatically in software, using the well-defined API and available, familiar software libraries.
- Automated provisioning – Main provisioning processes need to be automated in order to enable dynamic, on-demand configuring and operating software-defined IoT systems, on a large-scale (e.g, hundreds gateways).
- Cost awareness – We need to be able to assign and control costs of delivered IoT resources and capabilities in order to enable their utility-oriented consumption.
- Elasticity support – They should support elasticity governance [9], by exposing runtime control of elastic capabilities through well-defined API.

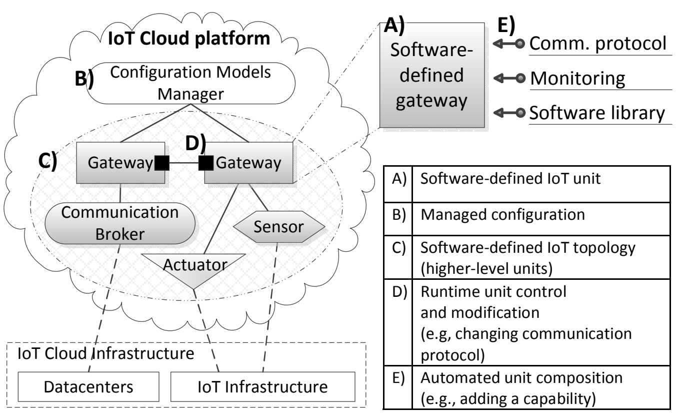


Fig. 1. Main enablers of software-defined IoT cloud systems

Table I summarizes how we translate the aforementioned high-level principles into concrete enablers. For example, to allow for flexible system customization, we need to enable fine-grained resource consumption, well-defined API encapsulation and provide support for policy-based specification and configuration. These principles are enabled by our software-defined IoT units and support for centrally managed configuration. Figure 1 gives high-level graphical overview of the main building blocks and enabling techniques, needed to support the main principles of software-defined IoT systems. Subsequently, we describe them in more detail.

B. Conceptual Model of Software-defined IoT Units

Figure 2 illustrates the conceptual model of our software-defined IoT units. The units encapsulate functional aspects (e.g., communication capabilities or sensor poll frequencies) and non-functional aspects (e.g., quality attributes, elasticity capabilities, costs and ownership information) of the IoT resources and expose them in the IoT cloud. The functional, provisioning and governance capabilities of the units are exposed via *well-defined APIs*, which enable provisioning and controlling the units at runtime, e.g., start/stop. Our conceptual model also allows for composing and interconnecting software-defined IoT units, in order to dynamically deliver the IoT resources and capabilities to the applications. The runtime provisioning and configuration is performed by specifying late-bound policies and configuration models. Naturally, the software-defined IoT units support mechanisms to map the virtual resources with the underlying physical infrastructure.

To technically realize our unit model we introduce a concept of *unit prototypes*. They can be seen as resource containers, which are used to bootstrap more complex, higher-level units. Generally, they are hosted in the cloud and enriched with functional, provisioning and governance capabilities, which are exposed via software-defined APIs. The unit prototypes can be based on OS-level virtualization, e.g., VMs, or more finer-grained kernel supported virtualization, e.g., Linux containers. Conceptually, virtualization choices do not pose any limitations, because by utilizing the well-defined API, our unit prototypes can be dynamically configured, provisioned, interconnected, deployed, and controlled at runtime.

Given our conceptual model (Figure 2), by utilizing the *provisioning API*, the unit prototypes can be dynamically coupled with late-bound runtime mechanisms. These can be any software components (custom or stock), libraries or clients that can be configured and whose binding with the unit prototypes is differed to the runtime. For example, the mechanisms can be used to dynamically add communication capabilities, new functionality or storage to our software-defined IoT units. Therefore, by specifying policies, which are bound later during runtime, system designers or operations managers can flexibly manage unit configurations and customize their capabilities, at *fine granularity levels*. Our conceptual model also allows for composing the software-defined IoT units at higher levels. By selecting dependency units, e.g., based on

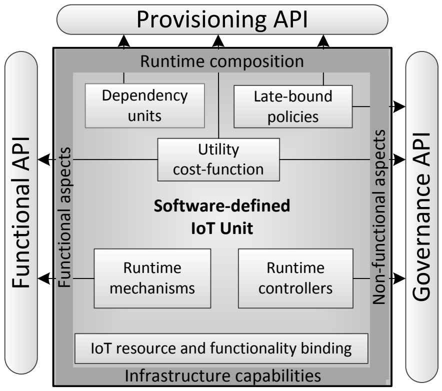


Fig. 2. Conceptual model of software-defined IoT units.

their costs, analytics or elasticity capabilities, and linking them together, we can dynamically build more complex units. This enables flexible *policy-based specification and configuration* of complex relationships between the units. Therefore, by carefully choosing the granularity of our units and providing configuration policies we can *automate the units composition process* at different levels and in some cases completely defer it to the runtime. This makes the provisioning process flexible, traceable and repeatable across different cloud instances and IoT infrastructures, thus reducing time, errors and costs.

The runtime *governance API*, exposed by the units, enables us to perform runtime control operations such as starting or stopping the unit or change the topological structure of the dependency units, e.g., dynamically adding or removing dependencies at runtime. Therefore, one of the most important consequences of having software-defined IoT unit is that the functionality of the virtual IoT infrastructure can be (re)defined and customized after it has been deployed. New features can be added to the units and the topological structure of the dependency units can be customized at runtime. This enables automating provisioning and governance processes, e.g., by utilizing the governance API and providing monitoring at unit level, we can enable *elastic horizontal scaling* of our units.

Therefore, most important features of software-defined IoT units which enable the general principles of software-defined IoT (see Section III-A) are:

- They provide software-defined API, which can be used to access, configure and control the units, in a unified manner.
- They support fine-grained internal configurations, e.g, adding functional capabilities like different communication protocols, at runtime.
- They can be composed at higher-level, via dependency units, creating virtual topologies that can be (re)configured at runtime.
- They enable decoupled and managed configuration (via late-bound policies) to provision the units dynamically and on-demand.
- They have utility cost-functions that enable pricing the IoT resources as utilities.

C. Units Classification

Depending on their purpose and capabilities, our software-defined IoT units have different granularity and internal topological structure. Therefore, conceptually we classify them into: (i) *atomic*, (ii) *composed* and (iii) *complex software-defined IoT units*. Depending on their type, the units require specific runtime mechanisms and expose specific provisioning API. Next we describe each unit type in more detail.

The *atomic software-defined IoT units* are the finest-grained software-defined IoT units, which are used to abstract the core capabilities of an IoT resource. They provide software-defined API and need to be packaged portably to include components and libraries, that are needed to provide desired capabilities. Figure 3 depicts some examples of the atomic software-defined units. We broadly classify them into functional and non-functional atomic software-defined IoT units, based on the capabilities they provide. Functional units encapsulate capabilities such as communication or IoT compute and storage. Non-functional units encapsulate configuration models and capabilities such as elasticity controllers or data-quality enforcement mechanisms. Therefore, the atomic units are used to identify fine-grained capabilities needed by an application. For example, the application might require the communication to be performed via a specific transport protocol, e.g., MQTT or it might need a specific monitoring component, e.g., Ganglia[1]. Classifications similar to the one presented in Figure 3 can be used to guide the atomic units selection process, in order to easily identify the exact capabilities, needed by the application.

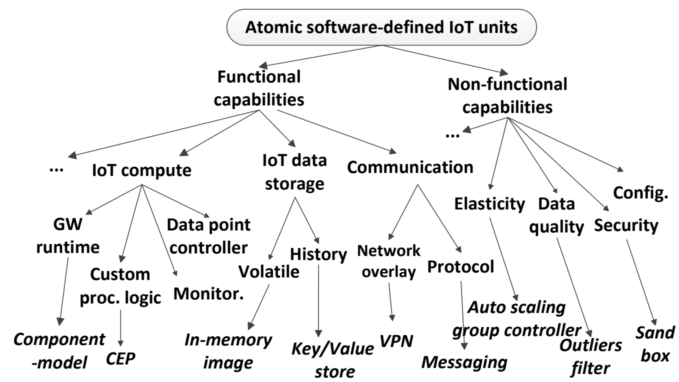


Fig. 3. Example classification of atomic software-defined IoT units.

The *composed software-defined IoT units* have multiple functional and non-functional capabilities, i.e., they are composed of multiple atomic units. Similarly to the atomic units they provide well-defined API, but require additional functionality such as mechanisms to support declaratively composing and binding the atomic units, at runtime (Section IV-B). Example of composed unit is a software-defined IoT gateway.

The *complex software-defined IoT units* enable capturing complex relationships among the finer-grained units. Internally, they are represented as a topological network, which can be configured and deployed, e.g., on the cloud. They define an API and can integrate (standalone) runtime controllers to dynamically (re)configure the internal topology, e.g., to enable elastic horizontal scaling of the units. Finally, they rely on runtime mechanism to manage the references, e.g., IP addresses and ports, among the dependency units.

We notice that the software-defined API and our units offer different advantages to the stakeholders involved into

[1] http://ganglia.info/

designing, provisioning and governing of software-defined IoT systems. For example, IoT infrastructure providers can offer their resources at fine-granularity, on-demand. This enables specifying flexible pricing and cost models and allows for offering the IoT resources as elastic utilities in a pay-as-you-go manner. Because our units support *dynamic and automated composition* on multiple levels, consumers of IoT cloud resources can provision the units to exactly match their functional and non-functional requirements, while still taking advantage of the existing systems and libraries. Further, system designers and operations managers, use late-bound policies to specify and configure the unit's capabilities. Because we treat the functional and configuration units in a similar manner (see Section IV-B), configuration models can be stored, reused, modified at runtime and even shared among different stakeholders. This means that we can support *managed configuration models*, which can be centrally maintained via configuration management solutions for IoT cloud, e.g., based on OpsCode Chef[2], Bosh[3] or Puppet[4].

IV. Provisioning software-defined IoT cloud systems

A. Automated composition of software-defined IoT units

Generally, building and deploying software-defined IoT cloud systems includes creating and/or selecting suitable software-defined IoT units, configuring and composing more complex units and building custom business logic components. The deployment phase includes deploying the software-defined IoT units together with their dependency units and required (possibly standalone) runtime mechanisms (e.g., a message broker). In this paper we mostly focus on provisioning reusable stock components such as gateway runtime environments or available communication protocols. Developing custom business logic components is out of scope of this paper and we address it elsewhere [15].

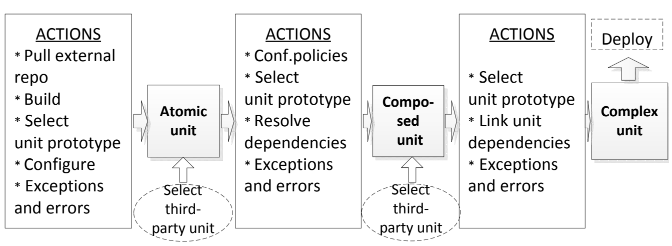


Fig. 4. Automated composition of software-defined IoT units.

Figure 4 illustrates most important steps to compose and deploy our IoT units. There are three levels of configuration that can be performed: (i) Building/selecting atomic units; (ii) Configuring composed units; (iii) Linking into complex units. Each of the phases includes selecting and provisioning suitable unit prototypes. For example, the unit prototypes can be based on different resource containers such as VMs, Linux Containers (e.g., Docker) or OSGi runtime.

The atomic units are usually provided as stock components, e.g., by a third-party, possibly in a market-like fashion. Therefore, this phase usually involves selecting and configuring stock components (e.g., Sedona[5] or Niagara^{AX}[6] execution environments). Classifications similar to the one presented in Figure 3 can be used to guide the atomic units selection process. In case we want to perform custom builds of the existing libraries and frameworks, there are many established build tools which can be used, e.g., for Java-based components, Apache Ant or Maven.

On the second level, we configure the composed units, e.g., a software-defined IoT gateway. This is performed by adding the atomic units (e.g., runtime mechanisms and/or software libraries) to the composed unit. For example, we might want to enable the gateway to communicate over a specific transport protocol, e.g., MQTT and add a monitoring component to it, e.g., a Ganglia agent. To perform this composition seamlessly at runtime, additional mechanisms are required. We describe them in Section IV-B.

Third level includes defining the dependencies references between the composed units, which "glue together" the complex units. These links specify the topological structure of the desired complex units. For example, to this end we can set up a virtual private network and provide each unit with a list of IP addresses of the dependency units. In this phase, we can use frameworks (e.g., TOSCA-based, OpenStack Heat, Amazon CloudFormation, etc.) to specify the runtime topological structure of our units and utilize mechanisms (e.g., Ubuntu CloudInit[7]) to bootstrap the composition, e.g, pass the references to the dependency units.

B. Centrally managed configuration models and policies

An important concept behind software-defined IoT cloud systems is the late-bound runtime policies. Our units are configured declaratively, via the policies by utilizing the exposed software-defined API, without worrying about internals of the runtime mechanisms, i.e, the atomic units. To enable seamless binding of the atomic units we provide a special unit prototype, called *bootstrap container*. The bootstrap container acts as a plug-in system, which provides mechanisms to define (bind) the units based on supplied configurations or to redefine them when configuration policies are changed. For example, runtime changes of the units are achieved by invalidating affected parts of the existing dependency tree and dynamically rebuilding them, based on the new configuration directives. Therefore, the units can be simply "droped in" and our bootstrap container (re)binds them together, at runtime without rebooting system.

We decouple the configuration models (late-bound policies) from the functional units. Therefore, we can treat configuration policies as any software-defined IoT unit, which adheres to the general principles of software-defined IoT (Section III-A). By encapsulating the configuration policies in separate units, we can manage them at runtime via centralized configuration managements solutions for IoT cloud. Our framework provides mechanisms to specify and propagate the configuration models to the edge of IoT cloud infrastructure (e.g., gateways) and our bootstrap container enforces the provided directives. To this end, our bootstrap container initially binds functional and configuration units and continuously listens for configuration changes and applies them on the affected functional units, accordingly. To enable performing runtime modifications without

[2]http://opscode.com/chef
[3]http://docs.cloudfoundry.org/bosh/
[4]http://puppetlabs.org
[5]http://www.sedonadev.org/
[6]http://www.niagaraax.com/
[7]http://help.ubuntu.com/community/CloudInit/

worrying about any side-effects we require the configuration actions to be idempotent. The usual approach to achieve this is to wrap the units as OS services. Among other things the late-bound policies and our mechanisms for managed configuration enable flexible customization and dynamic configuration changes, at runtime.

V. Prototype and experiments

A. Prototype implementation

The main aim of our prototype is to enable developers and operations managers to dynamically, on-demand provision and deploy software-defined IoT systems. This includes providing software-defined IoT unit prototypes, enabling automated unit composition, at multiple levels and supporting centralized runtime management of the configuration models.

In Section III we introduced the conceptual model of our software-defined IoT units. To technically realize our units, we utilize the concept of virtual resource containers. More precisely, we provide different *unit prototypes* that can be customized and/or modified during runtime by adding required runtime mechanisms encapsulated in our atomic units. The unit prototypes provide resources with different granularity, e.g., VM flavors, group quotas, priorities, etc., and boilerplate functionality to enable automated provisioning of custom software-defined IoT units.

Figure 5 provides a high-level overview of the framework's architecture. Our framework is completely hosted in the cloud and follows a modular design which guarantees flexible and evolvable architecture. The current prototype is implemented atop OpenStack [2], which is an open source Infrastructure-as-a-Service (IaaS) cloud computing platform. *Presentation layer* provides an user interface via Web-based UI and RESTful API. They allow a user to specify various configuration models and policies, which are used by the framework to compose and deploy our units in the cloud. *Cloud core services layer* contains the main functionality of the framework. It includes the *PolicyProcessor* used to read the input configurations and transform it to the internal model defined in our framework. *Units management services* utilize this model for composing and managing the units. The *InitializationManager* is responsible for configuring and composing more complex units. It translates the directives specified in configuration models into concrete initialization actions on the unit level. In our current implementation, the core of the *InitalizationManager* is an OpsCode Chef client, which is passed to the VMs during initialization via Ubuntu cloud-init. *InitalizationManager* also provides mechanisms for configuration management. The *DeploymentManager* is used to deploy the software-defined IoT units in the cloud. Our prototype relies on SALSA[8], a deployment automation framework we developed. It utilizes the API exposed by the *CloudSystemWrapper* to enable deployment across various cloud providers, currently implemented for OpenStack cloud. *DeploymentManager* is responsible to manage and distribute the dependency references for the complex units (Section III-C). *Units persistence layer* provides functionality to store and manage our software-defined units and policies.

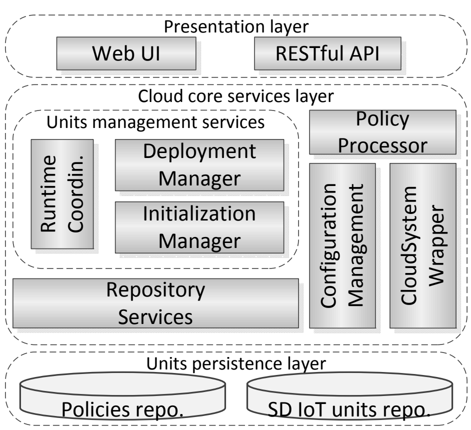


Fig. 5. Framework architecture overview.

B. Experiments

1) Scenario analysis: We now show how our prototype is used to provision a complex software-defined IoT unit, which provides functionality for the real-life FM location tracking service (Section II-A). The service reports vehicle location in near real-time on the cloud. To enable remote access, the monitored vehicles have an on-board device, acting as a gateway to its data and control points. To improve performance and reliability, the golf course provides on-site gateways, which communicate with the vehicles, provide additional processing and storage capabilities and feed the data into the cloud. Therefore, the physical IoT infrastructure comprises network connected vehicles, on-board devices and local gateways.

Typically, to provision the FM service system designers and operations manager would need to directly interact with the rigid physical IoT infrastructure. Therefore, they at least need to be aware of its topological structure and devices' capabilities. This means that the FM service also needs to have understanding of the IoT infrastructure, instead of being able to customize the infrastructure to its needs. Due to inherent inflexibility of IoT infrastructure, its provisioning usually involves long and tedious task such as manually logging into individual gateways, understanding gateway internals or even on site presence. Therefore, provisioning even a simple FM location tracking service involves performing many complex tasks. Due to a large number of geographically distributed vehicles and involved stakeholders IoT infrastructure provisioning requires a substantial effort prolonging service delivery and increasing costs. Subsequently, we show the advantages our units (Section III-B) and the provisioning techniques (Section IV) have to offer to operations managers and application designers in terms of: a) *Simplified provisioning* to reduce time, costs and possible errors; b) *Flexibility* to customize and modify the IoT units and their runtime topologies.

To enable the FM system we developed a number of atomic software-define IoT units[9] such as: a software-defined sensor that reports vehicle location in realtime, messaging infrastructure based on Apache ActiveMQ[10], software-defined protocol based on MQTT and JSON, the bootstrap container based on the Spring framework[11], and corresponding configuration units. The experiments are simulated on our OpenStack (Folsom) cloud and we used Ubuntu 12.10 cloud image (Memory:

[8]https://github.com/tuwiendsg/SALSA/

[9]https://github.com/tuwiendsg/SDM

[10]http://activemq.apache.org/

[11]http://projects.spring.io/spring-framework/

2GB, VCPUS: 1, Storage: 20GB). To display location changes we develop a Web application which displays changes of vehicles' location on Google Maps.

2) Simplified provisioning: To demonstrate how our approach simplifies provisioning of the virtual IoT infrastructure, we show how a user composes the FM complex software-defined IoT unit, using our framework. Figure 6 shows the custom deployment of the topological structure of the FM vehicle tracking unit, deployed in the cloud. The unit contains two gateways for the vehicles it tracks, a web server for the Web application and a message broker that connects them.

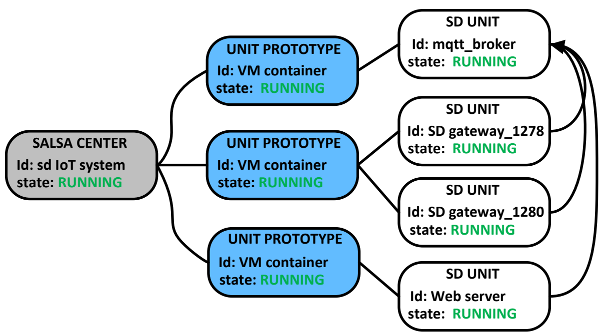


Fig. 6. Topological structure of FM vehicle tracking unit (a screen shot).

In order to start provisioning the complex unit, system designer only needs to provide a policy describing the required high-level resources and capabilities required by the FM service. For example, Listing 1 shows a snippet from the configuration policy for FM location tracking unit, that illustrates specifying a software-defined gateway, for the on-board device.

```
...
<tosca:NodeTemplate id="SD-Gateway"
  name="car_1278" type="vm">
  <tosca:Properties>
   <MappingProperties>
    <MappingProperty type="vm">
     <property name="instanceType">m1.small</property>
     <property name="provider">openstack@dsg</property>
     <property name="baseImage">ami-00000163</property>
    </MappingProperty>
   </MappingProperties>
  </tosca:Properties>
  <tosca:Requirements>
   <tosca:Requirement name="MQTT-broker-IP" type="String"
   id="brokerIp_Requirement"/>
  </tosca:Requirements>
  <tosca:DeploymentArtifacts>
   <tosca:DeploymentArtifact artifactType="chef"
   artifactRef="deployClient"/>
  </tosca:DeploymentArtifacts>
</tosca:NodeTemplate>
...
```

Listing 1. Partial TOSCA-like complex unit description.

The policy describes gateway's initial configuration and the cloud instance where it should be deployed. Additionally, it defines a dependency unit, i.e. the MQTT broker and specifies vehicle's Id that can be used to map it on the underlying device, as shown in [15]. Our framework takes the provided policy, spawns the required unit prototypes and provides them with references to the dependency units. At this stage the virtual infrastructure comprises solely of unit prototypes (VM-based). After performing the high-level unit composition and establish the dependencies between the units, the user continues composing on the finer granularity level. By applying the top-down approach we enable differing design decisions and enable early automation of known functionality, to avoid over-engineering and provisioning redundant resources.

In the next phase, the user provisions individual unit prototypes. To this end, he/she provides policies specifying desired finer-grained capabilities. Listing 2 shows example capabilities, that can be added to the gateway. To enable asynchronous pushing of the location changes it should communicate over the MQTT protocol. Listing 3 shows a part of Chef recipe used to add MQTT client to the gateway. Our framework fetches the atomic units, that encapsulate the required capabilities, from the repository and composes them automatically, relying on the software-defined API and our bootstrap container.

```
{"run_list":
  ["recipe[bootstrap_container]",
  "recipe[mqtt-client]",
  "recipe[protocol-config-unit]",
  "recipe[sd-sensor]"]
}
```

Listing 2. Run list for software-defined gateway.

```
include_recipe 'bootstrap_container::default'
remote_file "mqtt-client-0.0.1-SNAPSHOT.jar" do
 source "http://128.130.172.215/salsa/upload/files/..."
 group "root"
 mode 00644
 action :create_if_missing
end
```

Listing 3. Chef recipe for adding the MQTT protocol.

Therefore, compared to the traditional approaches, which require gateway-specific knowledge, using proprietary API, manually logging in the gateways to set data points, our *automated units composition* (Section IV-A), based on declarative unit configuration policies, simplifies the provisioning process and makes it traceable and repeatable. Our units can easily be shared among the stakeholders and composed to provide custom functionality. This enables system designers and operations managers to rely on the existing, established systems, thus reducing provisioning time, potential errors and costs.

3) Flexible customization: To exemplify the flexibility of our approach let us assume that we need to change configuration of the FM unit to use CoAP instead of MQTT. This can be due to requirements change (Section II-A), reduced network connectivity or simply to reuse the unit for a golf course with different networking capabilities. To customize the existing unit, an operations manager only needs to change the `"recipe[protocol-config-unit]"`unit (Listing 2) and provide an atomic unit for the CoAP client. This is a nice consequence of our late-bound runtime mechanisms and support for *managed configuration models*, provided by our framework. We treat both functional and configuration units in the same manner and our bootstrap container manages their runtime binding (Section IV-B). Compared to traditional approaches that require addressing each gateway individually, firmware updates or even modifications on the hardware level, our framework enables flexible runtime customization of our units and supports operation managers to seamlessly enforce configuration baseline and its modifications on a large-scale.

VI. Related Work

Recently, many interesting approaches enabling convergence of IoT and cloud computing appeared. For example, in [5], [8], [11], [16], [18] the authors mostly deal with IoT infrastructure virtualization and its management on cloud platforms, [6], [13] utilize the cloud for additional computation resources, in [17], [19] the authors mostly focus on utilizing cloud's storage resources for Big-IoT-Data and [12], [16] integrating IoT devices with enterprise applications based on SOA paradigm. Due to limited space in the following we only mention related work exemplifying approaches that deal with IoT infrastructure virtualization and management.

In [11] the authors develop an infrastructure virtualization framework for wireless sensor networks. It is based on a content-based pub/sub model for asynchronous event exchange and utilizes a custom event matching algorithm to enable delivery of sensory events to subscribed cloud users. In [18] the authors introduce sensor-cloud infrastructure that virtualizes physical sensors on the cloud and provides management and monitoring mechanisms for the virtual sensors. However, their support for sensor provisioning is based on static templates that, contrary to our approach, do not support dynamic provisioning of IoT capabilities such as communication protocols. SenaaS [5] mostly focuses on providing a cloud semantic overlay atop physical infrastructure. They define an IoT ontology to mediate interaction with heterogeneous devices and data formats, exposing them as event stream services to the upper layers. OpenIoT framework [16] utilizes semantic web technologies and CoAP [10] to enable web of things and linked sensory data. They mostly focus on discovering, linking and orchestrating internet connected objects, thus conceptually complementing our approach. In [8] the authors focus on developing a virtualization infrastructure to enable sensing and actuating as a service on the cloud. They propose a software stack which includes support for management of device identification, selection and aggregation, all of which can be seen as enablers of our approach. Also there are various commercial solutions such as Xively [4] and ThingWorx [3], which allow users to connect their sensors to the cloud and enable remote access and management.

Most of these approaches focus on different virtualization techniques for IoT devices and data format mediation. They also enable some form of configuration, e.g., setting sensor poll rates. These approaches can be seen as complementary to our own, as device virtualization sets the corner stone for achieving software-defined IoT systems. We rely on the advances in convergence of IoT and cloud, introduced by the previous work and extend it with novel concepts for abstracting and encapsulating IoT resources and capabilities, exposing them via software-defined API on the cloud and enabling fine-grained provisioning. Therefore, our approach can be seen as a natural step in evolution of IoT cloud systems.

VII. Conclusion

In this paper, we introduced the conceptual model of software-defined IoT units. To our best knowledge this is the first attempt to apply software-defined principles on IoT systems. We showed how they are used to abstract IoT resources and capabilities in the cloud, by encapsulating them in software-defined API. We presented automated unit composition and managed configuration, the main techniques for provisioning software-defined IoT systems. The initial results are promising in the sense that software-defined IoT system enable sharing of the common IoT infrastructure among multiple stakeholders and offer advantages to IoT cloud system designers and operations managers in terms of simplified, on-demand provisioning and flexible customization. Therefore, we believe that software-defined IoT systems can significantly contribute the evolution of the IoT cloud systems.

In the future we plan to continue developing the prototype and extend it in several directions: a) Providing techniques and mechanisms to support runtime governance of software-defined IoT systems; b) Enabling our software-defined IoT systems to better utilize the edge of infrastructure, e.g., by providing code distribution techniques between cloud and IoT devices; c) Enabling policy-based automation of data-quality, security and safety aspects of software-defined IoT systems.

Acknowledgment

This work is sponsored by Pacific Controls Cloud Computing Lab (PC3L[12])

References

[1] Mq telemetry transport. http://mqtt.org/. March 2014.

[2] Openstack. http://www.openstack.org/. March 2014.

[3] Thingworks. http://thingworx.com. Last accessed: March 2014.

[4] Xively. https://xively.com. Last accessed: March 2014.

[5] S. Alam, M. Chowdhury, and J. Noll. Senaas: An event-driven sensor virtualization approach for internet of things cloud. In *NESEA*, 2010.

[6] B.-G. Chun, S. Ihm, P. Maniatis, M. Naik, and A. Patti. Clonecloud: elastic execution between mobile device and cloud. In *Proceedings of the sixth conference on Computer systems*. ACM, 2011.

[7] S. Dawson-Haggerty, X. Jiang, G. Tolle, J. Ortiz, and D. Culler. smap: a simple measurement and actuation profile for physical information. In *SenSys*, pages 197–210, 2010.

[8] S. Distefano, G. Merlino, and A. Puliafito. Sensing and actuation as a service: a new development for clouds. In *NCA*, pages 272–275, 2012.

[9] S. Dustdar, Y. Guo, B. Satzger, and H.-L. Truong. Principles of elastic processes. *Internet Computing, IEEE*, 15(5):66–71, 2011.

[10] B. Frank, Z. Shelby, K. Hartke, and C. Bormann. Constrained application protocol (coap). *IETF draft, Jul*, 2011.

[11] M. M. Hassan, B. Song, and E.-N. Huh. A framework of sensor-cloud integration opportunities and challenges. In *ICUIMC*, 2009.

[12] M. Kovatsch, M. Lanter, and S. Duquennoy. Actinium: A restful runtime container for scriptable internet of things applications. In *Internet of Things*, pages 135–142, 2012.

[13] K. Kumar and Y.-H. Lu. Cloud computing for mobile users: Can offloading computation save energy? *Computer*, 43(4):51–56, 2010.

[14] B. Lantz, B. Heller, and N. McKeown. A network in a laptop: rapid prototyping for software-defined networks. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*. ACM, 2010.

[15] S. Nastic, S. Sehic, M. Voegler, H.-L. Truong, and S. Dustdar. Patricia - a novel programing model for iot applications on cloud platforms. In *SOCA*, 2013.

[16] J. Soldatos, M. Serrano, and M. Hauswirth. Convergence of utility computing with the internet-of-things. In *IMIS*, pages 874–879, 2012.

[17] P. Stuedi, I. Mohomed, and D. Terry. Wherestore: Location-based data storage for mobile devices interacting with the cloud. In *ACM Workshop on Mobile Cloud Computing & Services*. ACM, 2010.

[18] M. Yuriyama and T. Kushida. Sensor-cloud infrastructure-physical sensor management with virtualized sensors on cloud computing. In *NBiS*, 2010.

[19] A. Zaslavsky, C. Perera, and D. Georgakopoulos. Sensing as a service and big data. *arXiv preprint arXiv:1301.0159*, 2013.

[12] http://pcccl.infosys.tuwien.ac.at/

rtGovOps: A Runtime Framework for Governance and Operations in Large-scale IoT Cloud Systems

Conference Paper · January 2015

DOI: 10.1109/MobileCloud.2015.38

5 authors, including:

Stefan Nastic
TU Wien

Christian Inzinger
University of Zurich

Hong-Linh Truong
TU Wien

Schahram Dustdar
TU Wien

rtGovOps: A Runtime Framework for Governance in Large-scale Software-defined IoT Cloud Systems

Stefan Nastic, Michael Vögler, Christian Inzinger, Hong-Linh Truong, and Schahram Dustdar
Distributed Systems Group, Vienna University of Technology, Austria
Email: {*lastname*}@*dsg.tuwien.ac.at*

***Abstract*—The ongoing convergence of cloud computing and the IoT gives rise to the proliferation of diverse, large-scale IoT and mobile cloud systems. Such novel IoT cloud systems offer numerous advantages for all involved stakeholders. However, due to scale, complexity, and inherent geographical distribution of such systems, governing new IoT cloud resources poses numerous challenges. In this paper we introduce rtGovOps, a novel framework for on-demand runtime operational governance of software-defined IoT cloud systems. To illustrate the feasibility of our framework and its practical applicability to implement and execute *operational governance processes* in large-scale software-defined IoT cloud system, we evaluate our approach using a real-world case study on managing fleets of electric vehicles.**

I. Introduction

Current advances in the Internet of Things (IoT) and mobile cloud computing research have enabled the creation of unified IoT and mobile cloud infrastructures [1]–[5] that offer large pools of IoT cloud resources. Recently, software-defined IoT cloud systems have been introduced [6] to abstract from low-level resources (i.e., hardware) and enable their programmatic management through well-defined APIs. This enables refactoring the underlying IoT cloud infrastructure into finer-grained resource components whose functionality can be (re)defined after they have been deployed. While such systems create numerous opportunities by exploiting novel IoT and mobile cloud resources, they also introduce a number of challenges not previously encountered in traditional systems, to operate and govern such resources at runtime. Unfortunately, traditional governance approaches are hardly applicable for IoT cloud systems, mainly due to their dynamicity, heterogeneity, geographical distribution, and large scale. Supporting tools and mechanisms for runtime operational governance of IoT cloud systems remain largely undeveloped, thus placing much of the burden on operations managers to perform operational governance processes.

This calls for a systematic approach to govern IoT cloud resources throughout their entire lifecycle. In our previous work [7], we introduced the GovOps methodology to effectively manage runtime governance in software-defined IoT cloud systems. The main purpose of GovOps is to close the gap between high-level governance objectives (e.g., costs, legal issues or compliance) and underlying operations processes that support such objectives. Therefore, GovOps mostly focuses on designing and realizing *operational governance processes* (e.g., similar to [8], [9]), which represent a subset of the overall IoT cloud governance and incorporate relevant aspects of both high-level governance strategies and underlying operations management. Continuing along this line of research, in this paper, we introduce the rtGovOps framework for dynamic, on-demand operational governance of software-defined IoT cloud systems at runtime. The rtGovOps framework provides runtime mechanisms and enabling techniques to reduce the complexity of IoT cloud operational governance, thus enabling operations managers to perform custom operational governance processes more efficiently in large-scale IoT cloud systems.

The remainder of the paper is structured as follows: Section II presents a motivating case study and summarizes our background work; Section III outlines main concepts and the design of the rtGovOps framework; In Section IV, we explain major runtime mechanisms of rtGovOps; Section V describes preliminary experimental results and outlines the current prototype implementation; Section VI discusses related work; Finally, Section VII concludes the paper and gives an outlook of our future research.

II. Motivation and background

A. Scenario

Let us consider a realistic application scenario in the domain of vehicle management that we will refer to throughout the paper. This scenario is based on an ongoing collaboration with industry partners[1] and our current effort in software-defined IoT cloud systems.

The Fleet Management System (FMS) is a real-world software-defined IoT cloud system responsible for managing fleets of zero-emission, electric vehicles deployed worldwide, e.g., on different golf courses. Vehicles communicate with the cloud via 3G or Wi-Fi networks to exchange telematic and diagnostic data. On the cloud, FMS provides different applications and services to manage this data. Relevant services include realtime vehicle status, remote diagnostics, and remote control. The FMS is currently used by the following three types of stakeholders: vehicle manufacturers, distributors, and golf course managers. These stakeholders have different business models. For example, when a manufacturer only leases vehicles to customers, they are interested in the status and upkeep of the complete fleet, will perform regular maintenance, as well as monitor crashes and battery health. Golf course managers are mostly interested in vehicle security to prevent misuse and ensure safety on the golf course (e.g., using geofencing features). In general, the stakeholders rely on the FMS and its services to optimize their respective business tasks.

1) FMS IoT cloud infrastructure: The FMS runs atop a nontrivial IoT cloud infrastructure that includes a variety of IoT cloud resources. Figure 1 gives a high-level overview of the FMS infrastructure. For our discussion, the two most relevant types of IoT cloud resources are on-board physical gateways (G) and cloud virtual gateways (VG). Most of

[1] http://pcccl.infosys.tuwien.ac.at/

the vehicles are equipped with on-board gateways that are capable to host lightweight services such as geofencing or local diagnostics services. For legacy cars that are not equipped with such gateways, a device acting as a CAN-IP bridge is used (e.g, Teltonika FM5300[2]). In this case FMS hosts virtual gateways on the cloud that execute the aforementioned services on behalf of the vehicles.

We notice that the FMS is a large-scale system that manages thousands of vehicles and relies on diverse cloud communication protocols. Further, the FMS depends on IoT cloud resources that are geographically distributed on different golf courses around the globe. Jurisdiction over these resources can change over time, e.g., when a vehicle is handed over from the distributor to a golf course manager. In addition, these resources are usually constrained. This is why the FMS heavily relies on cloud services, e.g., for computationally intensive data processing, fault-tolerance or to reliably store historical readings of vehicle data. While the cloud offers the illusion of unlimited resources, systems of such scale as FMS can incur very high costs in practice (e.g., of computation or networking). Finally, due to the large number of involved stakeholders, the FMS needs to enable runtime customizations of infrastructure resources in order to exactly meet stakeholder requirements and allow for operation within specified compliance and legal boundaries.

Therefore, the IoT cloud resources need to be managed and governed throughout their entire lifecycle. In our approach, this is captured and modeled as *operational governance processes*.

2) Example operational governance processes: Subsequently, we highlight some basic operational governance processes in FMS that are facilitated through our framework:

- Typically, the FMS polls diagnostic data from vehicles (e.g., with CoAP). However, a golf course manager could design an operational governance process that is triggered in specific situations such as in case of emergency. Such process could, for example, increase the update rate of the vehicle sensors and change the communication protocol to MQTT in order to satisfy a high-level governance objective, e.g., company's compliance policy to handle emergency updates in (near) real-time.
- To increase fault-tolerance and guarantee history preservation of vehicle data (e.g., due to governance objectives related to legal requirements), a distributor could decide to spin up additional virtual gateways in a different availability zone.
- After multiple complaints about problems with vehicles of type X, a manufacturer would need to add additional monitoring features to all vehicles of type X to perform more detailed inspections.

This is by no means a comprehensive list of operational governance processes in software-defined IoT cloud systems. However, due to dynamicity, heterogeneity, geographical distribution, and the large scale of IoT cloud systems, traditional approaches to realize even basic operational governance processes are hardly feasible in practice. This is mostly because such approaches implicitly make assumptions such as physical on-site presence, manually logging into gateways, understanding device specifics, etc., which are difficult, if not impossible, to meet in IoT cloud systems. Therefore, due to a lack of systematic approaches for operational governance in IoT cloud systems, operations managers currently have to rely on ad-hoc solutions to deal with the characteristics and complexity of IoT cloud systems when performing operational governance processes.

[2] http://www.teltonika.lt/en/pages/view/?id=1024

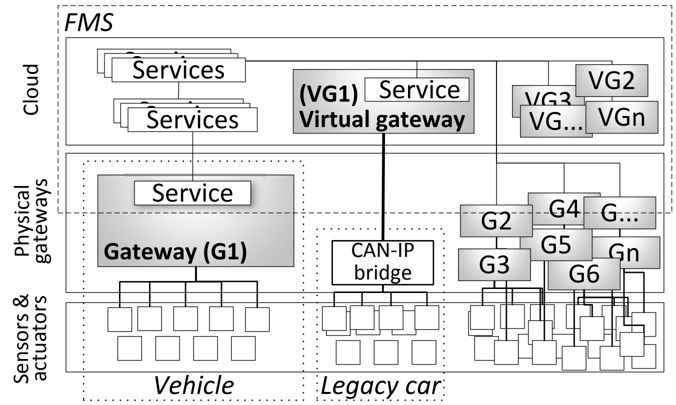


Fig. 1. Overview of FMS infrastructure.

B. Background

As we have shown earlier in [6], in software-defined IoT cloud systems, IoT cloud resources (e.g., gateways) are described as software-defined IoT units. The software-defined IoT units enable abstracting the underlying IoT cloud resources and allow for their management through well-defined APIs, exposed by these units. One of the main advantages of software-defined IoT cloud systems is opening up the traditional infrastructure silos and moving one step higher in the abstraction, i.e., effectively making applications independent of the underlying rigid infrastructure. The most important consequence is that the *functionality and states* of the underlying IoT cloud resources can be redefined in software during runtime. For example, new features such as additional cloud communication protocols can be added to the units at runtime.

In our previous work [7] we introduced the general GovOps approach for runtime governance in software-defined IoT cloud systems, as well as the concepts of *operational governance processes* that manipulate the states of IoT units (IoT cloud resources) at runtime. Such processes can be seen as a sequence of operations, which perform runtime state transitions from a current state to some desired target state (e.g., that satisfies some non-functional properties, enforces compliance, or exactly meets custom requirements). We presented a GovOps reference model that provides suitable abstractions to specify such operational governance processes. We also outlined the GovOps methodology on how to design the operational governance processes and realize governance strategies. At this point it is worth reminding that, from a technical perspective, GovOps does not make any assumptions about the implementation of operational governance processes, in the sense that such processes can be realized as business processes (e.g., using BPMN), with governance policies, via Domain Specific Languages (DSLs), or even as dedicated governance applications or services.

III. Overview of the rtGovOps Framework

The main aim of our rtGovOps (*runtime GovOps*) framework is to facilitate operational governance processes for software-defined IoT cloud systems. To this end, rtGovOps provides a set of runtime mechanisms and does most of the "heavy lifting" to support operations managers in implementing and executing operational governance processes in large-scale software-defined IoT cloud systems, without worrying about scale, geographical distribution, dynamicity, and other characteristics inherent to such systems that currently hinder operational governance in practice.

To facilitate performing the operational governance processes, while considering the characteristics of the software-defined IoT cloud systems, the rtGovOps framework follows a set of design principles, which include:

Central point of operation (R1) – Enable conceptually centralized interaction with the software-defined IoT cloud system to enable a unified view on the system's operations and governance capabilities (available at runtime), without worrying about low-level infrastructure details.

Automation (R2) – Allow for dynamic, on-demand governance of software-defined IoT cloud systems on a large scale and enable governance processes to be easily repeatable, i.e., enforced across the IoT cloud, without manually logging into individual gateways.

Fine-grained control (R3) – Expose the control functionality of IoT cloud resources at fine granularity to allow for precise definition of governance processes (to exactly meet requirements) and flexible customization of IoT cloud system governance capabilities.

Late-bound directives (R4) – Support declarative directives that are bound later during runtime in order to allow for designing generic and flexible operational governance processes.

IoT cloud resources autonomy (R5) – Provide a higher degree of autonomy to IoT cloud resources to reduce communication overhead, increase availability (e.g., in case of network partitions), enable local exception and fault handling, support protocol independent interaction, and increase system scalability.

Figure 2 gives a high-level architecture and deployment overview of the rtGovOps framework. Generally, the rtGovOps framework is distributed across the cloud and IoT devices. It is designed based on the microservices architecture[3], which among other things enables flexible, evolvable, and fault-tolerant system design, while allowing for flexible management and scaling of individual components. The main components of rtGovOps include: i) the *governance capabilities*, ii) the *governance controller* that runs on the cloud, and iii) the *rtGovOps agents* that run in IoT devices. In the remainder of this section, we will discuss these components in more detail.

A. Operational governance capabilities

As we described in Section II, operational governance processes govern software-defined IoT units throughout their entire lifecycle. Generally, *Governance capabilities* represent the main building blocks of operational governance processes and they are usually executed in IoT devices. The governance capabilities encapsulate governance operations which can be applied on deployed IoT units, e.g., to query the current version of a service, change a communication protocol, or spin up a virtual gateway. Such capabilities are described via well-defined APIs and are usually provided by domain experts who develop the IoT units. The rtGovOps framework enables such

[3]http://martinfowler.com/articles/microservices.html

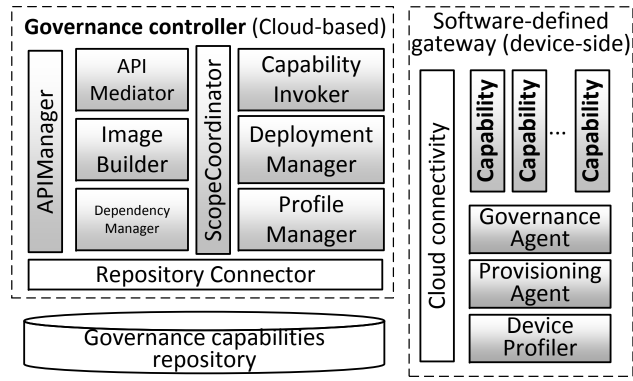


Fig. 2. Overview of rtGovOps architecture and deployment.

capabilities to be dynamically added to the system (e.g, to gateways), and supports managing their APIs. From a technical perspective, they behave like add-ons, in the sense that they extend the resources with additional operational functionality. Internally, IoT devices host rtGovOps agents that behave like an add-on manager, responsible for installing/enabling, starting/stopping a capability, and managing the APIs they expose. Generally, rtGovOps does not make any assumptions about concrete capability implementations. However, it requires them to be packaged as shown in Figure 3. Subsequently, we highlight relevant examples of governance capabilities related to our FMS application.

- *Configuration-specific capabilities* include changes to the configuration models of software-defined IoT cloud systems at runtime. For example: setting sensor poll rate, changing communication protocol for cloud connectivity, configuring data point unit and type (e.g., temperature in Kelvin as unsigned 10-bit integer), mapping a sensor or CAN bus unit to a device's virtual pin, or activating a low-pass filter for an analog sensory input.
- *Topology-specific capabilities* address structural changes that can be performed on the deployment topologies of software-defined IoT systems. Examples include replicating a virtual gateway to increase fault-tolerance or data source history preservation and push data processing logic from the application space towards the edge of the infrastructure.
- *Stream-specific capabilities* deal with managing the runtime operation of sensory data streams and continuous Complex Event Processing (CEP) queries. Therefore, to enable features like scaling out or stream replaying, operations managers need capabilities such as: filter placement near the data source to reduce network traffic, allocation of queries to gateways, and stream splitting, i.e., sending events to multiple virtual gateways.
- *Monitoring-specific capabilities* deal with adding a general monitoring metric, e.g., CPU load, or providing an implementation of a custom metric to IoT cloud resources.

For the sake of simplicity, in this paper, we assume that the capabilities are readily available[4]. In reality, they can be obtained from a central repository, provided by a third-party in a market-like fashion, or custom developed in-house.

[4]We provide example governance capabilities under https://github.com/tuwiendsg/GovOps/

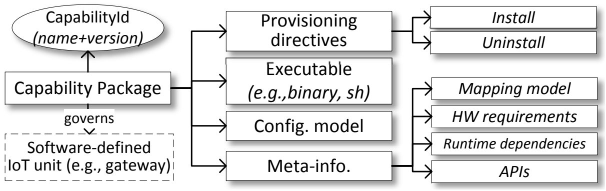


Fig. 3. Overview of capability package structure.

As mentioned above, governance capabilities are dynamically added to the IoT cloud resources. There are several reasons why such behavior is advantageous for operations managers and software-defined IoT cloud systems. For example, as we usually deal with constrained resources, static provisioning of such resources with all available functionality is rarely possible (e.g., factory defaults rarely contain the desired configuration for FMS vehicle gateways). Further, as we have seen in Section II, jurisdiction over resources (in this case FMS vehicles) can change during runtime, e.g., when a vehicle is handed over to a golf course manager. In such cases, because the governing stakeholder changes, it is natural to assume that the requirements regarding operational governance will also change, thus requiring additional or different governance capabilities. As opposed to updating the whole device image at once, we reduce the communication overhead, but also enable changing device functionality without interrupting the system, e.g., to reboot. This provides greater flexibility and enables on-demand governance tasks (e.g, by temporally adding a capability), which are often useful in systems with a high degree of dynamicity. Finally, executing capabilities in the IoT devices improves scalability of the operational governance processes and enables better resource utilization.

B. Operational governance processes and governance scopes

Operational governance processes represent a subset of the general IoT cloud governance and deal with operating and governing IoT cloud resources at runtime. Such processes are usually designed by operations managers in coordination with business stakeholders [7]. The main purpose of such processes is to enable supporting high-level governance objectives such as compliance and legal concerns, which influence system's runtime behavior. To be able to dynamically govern IoT cloud resources, the operational governance processes rely on the governance capabilities. This means that individual steps of such process usually invoke governance capabilities in order to enforce the behavior of IoT cloud resources in such manner that it complies with the governance objectives. In this context, our rtGovOps framework provides runtime mechanisms to enable execution of these operational governance processes.

As we have mentioned earlier (Section II-B), we use software-defined IoT units to describe IoT cloud resources. However, these units are not specifically tailored for describing non-functional properties and available meta information about IoT cloud resources, e.g., location of a vehicle (gateway) or its specific type and model. For this purpose, rtGovOps provides governance scopes. The governance scope is an abstract resource which represents a group of IoT cloud resources that share some common properties. For example, an operations manager can specify a governance scope to include all the vehicles of type X. The *ScopeCoordinator* (Figure 2) provides mechanisms to define and manage the governance scopes. The rtGovOps framework relies on the *ScopeCoordinator* to determine which IoT cloud resources need to be affected by an operational governance process. Generally, the governance scopes enable implementing the operational governance processes in a scalable and generic manner, since the IoT cloud resources do not have to be individually referenced within such process.

C. Governance controller and rtGovOps agents

The *Governance controller* (Figure 2) represents a central point of interaction with all available governance capabilities. It provides a mediation layer that enables operations managers to interact with IoT cloud systems in a conceptually centralized fashion, without worrying about geographical distribution of the underlying system. Internally, the governance controller comprises several microservices, among which the most important include: *DeploymentManager* and *ProfileManager* that are used to support dynamical provisioning of the governance capabilities, as well as *APIManager* and previously mentioned *ScopeCoordinator* that support operational governance processes to communicate with the underlying capabilities. The *APIManger* exposes governance capabilities to operational governance processes via well-defined APIs and handles all API calls from such processes. It is responsible to resolve incoming requests, map them to respective governance capabilities in the IoT devices and deliver results to the calling process. Among other things, this involves discovering capabilities by querying the capabilities repository, and parameterizing capabilities via input arguments or configuration directives.

Since governance capabilities are usually not "pre-installed" in IoT devices, the *DeploymentManager* is responsible to inject capabilities into such devices (e.g., gateways) at runtime. To this end it exposes REST APIs, which are used by the devices to periodically check for updates, as well as by the operational governance processes to push capabilities into the devices. Finally, the *ProfileManager* is responsible to dynamically build and manage device profiles. This involves managing static device meta-information and periodically performing profiling actions in order to obtain runtime snapshots of current device states.

Another essential part of the rtGovOps framework are the *rtGovOps agents*. They include: *ProvisioningAgent*, *GovernanceAgent* and *DeviceProfiler*. These agents are very lightweight components that run in all IoT cloud resources that are managed by rtGovOps such as the FMS vehicles. Figure 4 shows a high-level overview of the *GovernanceAgent* architecture. It is responsible to manage local governance capabilities, to wrap them in well-defined APIs and to expose them to the *Governance controller*. The rtGovOps agents offer advantages

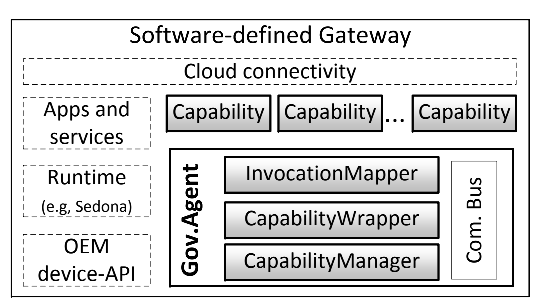


Fig. 4. Overview of the governance agent architecture.

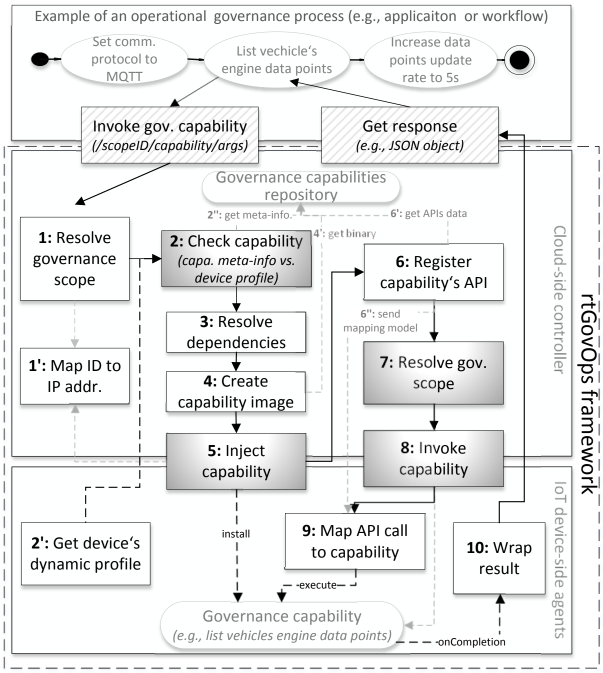


Fig. 5. Execution of an operational governance process.

in terms of general scalability of the system and provide a degree of autonomy to the IoT cloud resources.

IV. The rtGovOps framework main concepts and enabling techniques

Generally, the rtGovOps framework supports operations managers to handle two main tasks. First, the rtGovOps framework enables dynamic, on-demand *provisioning of governance capabilities*. For example, it allows for dynamically injecting capabilities into IoT cloud resources, and coordinating the dynamic profiles of these resources at runtime. Second, our framework allows for runtime management of governance capabilities throughout their entire lifecycle that, among other things, includes *remote capability invocation and managing dynamic APIs* exposed to users.

As we have mentioned earlier, in order to achieve a high-level governance objective such as enforce (part of) compliance policies for handling emergency situations an operations manager could design an operational governance process similar to the one shown in Figure 5 (top). Individual actions of such processes usually reference specific governance capabilities and rely on rtGovOps to support their execution. Figure 5 depicts a simplified sequence of steps executed by the rtGovOps framework when a governance capability gets invoked by an operational governance process. For the sake of clarity, we omit several steps performed by the framework and mainly focus on showing the most common interaction, i.e., we assume no errors or exceptions occur. We will discuss the most important steps performed by rtGovOps below. Note that all of these steps are performed transparently to operations managers and operational governance processes. The only thing that such processes observe is a simple API call (similar to REST service invocation) and a response (e.g., a JSON array in this case). Naturally, the process is responsible to provide arguments and/or configuration directives that are used by rtGovOps to parametrize the underlying capabilities.

A. Automated provisioning of governance capabilities

In order to enable dynamic, on-demand provisioning of governance capabilities whenever a new capability is requested (i.e., referenced in an operational governance process), the rtGovOps framework needs to perform the following steps: i) the *ScopeCoordinator* resolves the governance scope to get a set of devices to which the capability will be added; ii) the *ProfileManager* checks whether the governance capability is available and compatible with the device; iii) the *Dependency Manager* resolves runtime dependencies of the capability; iv) the *ImageBuilder* creates a capability image; v) Finally, the *DeploymentManager* injects the capability into devices; An overview of this process is also shown in steps $1-5$ in Figure 5.

Algorithm 1 shows the capability provisioning process in more detail. An operational governance process requests a capability by supplying a capability ID (currently consisting of capability name and version) and an operational governance scope (more detail in Section IV-B). After that rtGovOps tries to add the capability (together with its runtime dependencies) to a device. If successful, it continues along the steps shown in Figure 5. The algorithm performs in a similar fashion to a fail-safe iterator, in the sense that it works with snapshots of devices states. For example, if something changes on the device side inside $checkComponent$ (Algorithm 1, lines $2-5$) it cannot be detected by rtGovOps and in this case the behavior of rtGovOps is not defined. Since we assume that all the changes to the underlying devices are performed exclusively by our framework, this is a reasonable design decision. Other errors, such as failure to install a capability on a specific device, are caught by rtGovOps and delivered as notifications to the operational governance process, so that they do not interrupt its execution.

Algorithm 1: Governance capability provisioning.

```
   input : capaID : A capability ID.
           gscope : Operational governance scope.
   result: Capability added to device or error occurred.
 1 func checkComponent (component, device)
 2     capaMeta ← queryCapaRepo(component)
 3     devProfile ← getDeviceProfile(device)
 4     status ← isCompatible(capaMeta, devProfile)
 5     return status
 6 end
   /* Begin main loop.                         */
 7 components ← resolveDependencies(capaID)
 8 components ← add(capaID)
 9 for device in resolveGovScope(gscope) do
10     for component in components do
11         if not checkComponent(component, device) then
12             error
13         end
14     end
       /* Inject capability.                    */
15     capaImg ← createImg(components)
16     deployCapa(capaImg, device)
17     installCapa(capaImg) // On device-side
18 end
```

1) Capability checking: From the steps presented in Algorithm 1 $checkComponent$ (lines $1-6$) and $injectCapability$ (lines $15-17$) are the most interesting. The framework invokes $checkComponent$ for each governance capability and all of its dependencies for the currently considered device. At this point rtGovOps verifies that the component can be installed on this specific device. To this end, the *ProfileManager* first queries the *central capabilities repository*. Besides the capability binaries, the repository stores capability meta-information, such as required CPU instruction set (e.g., ARMv5 or x86), disk space and memory requirements, as well as installation and decommissioning directives. After obtaining the capability meta-information the framework starts building the current device profile. This is done in two stages. First, the gateway features catalog is queried to obtain relevant static information, such as CPU architecture, kernel version and installed userland (e.g., BusyBox[5]) or OS. Second, the *ProfileManager* in coordination with *DeviceProfiler* executes a sequence of runtime profiling actions to complete the dynamic device profile. For example, the profiling actions include: currently available disk space, available RAM, firewall settings, environment information, list of processes and daemons, and list of currently installed capabilities. Finally, when the dynamic device profile is completed, it is compared with the capability's meta information in order to determine if the capability is compatible with the device.

2) Capability injection: The rtGovOps *capability injection mechanism*, deals with uploading and installing capabilities on devices, as well as managing custom configuration models. This process is structured along three main phases: Creating a capability image, deploying the capability image on a device and installing the capability locally on the device.

i) After the *ProfileManager* determines a capability is compatible with the gateway, the *ImageBuilder* creates a capability image. The capability image is rtGovOps internal representation of the capability package (see Figure 3). In essence it is a compressed capability package containing component binaries and a dynamically created runlist. The runlist is an ordered list of components that need to be installed. It is created by the *DependencyManager* and its individual steps reference component installation or decommissioning directives that are obtained from the *capabilities repository*.

ii) In the second phase, *DeploymentManager* deploys the image to the device. We support two different deployment strategies. The first strategy is *poll-based*, in the sense that the image is placed in the update queue and remains there for a specified period of time (TTL). The *ProvisioningAgent* periodically inspects the queue for new updates. When an update is available, the device can poll the new image when it is ready, e.g., when the load on it is not too high. A governance process can have more control over the poll-based deployment by specifying a capability's priority in the update queue. Finally, on successful update the *DeploymentManager* removes the update from the queue. The second deployment strategy allows governance capabilities to be asynchronously *pushed* to gateways. Since the capability is forced onto the gateway, it should be used cautiously and for urgent updates only, such as increasing a sensor poll-rate in emergency situations. Finally, independent of the deployment strategy, the framework performs a sequence of checks to ensure that an update was performed correctly (e.g., compares checksums) and moves to the next phase.

iii) In the final phase, the *ProvisioningAgent* performs a local installation of the capability binaries and its runtime dependencies, and performs any custom configurations. Initially, *ProvisioningAgent* unpacks the previously obtained capability image and verifies that the capability can be installed based on the current device profile. In case the conditions are not satisfied, e.g., due to disk space limitation, the process is aborted and an error is sent to the *DeploymentManager*. Otherwise, the *ProvisioningAgent* reads the runlist and performs all required installation or decommissioning steps.

[5]http://busybox.net

A limitation of the current rtGovOps prototype is that it only provides rudimentary support to specify installation and decommissioning directives. Therefore, capability providers need to specify checks, e.g., if a configuration file already exists, as part of the installation directives. In the future we plan to provide a dedicated provisioning DSL to support common directives and interactions.

B. rtGovOps APIs and invocation of governance capabilities

When a new governance capability is injected into a gateway, the rtGovOps framework performs the following steps: i) register capability with *APIManager*; ii) *ScopeCoordinator* resolves the governance scope; iii) *APIMediator* provides a mapping model to the *GovernanceAgent*; iv) the *GovernanceAgent* wraps the capability into a well-defined API, dynamically exposing it to the outside world; v) *CapabilityInvoker* invokes the capability and deliver the result to the invoking operational governance process when the capability execution completes. A simplified version of this process is also shown in steps $6-10$ in Figure 5.

Before we dive into technical details of this process, it is worth mentioning that currently in the capabilities repository, besides aforementioned capability meta-information and binaries, we also maintain well-defined capability API descriptions, e.g., functional, meta and lifecycle APIs. These APIs are available to operations managers as soon as a capability is added to the repository and independent of whether the capability is installed on any device. Additionally, we provide a general rtGovOps API that is used to allow for more control over the system and its capabilities. It includes *CapabilityManager* API (e.g., list capabilities, check if capability installed/active), capability lifecycle API (e.g., start, stop or remove capability), and *ProvisioningAgent* API (e.g., install new capability). Listing 1 shows some examples of such APIs as REST-like services (version numbers are omitted for clarity).

```
/* General case of capability invocation.     */
/govScope/{capabilityId}/{methodName}/{arguments}?
arg1={first-argument}&arg2={second-argument}&...

/* Data points capability invocation example. */
/deviceId/DPcapa/setPollRate/arguments?rate=5s
/deviceId/DPcapa/list

/* Capabilities manager examples.             */
/deviceId/cManager/capabilities/list
/deviceId/cManager/{capabilityId}/stop
```

Listing 1. Examples of capabilities and rtGovOps APIs.

1) Single invocation of governance capabilities: In the following we mainly focus on explaining the steps that are performed by the rtGovOps framework when a capability is invoked on a single device. The more general case involving multiple devices and using operational governance scopes is discussed in the next section.

When a capability gets invoked by an operational governance process for the first time, *APIManager* does not know anything about it. Therefore, it first needs to check, based on the API call (e.g., see Listing 1), if the capability exists in the central capabilities repository. After the capability is found and provisioned (Section IV-A), the rtGovOps framework tries to invoke the capability. This involves the following steps: registering the capability, mapping the API call, executing the capability, and returning the result.

i) First, the *APIManager* registers the API call with the corresponding capability. This involves querying the capability repository to obtain its meta-information (such as expected arguments), as well as building a dynamic mapping model. Among other things, the mapping model contains the capability ID, a reference to a runtime environment (e.g., Linux shell), a sequence of input parameters, the result type, and further configuration directives. The *APIMediator* forwards the model to the device (i.e. *GovernanceAgent*) and caches this information for subsequent invocations. During future interactions, the rtGovOps framework acts as transparent proxy, since subsequent steps are handled by the underlying devices.
ii) In the next step, rtGovOps needs to perform a mapping between the API call and the underlying capability. Currently, there are two different ways to do this. By default, rtGovOps assumes that capabilities follow the traditional Unix interaction model, i.e., that all arguments and configurations (e.g., flags) are provided via the standard input stream (stdin) and output is produced to standard output (stdout) or standard error (stderr) streams. This means, if not specified otherwise in the mapping model, the framework will try to invoke the capability by its ID and will forward the provided arguments to its stdin. For capabilities that require custom invocation, e.g., property files, policies, or specific environment settings, the framework requires a custom mapping model. This model is used in the subsequent steps to correctly perform the API call.
iii) Finally, the *CapabilityInvoker* in coordination with the *GovernanceAgent* invokes the governance capability. As soon as the capability completes, the *GovernanceAgent* collects and wraps the result. Currently, the framework provides means to wrap results as JSON objects for standard data types and it relies on the mapping model to determine the appropriate return type. However, this can be easily extended to support more generic behavior, e.g., by using Google Protocol Buffers[6].

2) Operational governance scopes: When an operational governance process gets invoked on a governance scope, the aforementioned invocation process remains the same, with the only difference that rtGovOps performs all steps on a complete governance scope in parallel instead on an individual device. To this end, the *ScopeCoordinator* enables dynamic resolution of the governance scopes.

There are several ways how a governance scope can be defined. For example, an operations manager can manually assign a set of resources to a scope, such as all vehicles belonging to a golf course, or they can be dynamically determined depending on runtime features by querying governance capabilities to obtain dynamic properties such as current configuration model. To bootstrap defining the governance scopes, the *ScopeCoordinator*, defines a global governance scope that is usually associated with all the IoT cloud resources administered by a stakeholder at the given time. Governance scope specifications are implemented as composite predicates referencing device meta information and profile attributes, The predicates are applied to the global scope, filtering out all resources that do not match the provided attribute conditions. The *ScopeCoordinator* uses the resulting set of resources to initiate capability invocation with the *CapabilityInvoker*. The *ScopeCoordinator* is also responsible to provide support for gathering results delivered by the invoked capabilities. This is needed since the scopes are resolved in parallel and the results are asynchronously delivered by the IoT devices.

[6]http://code.google.com/p/protobuf/

V. Evaluation & Prototype implementation

A. Prototype implementation

In the current prototype, the rtGovOps *Governance controller* microservices are implemented in Java and Scala programming languages. The rtGovOps agents are based on lightweight httpd server and are implemented as Linux shell scripts. The complete source code and supplement materials providing more details about current rtGovOps implementation are publicly available in Git Hub[7].

B. Experiments setup

In order to evaluate how our rtGovOps framework behaves in a large-scale setup (hundreds of gateways), we created a virtualized IoT cloud testbed based on CoreOS[8]. In our testbed we use Docker containers to virtualize and mimic physical gateways in the cloud. These containers are based on a snapshot of a real-world gateway, developed by our industry partners. The Docker base image is publicly available in Docker Hub under dsgtuwien/govops-box[9].

For the subsequent experiments we deployed a CoreOS cluster on our local OpenStack cloud. The cluster consists of 4 CoreOS 444.4.0 VMs (with 4 VCPUs and 7GB of RAM), each running approximately 200 Docker containers. Our rtGovOps agents are preinstalled in the containers. The rtGovOps Governance controller and capabilities repository are deployed on 3 Ubuntu 14.04 VMs (with 2VCPUs and 3GB of RAM). The operational governance processes are executing on a local machine (with Intel Core i7 and 8GB of RAM).

C. Governing FMS at runtime

We first show how our rtGovOps framework is used to support performing operational governance processes on a real-world FMS application for monitoring vehicles (e.g., location and engine status) on a golf course (Section II). The application consists of several services. On the one side, there is a lightweight service running in the vehicle gateways that interfaces with vehicle sensors via the CAN protocol, and feeds sensory data to the cloud. On the cloud-side of the FMS application, there are several services that, among other things, perform analytics on the sensory data and offer data visualization support. In our example implementation of this FMS application, the gateway service is implemented as a software-defined IoT unit that among other things provides an API and mechanisms to dynamically change the cloud communication protocol without stopping the service.

The FMS application polls diagnostic data from vehicles with CoAP. However, in case of an emergency, a golf course

[7]http://github.com/tuwiendsg/GovOps
[8]http://coreos.com/
[9]https://registry.hub.docker.com/u/dsgtuwien/govops-box/

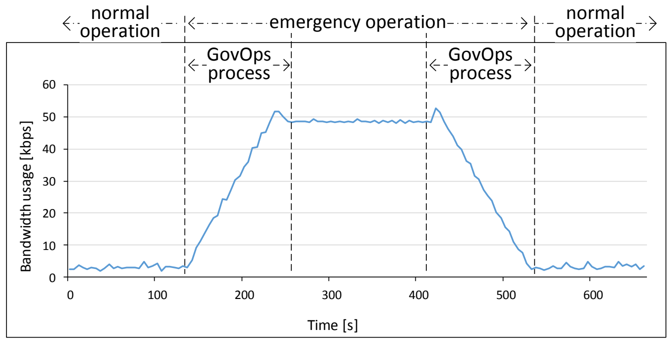


Fig. 6. Example execution of operational governance process in the FMS.

manager needs to increase the update rate and switch to MQTT in order to handle emergency updates in (near) real-time. This can be easily specified with an operational governance process that contains the following steps: change communication protocol to MQTT, list vehicle engine and location data points and set data points update rate, e.g, to 5 seconds. These steps are also depicted in Figure 5 (top). The golf course manager relies on rtGovOps governance capabilities to realize individual process steps and rtGovOps mechanisms (Section IV) to execute the operational governance process.

Figure 6 shows the bandwidth consumption of the FMS application that monitors 50 vehicles over a period of time. We notice two distinct operation modes: normal operation and operation in case of an emergency (emergency operation). Most notable are the two transitions: first, from normal to emergency operation and second, returning from emergency to normal operation. These transitions are described with the aforementioned operational governance process that is executed by the rtGovOps framework. The significant increase in bandwidth consumption happens during the execution of the operational governance process, because it changes the communication protocol from polling the vehicles approximately every minute with CoAP, to pushing the updates every 5 seconds with MQTT.

Typically, when performing processes such as the transition from normal to emergency operation without the rtGovOps framework, golf course managers (or generally operations managers) need to directly interact with vehicle gateways. This usually involves long and tedious tasks such as manually logging into gateways, dealing with device specific configurations or even an on-site presence. Therefore, realizing even basic governance processes, such as the one we presented above, involves performing many manual and error prone tasks, usually resulting in a significant increases in operations costs. Additionally, in order to be able to have a timely realization of governance processes and consistent implementation of governance strategies across the system, very large operations and support teams are required. This is mainly due to the large scale of the FMS system, but also due to geographical distribution of the governed IoT resources, i.e., vehicles.

Besides the increased efficiency, the main advantage that rtGovOps offers to operations managers is reflected in the flexibility of performing operational governance processes at runtime. For example, in Figure 6 the execution of the operational governance process took around 2 minutes. In our framework this is, however, purely a matter of operational governance process configuration (naturally with upper limits as we show in the next section). This means, the operational governance process can be easily customized to execute the protocol transition "eagerly", in the sense to force the change as soon as possible, even within seconds, or "lazy", to roll-out the change step-wise, e.g., 10 vehicles at the time. The most important consequence is the opportunity to effectively manage tradeoffs. For example, executing the process eagerly incurs higher costs, due to additional networking and computation consumption, but it is needed in most emergency situations. Conversely, executing the process in a lazy manner can be desirable for non-emergency situations, since operations managers can prevent possible errors to affect the whole system.

Figure 6 also shows that rtGovOps introduces a slight communication overhead. This is observed in the two peaks at the end of the first process execution, when the framework performs the final checks that the process completed successfully and also when the second process gets triggered, i.e., when the capabilities get invoked on the vehicles. However, in our experiments this overhead was small enough not to be statistically significant. An additional performance-related concern of using rtGovOps is that network latency can slow down the execution of the operational governance process. However, since rtGovOps follows the microservices architecture style, it is possible to deploy relevant services (*API-* and *DeploymentManager*) on Cloudlets [3] near the vehicles, e.g., on golf courses, where they can utilize local wireless networks.

D. Experiments results

To demonstrate the feasibility of using rtGovOps to facilitate operational governance processes in large-scale software-defined IoT cloud systems, we evaluate its performance to govern approximately 800 vehicle gateways that are simulated in the previously described test-bed. In our experiments, we mainly focus on showing the scalability of the two main mechanisms of the rtGovOps framework: (i) capability invocation and (ii) automated capability provisioning. We also consider the performance of capability checking and governance scope resolution. The reason why we put an emphasis on the scalability of our framework is that it is one of the key factors to enable consistent implementation of governance objectives across a large-scale systems. For example, if the execution of an operational governance process were to scale exponentially with the size of the resources pool, theoretically it would take infinitely long to have a consistent enforcement of the governance objectives in the whole system, with sufficiently large resource pool. The results of the experiments are averaged results of 30 repetitions and we have experimented with 5 different capabilities that have different properties related to their size and computational overhead.

Figure 7 shows the execution time of the first invocation of a capability (stacked bar) and an average invocation time of capability execution (plain bar). We notice that the first invocation took approximately between 10 and 15 seconds and average invocation varied between 4 and 6 seconds depending on the scope size (measured in the number of gateways). The main reason for such a noticeable difference is the invocation caching performed by rtGovOps. This means that most of the steps, e.g., building capability image and building the mapping model are only performed when a capability is invoked for the first time, since in the subsequent invocations the capability is already in the gateways and the mapping can be done in cache. In Figure 8, we present the average execution time of a capability (as it is observed by an invoking operational governance process on the locale machine), average execution

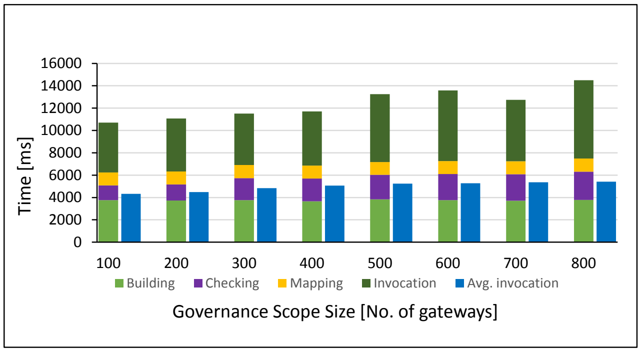


Fig. 7. Capabilities first invocation.

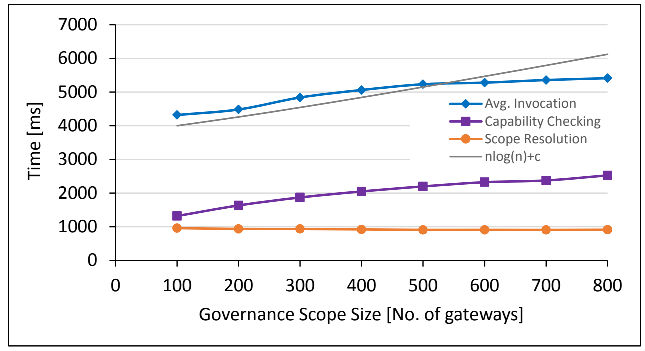


Fig. 8. Average invocation time of capabilities on a governance scope.

of capability checking mechanism and governance scope resolution. As a reference, the diagram also shows a plot of a $nlog(n) + c$ function. We can see that the mechanisms scale within $O(nlog(n))$ for relatively large governance scopes (up to 800 gateways), which can be considered a satisfactory result. We also notice that computational overheads of the capabilities have no statistically significant impact on the results, since they are distributed among the underlying gateways. Finally, it is interesting to notice that the scope resolution time actually decreases with increasing scope size. The reason for this is that in the current implementation of rtGovOps, scope resolution always starts with the global governance scope and applies filters (lambda expressions) on it. After some time Java JIT "kicks-in" and optimizes filters execution, thus reducing the overall scope resolution time.

In Figure 9, we show the general execution times of the rtGovOps capability provisioning mechanism (push-based deployment strategy) for two different capabilities. The first one has a size order of magnitude in MB and second capability size is measured in KB. There are several important things to notice here. First, the capability provisioning also scales similarly to $O(nlog(n))$. Second, after the governance scope size reaches 400 gateways there is a drop in the capability provisioning time. The reason for this is that the rtGovOps load balancer spins-up additional instances of the *DeploymentManger* and *ImageBuilder*, naturally reducing provisioning time for subsequent requests. Finally, the provisioning mechanism behaves in a similar fashion for both capabilities. The reason for this is that all gateways are in the same network, what can be seen as an equivalent to vehicles deployed on one golf course.

E. Discussion and lessons learned

The observations and results of our experiments show that rtGovOps offers advantages in terms of realizing operational governance processes with greater flexibility, and also makes such processes easily repeatable, traceable and auditable, which is crucial for successful implementation of governance strategies. Generally, by adopting the notion of governance capabilities and by utilizing resource agents, rtGovOps allows for operational governance processes to be specified with *finer granularity (R3)*, but also give a *degree of autonomy (R5)* to the managed IoT cloud resources. Therefore, by selecting suitable governance capabilities, operations managers can precisely define desired states and runtime behavior of software-defined IoT cloud systems. Further, since the capabilities are executed locally in IoT cloud resources (e.g., in the gateways), our framework enables better utilization of the "edge of infrastructure" and allows for local error handling, thus increasing system availability and scalability. Further, the main advantage of approaching provisioning and management of governance capabilities in the described manner is that operation managers do not have to worry about geographically-distributed IoT cloud infrastructure nor deal with individual devices, e.g., key management or logging in. They only need to *declare (R4)* which capabilities are required in the operational governance process and specify a governance scope. The rtGovOps framework takes care of the rest, effectively giving a *logically centralized view (R1)* on the management of all governance capabilities. Further, by *automating (R2)* the capability provisioning, rtGovOps enables installing, configuring, deploying, and invoking the governance capabilities in a scalable and easily repeatable manner, thus reducing errors, time, and eventually costs of operational governance.

It should be also noted that there is a number of technical limitations of and possible optimizations that can be introduced in the current prototype of the rtGovOps framework. As we have already mentioned, rtGovOps currently offers limited support for specifying provisioning directives. Additionally, while experimenting with different types of capabilities, we noticed that in many cases a better support to deal with streaming capabilities would be useful. Regarding possible optimizations, in the future we plan to introduce support for automatic composition of capabilities on the device level, e.g., similar to Unix piping. This should reduce the communication overhead of rtGovOps and improve resource utilization in general. In spite of the current limitations, the initial results are promising, in the sense that rtGovOps *increases flexibility* and enables *scalable execution of operational governance processes* in software-defined IoT cloud systems.

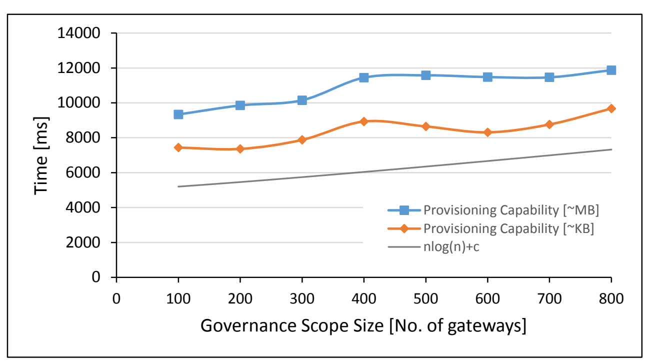


Fig. 9. Average capability provisioning duration (push-based strategy).

VI. Related work

Recently, IoT governance has been receiving a lot of attention. For example, in [10] the author evaluates various governance aspects, such as privacy, security, ethics, etc., and defines main principles of IoT governance, e.g., legitimacy, transparency, openness, and accountability. Governance approaches such as CMMI [11] or COBIT [12] also provide models and methodologies to manage governance objectives. Such approaches are complementary to our own and can be used along with rtGovOps to specify, manage and trace such high-level governance objectives.

Also approaches dealing with IoT mobile cloud operations and resources management have recently emerged. For example, in [1], [13]–[16] the authors mostly deal with IoT infrastructure virtualization and its management on the clouds. Approaches presented in [3]–[5], [17]–[19] address the issues to aggregate and manage the computational resources provided by various IoT devices, mobile devices and the clouds. In [2], [20] the authors focus on utilizing clouds for additional storage resources. In [16] the authors develop an infrastructure virtualization framework, based on asynchronous event exchange. They provide an event matching algorithm to enable coordination and management of sensory event streams. In [1] the authors propose virtualizing physical sensors on the cloud and provide template-based management and monitoring mechanisms for the virtual sensors. SenaaS [13] mostly focuses on providing a cloud semantic overlay atop physical IoT infrastructure. It defines an ontology to manage interaction with heterogeneous devices and mediate different data formats. OpenIoT framework [14] utilizes semantic web technologies and CoAP, to enable discovering, linking and orchestrating Internet connected objects. In [15] the authors focus on enabling sensing and actuating as a service, providing support for device management and identification, as well as their selection and aggregation. Edge Cloud Composits [5] provide support to enable mobile devices to enhance their resources, by utilizing additional near-by devices or remote clouds. Such approaches provide various governance capabilities such as template-based controlling of sensor groups, registering and decommissioning sensors, orchestrating IoT devices, as well as monitoring the IoT cloud systems QoS. Aforementioned approaches provide techniques for optimizing IoT cloud resources utilization, such as computation offloading, service migration or context-aware resource aggregation. Our rtGovOps conceptually builds on such approaches, and goes one step further by providing support for dynamic, on-demand provisioning of the governance capabilities and enabling their management throughout the entire lifecycle. By doing so rtGovOps increases flexibility and facilitates execution of operational governance processes in large-scale IoT cloud systems.

VII. Conclusion and future work

In this paper, we introduced the rtGovOps framework for runtime operational governance in software-defined IoT cloud systems. We presented rtGovOps runtime mechanisms and enabling techniques that support operations managers to handle two main tasks: (i) perform dynamic, on-demand provisioning of governance capabilities and (ii) remotely invoke such capabilities in IoT cloud resources remotely, via dynamic APIs. We demonstrated, on a real-world case study, how our framework can be used to facilitate execution of operational governance processes in large-scale software-defined IoT cloud systems. The initial results are promising in several aspects. We showed that the rtGovOps framework enables operational governance processes to be executed in a scalable manner across relatively large IoT cloud resource pools. Additionally, we discussed how rtGovOps enables flexible execution of operational governance processes by automating the execution of such processes to a large extent, offering finer-grained control over IoT cloud resources and providing a logically centralized interaction with IoT cloud resource pools.

In the future we plan to address the current limitations of rtGovOps, described in Section V. We also plan to extend the rtGovOps framework to support specifying and managing high-level governance objectives.

Acknowledgment

This work is sponsored by Pacific Controls Cloud Computing Lab (PC3L).

References

[1] M. Yuriyama and T. Kushida, "Sensor-cloud infrastructure-physical sensor management with virtualized sensors on cloud computing," in *NBiS*, 2010.

[2] P. Stuedi, I. Mohomed, and D. Terry, "Wherestore: Location-based data storage for mobile devices interacting with the cloud," in *MCS*, 2010.

[3] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *Pervasive Computing*, vol. 8, no. 4, pp. 14–23, 2009.

[4] E. Cuervo, A. Balasubramanian, D.-k. Cho, A. Wolman, S. Saroiu, R. Chandra, and P. Bahl, "Maui: making smartphones last longer with code offload," in *MobiSys '10*, pp. 49–62, ACM, 2010.

[5] K. Bhardwaj, S. Sreepathy, A. Gavrilovska, and K. Schwan, "Ecc: Edge cloud composites," in *MobileCloud 2014*, pp. 38–47, IEEE, 2014.

[6] S. Nastic, S. Sehic, D.-H. Le, H.-L. Truong, and S. Dustdar, "Provisioning software-defined iot systems in the cloud," in *FiCloud*, 2014.

[7] S. Nastic, C. Inzinger, H.-L. Truong, and S. Dustdar, "Govops: The missing link for governance in software-defined iot cloud systems," in *WESOA14*, 2014.

[8] SOA Software, "Integrated SOA governance." URL: http://www.soa.com/solutions/integrated_soa_governance. [Online; accessed June-2014].

[9] IBM , "SOA pages - Definition of SOA governance." URL: http://ibm.com/software/solutions/soa/gov/. [Online; accessed June-2014].

[10] R. H. Weber, "Internet of things–governance quo vadis?," *Computer Law & Security Review*, vol. 29, no. 4, pp. 341–347, 2013.

[11] D. M. Ahern, A. Clouse, and R. Turner, *CMMI distilled: a practical introduction to integrated process improvement*. Addison-Wesley Professional, 2004.

[12] G. Hardy, "Using IT governance and COBIT to deliver value with IT and respond to legal, regulatory and compliance challenges," *Information Security technical report*, vol. 11, no. 1, pp. 55–61, 2006.

[13] S. Alam, M. Chowdhury, and J. Noll, "Senaas: An event-driven sensor virtualization approach for internet of things cloud," in *NESEA*, 2010.

[14] J. Soldatos, M. Serrano, and M. Hauswirth, "Convergence of utility computing with the internet-of-things," in *IMIS*, pp. 874–879, 2012.

[15] S. Distefano, G. Merlino, and A. Puliafito, "Sensing and actuation as a service: a new development for clouds," in *NCA*, pp. 272–275, 2012.

[16] M. M. Hassan, B. Song, and E.-N. Huh, "A framework of sensor-cloud integration opportunities and challenges," in *ICUIMC*, 2009.

[17] B.-G. Chun, S. Ihm, P. Maniatis, M. Naik, and A. Patti, "Clonecloud: elastic execution between mobile device and cloud," in *Conference on Computer systems*, ACM, 2011.

[18] K. Kumar and Y.-H. Lu, "Cloud computing for mobile users: Can offloading computation save energy?," *Computer*, vol. 43, no. 4, pp. 51–56, 2010.

[19] R. Kemp, N. Palmer, T. Kielmann, and H. Bal, "Cuckoo: a computation offloading framework for smartphones," in *Mobile Computing, Applications, and Services*, pp. 59–79, Springer, 2012.

[20] A. Zaslavsky, C. Perera, and D. Georgakopoulos, "Sensing as a service and big data," *arXiv preprint arXiv:1301.0159*, 2013.

GovOps: The Missing Link for Governance in Software-Defined IoT Cloud Systems

Stefan Nastic(✉), Christian Inzinger, Hong-Linh Truong, and Schahram Dustdar

Distributed Systems Group, Vienna University of Technology, Vienna, Austria
{nastic,inzinger,truong,dustdar}@dsg.tuwien.ac.at

Abstract. Cloud computing and the IoT are converging ever stronger, enabling the proliferation of diverse large-scale IoT cloud systems. Such novel IoT cloud systems offer numerous advantages for the variety of involved stakeholders. However, due to scale, complexity, and inherent geographical distribution of IoT cloud systems, governing new IoT cloud resources and capabilities poses numerous challenges. In this paper, we introduce GovOps – a novel approach and a conceptual model for cloud-based, dynamic governance of software-defined IoT cloud systems. By introducing a suitable *GovOps reference model* and a dedicated *GovOps manager*, it simplifies realizing governance processes and enables performing custom governance tasks more efficiently in practice. We introduce real-world case studies in the building automation and vehicle management domains, to illustrate the main aspects and principles of our approach to governance of large-scale software-defined IoT cloud systems.

1 Introduction

To date, cloud computing models and techniques, such as infrastructure virtualization and management, Compute-, Storage- and Network-as-a-Service, etc., have been intensively exploited for large-scale Internet of Things (IoT) systems [7,14,18]. Recently, software-defined IoT cloud systems have been introduced [10] in order to enable easier provisioning and management of IoT cloud resources and capabilities. Generally, software-defined denotes a principle of abstracting low-level components (e.g., hardware) and enabling their management, programmatically through well-defined APIs [8]. This enables refactoring the underlying infrastructure into finer-grained resource components whose functionality can be (re)defined after they have been deployed. While IoT cloud systems introduce numerous possibilities, a plethora of challenges to govern and operate these new IoT cloud resources and capabilities emerge.

Various domains, such as smart building and vehicle management, increasingly rely on IoT cloud resources and capabilities. Consequently, governance issues such as security, safety, legal boundaries, compliance, and data privacy concerns are ever stronger being addressed [4,5,17], mainly due to their potential impact on the variety of involved stakeholders. However, existing approaches are mostly

F. Toumani et al. (Eds.): ICSOC 2014, LNCS 8954, pp. 20–31, 2015.
DOI: 10.1007/978-3-319-22885-3_3

intended for high-level business stakeholders, neglecting support, e.g., tools and frameworks, to realize governance strategies in large-scale, geographically distributed IoT cloud systems. Approaching IoT cloud from the operations management perspective, different approaches have been presented, e.g. [2,14,15,18]. Such approaches deal with IoT cloud infrastructure virtualization and its management, enabling utilization of cloud computation resources and operating cloud storage resources for big IoT data. However, most of these approaches do not consider high-level governance objectives such as legal issues and compliance. This increases the risk of lost requirements or causes over-regulated systems, potentially increasing costs and limiting business opportunities.

Currently, IoT governance mostly addresses the *Internet* part of the IoT, e.g., in the context of the Future Internet services[1], while IoT operations processes mostly deal with *Things* (e.g., in [3]) as additional resources that need to be operated. Therefore, governance objectives (law, compliance, etc.) are not easily mapped to operations processes (e.g., querying sensory data streams or adding/removing devices). Contemporary models, which assume that business stakeholders define governance objectives, and operations managers implement and enforce them, are hardly feasible in IoT cloud systems. In practice, bridging the gap between governance and operations management of IoT cloud systems poses significant challenges, because traditional management and governance approaches are hardly applicable for IoT cloud systems, mainly due to the large number of involved stakeholders, novel requirements for shared resources and capabilities, dynamicity, geographical distribution, and the sheer scale of IoT cloud systems.

This calls for a systematic approach to govern and operate IoT cloud resources and capabilities. Extending the previously developed concepts [10], in this paper we introduce GovOps – a novel approach for cloud-based dynamic governance and operations management in software-defined IoT cloud systems. The main objectives of GovOps are twofold. On the one side, it aims to enable seamless integration of high-level governance objectives with concrete operations processes. On the other side, it enables performing operational governance processes for IoT cloud systems in such manner that they are feasible in practice. We present a GovOps reference model that defines required roles, concepts, and techniques to reduce the complexity of realizing IoT cloud governance processes. GovOps enables performing custom governance tasks more efficiently, thus reduces time, costs, and potential consequences of insufficient or ineffective governance.

The remainder of this paper is structured as follows: Sect. 2 presents motivating scenarios that will be used throughout the paper. In Sect. 3, we present the GovOps approach to governance and operations management in software-defined IoT cloud systems; Sect. 4 outlines the GovOps reference model; Sect. 5 discusses the related work; Finally, Sect. 6 concludes the paper and gives an outlook of our future work.

[1] http://ec.europa.eu/digital-agenda/en/internet-things.

2 Scenarios: Governing Software-Defined IoT Systems

Consider the following scenarios in the Building Automation and Vehicle Management domains that we will refer to throughout the rest of this paper. The scenarios are derived from our work conducted in the P3CL lab[2].

2.1 Scenario 1 – Fleet Management System

General Description. Fleet Management System (FMS) is responsible for managing electric vehicles deployed worldwide, e.g., on different golf courses. We have identified three stakeholders who rely on the FMS to optimize their business tasks: vehicle manufacturer, distributors and golf course managers. The stakeholders have different business models. For example, as the manufacturer only leases the vehicles, he is interested in the complete fleet, e.g., regular maintenance, crash reports and battery health. On the other side, golf course managers are mostly interested in vehicles security (e.g., geofencing features), preventing misuse, and safety on the golf course.

Infrastructure Setup. The FMS is an IoT cloud system comprising vehicles' proprietary on-board gateways, network and cloud infrastructure. The on-board gateway is capable to host lightweight applications for: vehicle maintenance, tracking, monitoring and club set-up. Vehicles communicate with the cloud via 3G, GPRS or Wi-Fi network to exchange telematic and diagnostic data. On the cloud we host different FMS subsystems and services to manage and analyze this data, e.g., determine vehicle status, perform remote diagnostics, batch configuration and software updates. Legacy vehicles that are not capable to host applications are integrated using a CAN-IP bridge, and any custom business logic needs to be executed in the cloud.

2.2 Scenario 2 – Building Automation System

General Description. Building Automation System (BAS) is responsible to monitor and control various building assets, such as HVAC, lighting, elevators and humidity control systems, as well as to handle fault events and alarms (e.g., fire or gas leakage). For safety-critical services (e.g., alarm handling), timely processing of the events and the availability of the BAS play a crucial role and need to be ascertained.

Infrastructure Setup. Generally, BAS comprises a set of cloud-based services, gateways and various sensors and actuators integrated with the building's assets. Gateways which support typical BAS device protocols (ModBus, BACnet, LonWorks and Fox), e.g., Niagra or Sedona[3], are used to communicate with sensors

[2] http://pcccl.infosys.tuwien.ac.at/.
[3] http://www.tridium.com/.

and actuators. For local processing, the gateways usually allow executing custom triggers, rules and some form of complex event processing (CEP) queries. For permanent storage and more resource-demanding processing, the gateways send streams of data to the remote cloud services.

2.3 System Characteristics

We notice that both the FMS and the BAS have large-scale, geographically distributed infrastructure. Additionally, the FMS utilizes virtualized IoT cloud infrastructure, such as virtual gateways (VGW), to support integrating legacy vehicles. Depending on stakeholder and task-at-hand our systems have different customization requirements and non-functional requirements (e.g., regarding fault-tolerance and availability). For example in BAS, while for safety-critical services, real-time delivery and processing is essential, for services such as HVAC controller, cost reduction is more important. Due to the multiplicity of the involved stakeholders, the FMS needs to allow for flexible runtime customizations in order to exactly meet the stakeholder's functional requirements, depending on the problem-at-hand and availability or accessibility of the vehicles, as well as desired system's non-functional properties.

3 GovOps – A Novel Approach to Governance and Operations Management in IoT Cloud

The main objective of our GovOps approach (Governance and Operations) is twofold. On the one side it aims to enable seamless integration of high-level governance objectives and strategies with concrete operations processes. On the other side, it enables performing operational governance processes for IoT cloud systems in such manner that they are feasible in practice.

Figure 1 illustrates how GovOps relates to IoT cloud governance and operations. It depicts the main idea of GovOps to bring governance and operations closer together and bridge the gap between governance objectives and operations processes, by incorporating the main aspects of both IoT cloud governance and operations management. To this end, we define *GovOps principles and design process* of GovOps strategies (Sect. 4) that support determining what can and needs to be governed, based on the current functionality and features of an IoT cloud system, and that allow for aligning system's capabilities with regulations and standards. Additionally, we introduce a novel role, *GovOps manager* (Sect. 3.3) responsible to guide and manage designing GovOps strategies,

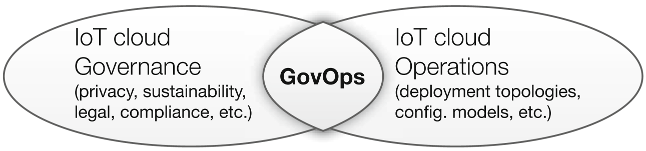


Fig. 1. GovOps in relation to IoT cloud governance and operations.

because in practice it is very difficult, risky, and ultimately very costly to adhere to traditional organizational silos, separating business stakeholders from operations managers. Therefore, GovOps integrates business rules and compliance constraints with operations capacities and best-practices, from early stages of designing governance strategies in order to counteract system over-regulation and lost governance requirements.

It is worth noting that GovOps does not attempt to define a general methodology for IoT cloud governance. There are many approaches (Sect. 5), which define governance models and accountability frameworks for managing governance objectives and coordinating decision making processes. Most of these approaches can be applied within GovOps without substantial modifications.

3.1 Governance Aspects

From our case studies, we have identified various business stakeholders such as building residents, building managers, governments, vehicle manufacturers and golf course managers. Typically, they are interested in energy efficient and greener buildings, sustainability of building assets, legal and privacy issues regarding sensory data, compliance (e.g., regulatory or social), health of the fleet, as well as security and safety issues related to the environments under their jurisdiction.

Depending on the concrete (sub)system and the involved stakeholders, governance objectives are realized via different governance strategies. Generally, we identify the following governance aspects: (i) *environment-centric*, (ii) *data-centric* and (iii) *infrastructure-centric governance.*

Environment-centric governance deals with issues of overlapping jurisdictions in IoT cloud managed environments. For example, in our BAS, we have residents, building managers and the government that can provide governance objectives, which directly or indirectly affect an environment, e.g., a residential apartment. In this context, we need to simultaneously articulate multiple governance objectives related to comfort of living, energy efficiency, safety, health and sustainability.

Data-centric governance mostly deals with implementing the governance strategies related to the privacy, quality, and provenance of sensory data. Examples include addressing legal issues, compliance, and user preferences regarding the sensory data.

Infrastructure-centric governance addresses issues about designing, installing, and deploying IoT cloud infrastructure. This mostly affects the early stages of introducing an IoT cloud system and involves feasibility studies, cost analysis, and risk management. For example, it supports deciding between introducing new hardware or virtualizing the IoT cloud infrastructure.

3.2 Operations Management Aspects

Operations managers implement various processes to manage BAS and FMS at runtime. Generally, we distinguish following operational governance aspects: (i) *configuration-centric*, (ii) *topology-centric*, and (iii) *stream-centric governance.*

Configuration-centric governance includes dynamic changes to the configuration models of deployed software-defined IoT cloud systems at runtime. Example processes include (a) enabling/disabling an IoT resource or capability (e.g., start/stop a unit), (b) changing an IoT capability at runtime (e.g., communication protocol), and (c) configuring an IoT resource (e.g., setting sensor poll rate).

Topology-centric governance addresses structural changes that can be performed on software-defined IoT systems at runtime. For example, (a) Pushing processing logic from the application space towards the edge of the infrastructure; (b) Introducing a second gateway and an elastic load balancer to optimize resource utilization; (c) Replicating a gateway, e.g., for fault-tolerance or data-source history preservation.

Stream-centric governance addresses runtime operation of sensor data streams and continuous queries, e.g., to perform custom filtering, aggregation, and querying of the available data streams. For example, to perform local filtering the processing logic is executed on physical gateways, while complex queries, spanning multiple data streams are usually executed on VGWs. Therefore, operations managers perform processes like: (a) Placement of custom filters (e.g., near the data source to reduce network traffic); (b) Allocation of queries to VGWs; and (c) Stream splitting, i.e., sending events to multiple VGWs.

3.3 Integrating Governance Objectives with Operations Processes

The examples presented in Sects. 3.1 and 3.2 are by no means a comprehensive list of IoT cloud governance processes. However, due to dynamicity, heterogeneity, geographical distribution and the sheer scale of IoT cloud, traditional approaches to realize these processes are hardly feasible in practice. This is mostly because such approaches implicitly make assumptions such as physical on-site presence, manually logging into gateways, understanding device specifics, etc., which are difficult, if not impossible, to meet in IoT cloud systems. Therefore, due to a lack of a systematic approach for operational governance in IoT systems, currently operations managers have to rely on ad hoc solutions to deal with the characteristics and complexity of IoT cloud systems when performing governance processes.

Further, Table 1 lists examples of governance objectives and according operations management processes to enforce these objectives. The first example comes from the FMS, since many of the golf courses are situated in countries with specific data regulations, e.g., the US or Australia. In order to enable monitoring of the whole fleet (as required by the manufacturer) the operations managers need to understand the legal boundaries regarding data privacy. For example, in Australia, the Office of the Australian Information Commissioner (OAIC) has issued an extensive guidance[4] as to what reasonable steps to protect personal information might include, that in practice need to be interpreted by operations managers. The second example contains potentially conflicting objectives supplied by stakeholders, e.g., building manager, end user, and the government,

[4] http://www.oaic.gov.au/privacy/applying-privacy-law/app-guidelines/.

Table 1. Example governance objectives and operations processes.

	Governance objectives	Operations processes
1	Fulfill legal requirements w.r.t. sensory data in country X. Guarantee history preservation	Spin-up an aggregator gateway. Replicate VGW, e.g., across different availability zones.
2	Reduce GHG emission. User preferences regarding living comfort. Consider health regulations	Provide configuration directives for an IoT cloud resource (e.g., HVAC).
3	Data quality compliance regarding location tracking services	Choose among available services, e.g., GPS vs. GNSS (Global Navigation Satellite System) platform.

leaving it to the operations team to solve the conflicts at runtime. The third example hints that GNSS is usually better-suited to simultaneously work in both northern and southern high latitudes. Therefore, even for these basic processes, an operations team faces numerous difficulties, since in practice there is no one-size-fits-all solution to map governance objectives to operations processes.

To address these issues, GovOps proposes a novel role, *GovOps manager*, as a dedicated stakeholder responsible to bridge the gap between governance strategies and operations processes in IoT cloud systems. The main rationale behind introducing a GovOps manager is that in practice designing governance strategies needs to involve operations knowledge about the technical features of the system, e.g., physical location of devices, configuration models, placement of queries and component replication strategies. Reciprocally, defining systems configurations and deployment topologies should incorporate standards, compliance, and legal boundaries at early stages of designing operations processes. To achieve this, the GovOps manager is positioned in the middle, in the sense that he/she continuously interacts with both business stakeholders (to identify high-level governance issues) and operations team (to determine operations capacities).

The main task of a GovOps manager is to determine suitable tradeoffs between satisfying the governance objectives and the system's capabilities, as well as to continuously analyze and refine how high-level objectives are articulated through operations processes. In this context, a key success factor is to ensure effective and continuous communication among the involved parties during the decision making process, facilitating (i) openness, (ii) collaboration, (iii) establishment of a dedicated GovOps communication channel, along with (iv) early adoption of standards and regulations. This ensures that no critical governance requirements are lost and counteracts over-regulation of IoT cloud systems. On the other side, in order to support performing runtime operations processes in IoT cloud systems, while considering system characteristics (e.g., large-scale, geographical distribution and dynamicity), GovOps proposes a set of concepts that includes: (i) central point of operation, (ii) automation, (iii) fine-grained control, (iv) late-bound policies, and (v) resource autonomy.

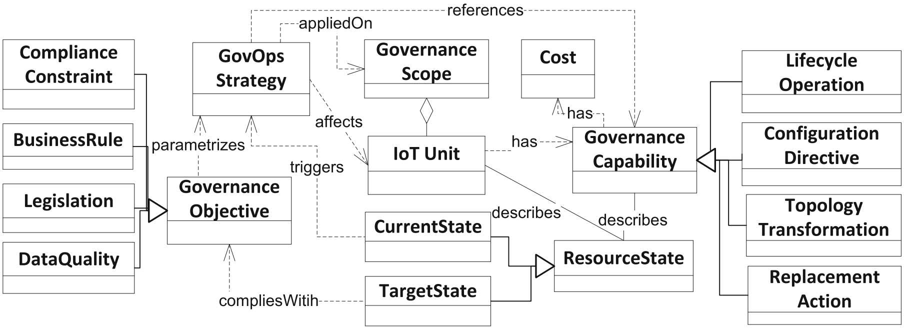


Fig. 2. Simplified UML diagram of GovOps model for IoT cloud governance.

4 A Reference Model for GovOps in IoT Cloud

4.1 Overview of GovOps Model for Software-Defined IoT Cloud Systems

To realize the GovOps approach we need suitable abstractions to describe IoT cloud resources that allow IoT cloud infrastructure to be (re)defined after it has been deployed. We show in [10] how this can be done with *software-defined IoT units*. The GovOps model (Fig. 2) builds on this premise and extends our previous work with fundamental aspects of operational governance processes: (i) describing states of deployed IoT resources, (ii) providing capabilities to manipulate these states at runtime, and (iii) defining governance scopes.

Within our model, the main building blocks of GovOpsStrategies are *GovernanceCapabilities*. They represent operations which can be applied on IoT cloud resources, e.g., query current version of a software, change communication protocol, and spin-up a virtual gateway. These operations manipulate IoT cloud resources in order to put an IoT cloud system into a specific (target) state. Governance capabilities are described via software-defined APIs and they can be dynamically added to the system, e.g., to a software-defined gateway. From a technical perspective, they behave like add-ons, in the sense that they extend resources with additional operational functionality. Generally, by adopting the notion of governance capabilities, we allow for processes to be automated to a great extent, and also give a degree of autonomy to IoT cloud resources.

Since the meaning of a resource state is highly task specific, we do not impose many constraints to define it. Generally, any useful information about an IoT cloud resource is considered to describe the *ResourceState*, e.g., a configuration model or monitoring data such as CPU load. Technically, there are many frameworks (e.g., Ganglia or Nagios) that can be used to (partly) describe resource states. Also configuration management solutions, such as OpsCode Chef, can be used to maintain and inspect configuration states. Finally, design best practices and reference architectures (e.g., AWS Reference Architectures[5]) provide a higher-level description of the desired target states of an IoT cloud system.

[5] http://aws.amazon.com/architecture/.

The *GovernanceScope* is an abstract resource, which represents a group of IoT cloud resources (e.g., gateways) that share some common properties. Therefore, our governance scopes are used to dynamically delimit IoT cloud resources on which a *GovernanceCapability* will have an effect. This enables writing the governance strategies in a scalable manner, since the IoT cloud resources do not have to be individually addressed. It also allows for backwards compatible GovOps strategies, which do not directly depend on the current resource capabilities. This means that we can move a part of the problem, e.g., fault and exception handling, inside the governance scope. For example, if a gateway loses a capability the scope simply will not invoke it i.e., the strategy will not fail.

4.2 Design Process of GovOps Strategies

As described in Sect. 3, the GovOps manager is responsible to oversee and guide the GovOps design process and to design concrete GovOps strategies. The design process is structured along three main phases: (i) identifying governance objectives and capabilities, (ii) formalizing strategy, and (iii) executing strategy.

Generally, the initial phase of the design process involves eliciting and formalizing governance objectives and constraints, as well as identifying required fine-grained governance capabilities to realize the governance strategy in the underlying IoT cloud system. GovOps does not make any assumptions or impose constraints on formalizing governance objectives. To support specifying governance objectives the GovOps manager can utilize various governance models and frameworks, such as 3P [13] or COBIT [6]. However, it requires tight integration of the GovOps manager into the design process and encourages collaboration among the involved stakeholders to clearly determine risks and tradeoffs, in terms of what should and can be governed in the IoT cloud system, e.g., which capabilities are required to balance building emission regulations and residents temperature preferences. To this end, the GovOps manager gathers available governance capabilities in collaboration with the operations team, identifies missing capabilities, and determines if further action is necessary. Generally, governance capabilities are exposed via well-defined APIs. They can be built-in capabilities exposed by IoT units (e.g., start/stop), obtained from third-parties (e.g., from public repositories or in a market-like fashion), or developed in-house to exactly reflect custom governance objectives. By promoting collaboration and early integration of governance objectives with operations capabilities, GovOps reduces the risks of lost requirements and over-regulated systems.

After the required governance capabilities and relevant governance objectives have been identified, the GovOps manager relies on the aforementioned concepts and abstractions (Sect. 4.1) to formally define the GovOps strategy and articulate the artifacts defined in the first phase of the design process. Governance capabilities are the main building blocks of the GovOps strategies. They are directly referenced in GovOps strategies to specify the concrete steps which need to be enforced on the underlying IoT cloud resources, e.g., defining a desired communication protocol or disabling a data stream for a specific region. Also in this context, the GovOps reference model does not make assumptions about the implementation of governance strategies, e.g., they can be realized as business processes, policies, applications, or domain specific languages. Individual steps, defined in the generic strategy, invoke governance capabilities that put the IoT cloud resources into desired target state, e.g., which satisfies a set of properties. Subsequently, the generic GovOps strategy needs to be parameterized, based on the concrete constraints and rules defined by the governance objectives. Depending on the strategy implementation these can be realized as process parameters, language constraints (e.g., Object Constraint Language), or application configuration directives. By formalizing the governance strategy, GovOps enables reusability of strategies, promotes consistent implementation of established standards and best practices, and ensures operation within the system's regulatory framework.

The last phase involves identifying the system resources, i.e. the governance scopes that will be affected by the GovOps strategy and executing the strategy in the IoT cloud system. It is worth mentioning that the scopes are not directly referenced in the GovOps strategies, rather the GovOps manager applies the strategies on the resource scopes. Introducing scopes at the strategy-level shields the operations team from directly referencing IoT cloud resources, thus enables designing declarative, late-bound strategies in a scalable manner. Furthermore, at this point additional capabilities identified in the previous phase will be acquired and/or provisioned, whereas unused capabilities will be decommissioned in order to optimize resource consumption.

5 Related Work

The IoT governance has been receiving a lot of attention recently. For example, in [17] the author evaluates various aspects of the IoT governance, such as privacy, security and safety, ethics, etc., and defines main principles of IoT governance, e.g., legitimacy and representation, transparency and openness, and accountability. In [16], the authors deal with issues of data quality management and governance. They define a responsibility assignment matrix that comprises roles, decision areas and responsibilities and can be used to define custom governance models and strategies. Traditional IT governance approaches, such as SOA governance [1,12] and governance frameworks like CMMI [9], the 3P model [13], and COBIT [6], provide a valuable insights and models which can be applied in GovOps processes, usually without substantial modifications. Compared to these approaches, GovOps does not attempt to define a general methodology for IoT cloud governance. Therefore, such approaches conceptually do not conflict with GovOps and can rather be seen as complementary to our approach.

Also approaches addressing operations management in IoT cloud system have recently emerged. For example, in [14,18] the authors deal with IoT infrastructure virtualization and its management on cloud, whereas [2] utilizes the cloud for additional computation resources. In [15] the authors focus on operating cloud storage resources for IoT data, and [11] present approaches for monitoring IoT systems and enforcing QoS aspects. Such approaches provide useful concepts and techniques, which can be used to support the GovOps processes in

IoT cloud systems. In [7] the authors develop an infrastructure virtualization framework, based on a content-based pub/sub model for asynchronous event exchange. In [18] the authors propose virtualizing physical sensors on the cloud and provide management and monitoring mechanisms for the virtual sensors. Such approaches provide various governance capabilities, e.g., template-based controlling of sensor groups, registering and decommissioning sensors and monitoring the QoS that can seamlessly be integrated with our GovOps approach.

The GovOps model builds on these approaches and addresses the issue of bridging the gap between governance objectives and operations processes, by introducing the GovOps manager as a dedicated stakeholder, as well as defining the suitable GovOps reference model to support early integration of governance objectives and operations processes.

6 Conclusion and Future Work

In this paper, we introduced the GovOps approach to governance of software-defined IoT cloud systems. We presented the GovOps reference model that defines suitable concepts and a flexible process to design IoT cloud governance strategies. We introduced the GovOps manager, a dedicated stakeholder responsible to determine tradeoffs between satisfying governance objectives and IoT cloud system capabilities, and ensure early integration of these objectives with operations processes, by continuously refining how the high-level objectives are articulated through operations processes. We showed how GovOps enables systematically approaching IoT cloud governance to counteract system over-regulation and lost requirements. Further, it allows for IoT cloud governance processes to be easily and flexibly realized in practice, without worrying about the complexity and scale of the underlying IoT cloud and diversities of various legal and compliance issues. In the future, in order to support GovOps managers, we will develop a comprehensive framework for GovOps that implements the presented concepts and required toolset.

Acknowledgments. This work is sponsored by Pacific Controls Cloud Computing Lab (PC3L), as well as the Austrian Science Fund under grant P23313-N23 (Audit 4 SOAs).

References

1. Charfi, A., Mezini, M.: Hybrid web service composition: business processes meet business rules. In: Second International Conference on Service-Oriented Computing - ICSOC 2004 Proceedings, pp. 30–38, New York, 15–19 November 2004
2. Chun, B., Ihm, S., Maniatis, P., Naik, M., Patti, A.: Clonecloud: elastic execution between mobile device and cloud. In: Proceedings of the Sixth European Conference on Computer systems (EuroSys 2011), pp. 301–314, Salzburg, 10–13 April 2011

3. Copie, A., Fortis, T., Munteanu, V.I., Negru, V.: From cloud governance to iot governance. In: 27th International Conference on Advanced Information Networking and Applications Workshops (WAINA 2013), pp. 1229–1234, Barcelona, 25–28 March 2013
4. DeLoach, D.: Internet of Things: Critical issues around governance for the Internet of Things. http://tinyurl.com/mxnq3ma. Accessed on July 2014
5. European Commission: Report on the public consultation on IoT governance. http://tinyurl.com/mx24d9o. Accessed on August 2014
6. Hardy, G.: Using IT governance and COBIT to deliver value with IT and respond to legal, regulatory and compliance challenges. Inf. Sec. Techn. Report **11**(1), 55–61 (2006)
7. Hassan, M.M., Song, B., Huh, E.: A framework of sensor-cloud integration opportunities and challenges. In: Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication (ICUIMC 2009), pp. 618–626, Suwon, 15–16 January 2009
8. Lantz, B., Heller, B., McKeown, N.: A network in a laptop: rapid prototyping for software-defined networks. In: Proceedings of the 9th ACM Workshop on Hot Topics in Networks. HotNets 2010, p. 19, Monterey, 20–21 October 2010
9. Lawler, B.: Review of cmmi distilled: a practical introduction to integrated process improvement. ACM SIGSOFT Softw. Eng. Notes **30**(1), 37–38 (2005). Addison Wesley, 2004, paperback. ISBN 0-321-18613-3
10. Nastic, S., Sehic, S., Le, D., Truong, H.L., Dustdar, S.: Provisioning software-defined iot cloud systems. In: 2014 International Conference on Future Internet of Things and Cloud, FiCloud 2014, pp. 288–295, Barcelona, 27–29 August 2014
11. Nef, M.A., Perlepes, L., Karagiorgou, S., Stamoulis, G.I., Kikiras, P.K.: Enabling qos in the internet of things. In: CTRQ 2012, The Fifth International Conference on Communication Theory, Reliability, and Quality of Service, pp. 33–38 (2012)
12. Niemann, M., Miede, A., Johannsen, W., Repp, N., Steinmetz, R.: Structuring SOA governance. IJITBAG **1**(1), 58–75 (2010)
13. Sandrino-Arndt, B.: People, portfolios and processes: the 3p model of it governance. Inf. Sys. Control J. **2**, 1–5 (2008)
14. Soldatos, J., Serrano, M., Hauswirth, M.: Convergence of utility computing with the internet-of-things. In: Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS 2012), pp. 874–879, Palermo, 4–6 July 2012
15. Stuedi, P., Mohomed, I., Terry, D.: Wherestore: Location-based data storage for mobile devices interacting with the cloud. In: Proceedings of the 1st ACM Workshop on Mobile Cloud Computing and Services: Social Networks and Beyond (MCS 2010), pp. 1:1–1:8. ACM, New York (2010)
16. Weber, K., Otto, B., Österle, H.: One size does not fit all-a contingency approach to data governance. J. Data Inf. Q. **1**(1), 4 (2009)
17. Weber, R.H.: Internet of things governance quo vadis? Comput. Law Secur. Rev. **29**(4), 341–347 (2013)
18. Yuriyama, M., Kushida, T.: Sensor-cloud infrastructure - physical sensor management with virtualized sensors on cloud computing. In: The 13th International Conference on Network-Based Information Systems (NBiS 2010), pp. 1–8, Takayama, September 14–16 2010

Missverständnis oder Königsweg?

Wissen spielt eine entscheidende Rolle: Wenn die Stadtbewohner ihre Daten aktiv bewirtschaften, kann Smart City eine Chance für alle sein. *Alain Thierstein*

Abstract Vernetzte Rechner, Big Data und individualisierte Mikrotechnologien verheissen die neue, bessere und ökologischere Stadt: die Smart City. Smart Buildings, Smart Homes und Smart Living wecken grosse Hoffnungen. Eine Koalition aus Big Government, Big Business, Big Advocacy hat sich innerhalb der EU aufgemacht, neue Lebensqualität zu erzeugen und die technologische Wettbewerbsfähigkeit Europas zu stärken. Der Alltag der Menschen in den städtischen Räumen sieht bis heute jedoch anders aus. Die Diskussion über die nutzerfreundliche Umsetzung und Aneignung von Smart City wird in der Nische geführt. Grosse Technologieunternehmen versprechen effizientere, integrierte Infrastruktursysteme für die Städte (Smart City 1.0); grosse Informationsverarbeiter aus dem Silicon Valley durchdringen alle Poren des Alltags (Smart City 2.0). Smart City 3.0 kann dann sein, wenn die Stadtnutzer nicht nur Datenspender sind, sondern sich auch als Produzenten von Stadt verstehen. Stadtmachen und Stadtumbauen können also zu einem gleichgewichtigeren Geben und Nehmen werden.

Be attractive! Be competitive! Be sustainable! Be smart! Wer möchte nicht in einer «Smart City» leben? Um das Konzept Smart City zu verstehen, hilft es, sich vorzustellen, wie das Leben in einer Nicht-Smart-City aussähe. Hier zeigt sich, dass es sich bei Smart City sowohl um ein funktionierendes Gesamtsystem, das unsere Zukunft neu umreisst, als auch um ein Vorhaben, das unsere Stadt lebensfähiger macht, handelt. Somit sehen die einen darin die Vorstellung eines besseren städtischen Lebens, das visionär vor uns liegt. Andere verstehen Smart City als ein Infrastruktur- und Technikprojekt, das der Stadt zur Zukunftsfähigkeit verhilft.

Städte sind die Wiegen der kreativen Menschheit. Als ökonomische sowie institutionelle Treiber sind sie unabdingbar für die Weiterentwicklung unserer Gesellschaften.[1] «Stadt» ist ein komplexes Phänomen von Netzen, das durch die Leistungsfähigkeit der technischen Infrastruktur und der dazu gehörenden sozialen Interaktion definierbar ist.[2] Im Zeitalter von vernetzten Rechnern und Big Data lassen sich die Grenzen dabei nicht mehr rein geografisch ziehen.[3] Aus der Perspektive von Akteuren steht im Vordergrund, was die Menschen aus ihrer Stadt machen: Wie wollen sie «smart» – geschickter, rücksichtsvoller, verträglicher als bisher – sein?

Eine zentrale Bedeutung kommt der Technologie zu: Durch den Einsatz von mehr vernetzter, digitaler Technologie und grossen Datenbeständen sollen unsere Städte ökonomisch produktiver und effizienter werden, mit der Folge eines geringeren ökologischen Fussabdrucks. Aus diesen Beweggründen werden in Asien und auf der Arabischen Halbinsel aktuell Stadtteile mit Smart Buildings, Smart Homes und Smart Living aus dem Boden gestampft.

Allerdings konnte die Hoffnung auf eine ökologischere Zukunft bisher nicht eingelöst werden, der Alltagsgebrauch erzählt eine andere Geschichte: Lebensqualität entsteht nicht alleine durch technologische Vernetzung und ökonomisches Wohlergehen, sondern durch die Identifikation mit der eigenen Tätigkeit, der Interaktion mit seinem persönlichen Umfeld sowie mit der Weise, wie die gebaute Umwelt den Menschen «willkommen» heisst. Diese Aneignungsprozesse gehen in der Debatte um Smart City meist vergessen.

Verkappte Industriepolitik

Ziel der EU-Kommission ist es, die Entwicklung und Implementierung intelligenter städtischer Technologien voranzutreiben. Zu diesem Zweck etablierte sie im Jahr 2012 die Initiative «Europäische Innovationspartnerschaft für Intelligente Städte und Gemeinschaften»[4], welche mit Geldern des EU-Forschungsprogramms Horizon 2020 Vorzeigeprojekte im städtischen Raum fördert.

Diese sogenannten Leuchtturmprojekte aus den Bereichen Energie, Verkehr und IT zielen alle auf eine höhere Lebensqualität. Die Hoffnung ruht dabei auf einer Reduzierung der Energiekosten für alle Akteure, auf einem beschleunigten, umweltfreundlichen Verkehr, auf einer verbesserten Mobilität, auf neuen Arbeitsplätzen, auf einer verbesserten Resilienz im Klimawandel – wie Reduktion von Hitzeinseln – sowie auf einer besseren Luftqualität.

1 Glaeser 2011, Storper 2013.
2 Bettencourt 2013, Batty 2013.
3 Kitchin 2014.
4 European Commission o. J.

DREAMSTIME

Sind wir alle nur Datenspender? Metro in Singapur.

Ein starker Treiber im Hintergrund ist Industriepolitik. In der EU scheint Smart City eine magische Verbindung von wichtigen Zielen zu bieten: Die internationalen Verpflichtungen zur CO_2-Reduktion kombiniert sie mit überfälligen Erneuerungen der bisherigen technischen Infrastrukturen, der Herausforderung der Digitalisierung aller Lebenswelten, also des Internets der Dinge, sowie der Wettbewerbsposition europäischer Technologiefirmen.

Die Begeisterung für Smart City erklärt sich auch durch die nach wie vor untergeordnete Rolle, welche Forschungs-, Technologie- und Innovationspolitik – im Gegensatz zu Agrarpolitik und Strukturhilfe für schwächere Regionen – spielt. Zudem erhofft man sich davon, gegenüber aufstrebenden Regionen in Asien den Anschluss halten zu können.

Die noch junge Geschichte von Smart City bietet drei Lesarten, die im Folgenden erläutert werden.

Smart City 1.0 als Strategie von Institutionen

Smart City startete als Projekt von Eliten. Die Akteure stammen dabei vorwiegend aus der zentralisierten Verwaltung, aus der leistungsstarken Industrie und aus Interessengruppen: Big Government, Big Business, Big Advocacy. Die zentralen wirtschaftlichen Interessen werden besonders deutlich, wenn man versteht, dass die Schaffung eines digitalen Binnenmarktes eine von zehn Prioritäten von Kommission, Rat und Parlament in Brüssel darstellt.

Vor diesem Hintergrund liest sich Smart City 1.0 als «Durchbruchsstrategie» von Grossinstitutionen, die einerseits in der Lage sind, ihre Steuerungs- und Regulationskompetenz einzusetzen, und anderseits über die Erfahrung verfügen, ihre Kompetenzen zu Technologieentwicklung, Systemintegration sowie Finanzierung und Betreibermodellen für Städte zu bündeln. Eine neue Runde von Public-Private-Partnerships wird damit eingeläutet, welche die Fähigen – Technologie- und Regulationsanbieter – zusammenbringen mit den Willigen – den von strukturellen Haushaltsschwächen gezeichneten Stadtkommunen in Europa.

Als Normalbürger könnten wir uns entspannt zurücklehnen und sarkastisch sagen, bei Smart City handle es sich nur um ein weiteres von «oben» geplantes, angebotsorientiertes Infrastrukturprojekt, das «unten» kaum mit nennenswerten Effekten ankomme. Wir alle haben aber Anlass, genauer hinzuschauen, denn die digitale

Durchdringung unseres Alltages bietet bereits ausreichend Anschauungsmaterial (siehe *Kasten*).

Ein ausschliesslich technologiebezogener Fokus verstellt aber den Blick für Herausforderungen und Chancen urbaner Transformation. Denn «Stadt» hat viele Gesichter und wird letztlich gemacht durch Interaktion. Je nach disziplinärer Herkunft lesen wir das Städtische anders: morphologisch-siedlungsstrukturell, funktional-wirtschaftlich, sozial-demografisch, institutionell-politisch, relational-vernetzt oder transformationsorientiert.

Stadtmachen wird zum Abenteuer des 21. Jahrhunderts.

Smart City 1.0 ist als Grundkonzept auf die Physis der Stadtfunktionen ausgerichtet. Grosse Technologiefirmen des Hard- und Softwaresektors wie ABB, Siemens, IBM oder Cisco versprechen mit ihren Systemlösungen signifikante Skalenvorteile und Wirkungssprünge. Städte wie Wien bündeln ihre urbanen Technologien in Dienstleistungsagenturen.[5] Dabei zeigt sich: Die 24-Stunden-Stadt funktioniert nur mit einem leistungsstarken, hoch vernetzten infrastrukturellen Rückgrat – «Stadtmachen» wird insgesamt zum Abenteuer des 21. Jahrhunderts.

Google & Co. dominieren bei Smart City 2.0

Smart City 2.0 fügt neu die «Systemmanager» hinzu, deren Kernkompetenz im Sammeln von Daten und dem Verarbeiten zu Information liegt: Google, Apple, Facebook, Amazon und andere lassen uns erahnen, dass sie die «Master of Ceremony» sein können und das Internet der Dinge letztlich die Stadt der Zukunft bestimmt. Diese Internetkonzerne verknüpfen die sich weiter ausdifferenzierenden Teile des Internets der Dinge schrittweise zu neuen, räumlich wirksamen Wertschöpfungsketten.

Für die Nutzer von Smart City 2.0 erscheinen diese neuen Angebotssysteme zunächst attraktiv, denn sie zahlen geringe Grenzkosten für Netzwerk-Dienstleistungen. Zugleich sind sie aber die zentralen individuellen Lieferanten von Bewegungs- und Attributsdaten und werden euphemistisch als «Datenspender» bezeichnet.

Smart City 3.0: Wissen entsteht in Interaktion

Chancen bestehen, dass sich daraus Smart City 3.0 entwickeln könnte. Neuartige Nutzerkooperationen entwickeln sich als Public-Private-Partnerships, wo das Geben und Nehmen von Daten und Leistungen ausgeglichener als heute erfolgt. Smart City 3.0 bedingt, dass wir Stadt als Phänomen der sozialen Interaktion begreifen.

Die relational verstandene Stadt hat sich seit je durch Austausch zwischen Akteuren und Standorten transformiert.[6] Als Ergebnis entsteht Wissen. Es ist das Kapital der Wissensökonomie, die wir als jenen wachsenden Teil der Wirtschaft definieren, der hoch spezialisiertes Wissen und Fähigkeiten aus verschiedenen Segmenten der Wertschöpfungskette strategisch miteinander kombiniert, um Innovation zu ermöglichen und Konkurrenzvorteile zu wahren.[7]

Wirtschaften ist ein ständiger Neuerungsprozess, der im Begriff Innovation zusammengefasst ist. Als strategischer Wettbewerbsfaktor spielt Wissen hier eine entscheidende Rolle – etwa bei der Steuerung von unternehmerischen Wertschöpfungsketten. Das Produzieren von Wissen ist heute ein zentraler Treiber der räumlichen Transformation von Stadt. In den Worten das US-Stadtökonomen Ed Malecki: «If knowledge is not found everywhere, then where it is located becomes a particularly significant issue.»[8]

5 Vgl. Tinavienna.at
6 Taylor 2013.
7 Bentlage, Thierstein und Lüthi 2014.
8 Malecki 2000.:110

Deutschland geht voraus

Die deutsche Regierung treibt Smart City voran. Laut dem *Bundesinstitut für Bau-, Stadt- und Raumforschung* dient in erster Linie die dabei eingesetzte Informations- und Kommunikationstechnologie (ICT) dazu, «die stadtplanerischen Leitbilder im Bereich Klima, Verkehr und Mobilität, Verwaltungsmodernisierung, Daseinsvorsorge und öffentliche Sicherheit umzusetzen».[a]

Ein Beispiel aus *München:* Hier untersucht ein weiteres Forschungsteam im Projekt «Gamification, Prognosemärkte, Wikis & Co: Neues Wissen für die Stadt?»[b] innovative Ansätze der Wissens- und Entscheidungsfindung, die Herausforderungen und Zukunftsaufgaben in den Kommunen begegnen helfen. Die Autoren versprechen sich davon neue Ansätze zur Sicherung der kommunalen Grundversorgung, zur Stärkung der Transparenz in Verwaltungsabläufen, für neue Möglichkeiten des Monitorings räumlicher Entwicklungsdynamiken, für ein verbessertes Risikomanagement komplexer urbaner Systeme, für neue Wege, dezentrales Steuerungswissen in der Kommune zu bündeln sowie für vertiefte Kenntnisse über wichtige Zielgruppen für die kommunale Entwicklung und Möglichkeiten deren gezielter Ansprache und Aktivierung.

a BBSR (2015b):7.
b BBSR (2015a).

Informations- und Kommunikationstechnologien sind verantwortlich für die gesteigerten Möglichkeiten der Entstehung und Nutzung von Wissen. Distanzunabhängige Technologien bedeuten aber mitnichten das Ende von lokalen Qualitäten. Vielmehr verdeutlichen empirische Arbeiten das scheinbare Paradoxon der Komplementarität von räumlicher und relationaler Nähe für die Wissensgenerierung. Wissen wird erst im Austausch zwischen Menschen geschaffen, die sich sowohl räumlich als auch in Netzwerken nahe und vertraut sowie gleichzeitig zu diesem Austausch bereit und in der Lage sind. Der systematische Einsatz von Wissen im Sinne einer Kombination von wissenschaftlichem und erfahrungsgestütztem Wissen im unternehme-

Songdo City in Südkorea. Kritiker bemängeln, die Stadt sei ohne die Bewohner geplant worden.

SHUTTERSTOCK

rischen Wertschöpfungsprozess trägt heute wesentlich zur Transformation des Raumes bei.[9]

Selbstreflexion der Stadtbewohner entscheidend

Das menschliche Bedürfnis des Austausches, Sich-Begegnens und Erkannt-Werdens kann als Konstante der Raumnutzung betrachtet werden: Das ist meine Hypothese. Die Smart City wird uns nur dann die erhofften Wirkungen bescheren, wenn smart nicht nur die Eigenschaft von Menschen oder Institutionen meint, sondern vor allem auch die Anordnung und Organisation von gebauter und nicht bebauter Umwelt. Ziel sollte sein, Smart City 3.0 als Instrument der Steuerung von Stadttransformation zu begreifen.

Das Konzept umfasst zwei Lesarten: Erstens führt die Implementierung zu besserer Performance der Städte. Zweitens muss diese von Nutzern reflektiert werden können und zu neuem Wissen über räumliche Entwicklungsdynamik selber führen, beispielsweise über die räumlichen Konsequenzen des eigenen Nutzerverhaltens. Solches Wissen wiederum ist Grundlage für eine wirkungsvollere Steuerung der Städte.

Zielt politisches Handeln letztlich auf wirkungsorientierte Stadttransformation[10] – also auf absichtsvolle Kenntnisse von Ursache-Wirkungs-Zusammenhängen –, dann dient Smart City der gesellschaftlichen Verständigung über Wissen. Smart City 3.0 könnte der Königsweg darstellen und die Chance bieten, den Stadtbewohner wieder als Master of Ceremony ins Stadtmachen zurückzubringen. Wir müssen uns gleichzeitig als Produzent und Nutzer unserer eigenen steuerungsrelevanten Information begreifen. Indem wir Steuerungswissen öffentlich herstellen, lässt sich die Entwicklung von Stadt wirkungsvoller diskutieren.

9 Bentlage, Thierstein und Lüthi 2014.
10 Förster 2014.

Alain Thierstein
Dr. oec. HSG, Professor für Raumentwicklung, Fakultät für Architektur, Technische Universität München, Partner beim Beratungsunternehmen Ernst Basler und Partner, Zürich

Literatur

Batty, Michael (2013). The New Science of Cities. Boston: MIT Press.
BBSR, Bundesinstitut für Bau-, Stadt- und Raumforschung (2015a): Smart Cities – Gamification, Prognosemärkte, Wikis & Co: Neues Wissen für die Stadt. *Weblink auf Dievolkswirtschaft.ch*
BBSR, Bundesinstitut für Bau-, Stadt- und Raumforschung (2015b). Smart Cities International – Strategien, Strukturen und Pilotvorhaben, Eva Schweitzer (Hrsg.). Bonn: BBSR, Bundesinstitut für Bau-, Stadt- und Raumforschung.
Bentlage, Michael, Alain Thierstein und Stefan Lüthi (2014). Knowledge Hubs: Poles of Physical Accessibility and Non-physical Connectivity. In: Ben Derudder, Sven Conventz, Alain Thierstein und Frank Witlox (Hrsg.): Hub Cities in the Knowledge Economy. Seaports, Airports, Brainports. London: Ashgate, 31–54.
Bettencourt, Luis M. A. (2013). The Origins of Scaling in Cities. In: Science 340(1438), 1438–1441.
European Commission (o. J.). Smart Cities and Communities. The European Innovation Partnership on Smart Cities and Communities. Operational Implementation Plan: First Public Draft. *Weblink auf Dievolkswirtschaft.ch*
Förster, Agnes (2014). Planungsprozesse wirkungsvoller gestalten – Wirkungen, Bausteine und Stellgrößen kommunikativer planerischer Methoden. Doktor der Ingenieurwissenschaften (Dr-Ing), Fakultät für Architektur, Lehrstuhl für Raumentwicklung. München: Technische Universität. *Weblink auf Dievolkswirtschaft.ch*
Glaeser, Edward (2011). Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier. New York: The Penguin Press.
Kitchin, Rob (2014). Big Data, New Epistemologies and Paradigm Shifts. In: Big Data & Society 1(1).
Malecki, Edward J. (2000). Creating and Sustaining Competitiveness. Local Knowledge and Economic Geography. In: John Bryson, Peter Daniels, Nick Henry und Jane Pollard (Hrsg.): Knowledge, Space, Economy. London, New York, 103–119.
Storper, Michael (2013). Keys to the City. How Economics, Institutions, Social Interactions, and Politics Shape Development Oxford: Princeton University Press
Taylor, Peter J. (2013). Extraordinary Cities: Millennia of Moral Syndrome, World-Systems and City/State Relations. Edward Elgar Pub.

Malentendu ou voie royale ?

Le savoir joue un rôle décisif. Si les citadins gèrent activement leurs données, la ville « intelligente » peut être une chance pour tous. *Alain Thierstein*

Abrégé Ordinateurs connectés, mégadonnées (« big data ») et microtechnologies personnalisées promettent une ville nouvelle, meilleure et plus écologique. Bâtiments, maisons et mode de vie deviendraient « intelligents », suscitant par là de grands espoirs. Une coalition de « big government », « big business » et « big advocacy » s'est constituée au sein de l'UE pour générer une nouvelle qualité de vie et renforcer la compétitivité de l'Europe en matière de technologie. Or, jusqu'à présent, le quotidien des citadins revêt un tout autre aspect. La discussion sur la mise en œuvre et l'appropriation de la ville « intelligente » se poursuit en circuit fermé. Les grandes entreprises de technologie promettent aux villes des infrastructures intégrées plus efficaces (« smart city » 1.0). Les grands informaticiens de la Silicon Valley s'infiltrent dans tous les pores de la vie quotidienne (« smart city » 2.0). Le stade « smart city » 3.0 sera atteint le jour où les utilisateurs de la ville ne seront plus seulement des fournisseurs de données, mais seront aussi assimilés à des producteurs d'urbanité. Faire et transformer la ville pourra alors devenir une forme d'échanges plus équilibrés entre offreurs et preneurs.

Attractive, compétitive, durable, « intelligente » : qui ne souhaiterait vivre dans une « smart city » ? Pour comprendre cette notion, il est utile d'imaginer de quoi la vie aurait l'air dans une ville qui ne le serait pas. Il s'avère alors que cette forme d'« intelligence » est aussi bien un système opérationnel global dessinant notre nouvel avenir qu'un projet qui rend notre ville plus viable. Les uns croient qu'elle améliorera leur vie. Les autres la considèrent comme un projet d'infrastructures et de techniques qui aidera la cité à affronter l'avenir.

Les villes sont les berceaux de l'humanité créative. Ce sont des moteurs économiques et institutionnels indispensables à l'évolution de nos sociétés[1]. Elles abritent des réseaux complexes, que l'on peut définir comme le produit des infrastructures techniques et des interactions sociales concomitantes[2]. À l'époque du maillage des ordinateurs et des mégadonnées (« big data »), les limites ne peuvent plus être fixées de manière purement géographique[3]. Du point de vue des acteurs, la priorité va à ce que les gens font de leur ville : comment veulent-ils devenir plus habiles, plus prévenants, plus tolérants, donc « smart » ?

La technologie revêt ici une importance capitale. En recourant à davantage de technologie numérique connectée et à de grands stocks de données, nos villes amélioreront leur productivité économique et leur efficacité, tout en réduisant leur empreinte écologique. C'est pour ces raisons que surgissent actuellement en Asie et dans la péninsule Arabique des quartiers de « smart buildings », « smart homes » et « smart living ».

L'espoir d'un avenir plus écologique ne s'est, toutefois, pas réalisé. La routine quotidienne est toute autre : la qualité de vie ne naît pas uniquement du maillage technologique et du bien-être économique, mais de l'identification à sa propre activité, des interactions avec son entourage personnel et de la manière dont l'environnement construit accueille l'homme. Or, ces processus d'appropriation sont la plupart du temps oubliés lors des débats que suscite la ville « intelligente ».

Une politique industrielle camouflée

La Commission européenne s'est fixée pour but de faire progresser le développement et la mise en œuvre de technologies urbaines « intelligentes ». À cet effet, elle a lancé en 2012 un *Partenariat d'innovation pour les villes et communautés intelligentes*[4], qui soutient des « projets de démonstration » en milieu urbain avec des fonds provenant du programme de recherche *Horizon 2020.*

Ces projets dits « phares » dans les secteurs de l'énergie, des transports et des TIC visent tous à améliorer la qualité de la vie. On espère ainsi réduire les coûts de l'énergie pour l'ensemble des acteurs, accélérer les transports et les rendre plus respectueux de l'environnement, améliorer la mobilité, créer de nouveaux emplois, renforcer la résilience face au changement climatique (réduction des îlots de chaleur urbains, par exemple) et améliorer la qualité de l'air.

La politique industrielle constitue un puissant moteur à cet égard. Dans l'UE, la ville « intelligente » paraît offrir une combinaison magique d'objectifs importants : engagements internatio-

1 Glaeser (2011), Storper (2013).
2 Bettencourt (2013), Batty (2013).
3 Kitchin (2014).
4 Commission européenne (s.d.)

Sommes-nous tous des fournisseurs de données ? Pendulaires à Singapour.

naux de réduction du CO_2, renouvellement trop longtemps ajourné des vieilles infrastructures techniques, défi de la numérisation de tous les domaines de l'existence (Internet des objets), compétitivité des compagnies européennes de technologie, etc.

L'enthousiasme pour la ville « intelligente » s'explique aussi par le rôle toujours subalterne que joue la politique de la recherche, de la technologie et de l'innovation – contrairement à la politique agricole et à l'aide accordée aux régions structurellement faibles. On espère également pouvoir ainsi rester dans la course avec les régions émergentes d'Asie.

L'histoire encore récente de la « smart city » se décline en trois étapes.

« Smart city » 1.0, une stratégie d'institutions

La ville « intelligente » a d'abord été imaginée par les élites. Ses concepteurs proviennent dans une large mesure d'administrations centralisées, d'industries de pointe et de grands groupes d'intérêts (« big government », « big business », « big advocacy »). La prédominance des intérêts économiques est particulièrement manifeste si l'on songe qu'à Bruxelles, la création d'un marché intérieur numérique est l'une des dix priorités de la Commission, du Conseil et du Parlement.

Dans ce contexte, « smart city » 1.0 peut être vu comme une stratégie adoptée par les grandes institutions pour percer. Celles-ci sont, d'une part, en mesure d'engager leur pouvoir de pilotage et de régulation. Elles disposent, d'autre part, de l'expérience nécessaire pour regrouper en faveur des villes leurs compétences en matière de développement technologique, de systèmes intégrés, de financement et de modèles de gestion. Un nouveau cycle de partenariats public-privé est ainsi lancé : il rassemble ceux qui peuvent – les fournisseurs de technologie et de régulation – et ceux qui veulent – les communes citadines d'Europe frappées de faiblesse budgétaire structurelle.

En tant que citoyens lambda, nous pourrions nous caler confortablement dans nos fauteuils et déclarer de manière sarcastique que la ville « intelligente » est seulement un nouveau projet d'infrastructures planifié « d'en haut » et axé sur l'offre, qui n'aurait guère d'effets sensibles « en bas ». Or, nous avons tout intérêt à y regarder de plus près, car l'imprégnation numérique de notre vie quotidienne offre déjà suffisamment de matière à réflexion (voir *encadré).*

Ce regard exclusivement technologique brouille, cependant, notre perception des défis et des chances qu'offre la transformation urbaine. La ville a, en effet, plusieurs visages et résulte d'interactions. Selon notre discipline d'origine, nous la voyons sous différents angles : morphologique/structurel, fonctionnel/économique, social/démographique, institutionnel/politique, relationnel/connecté ou transformiste.

La notion fondamentale de « smart city » 1.0 est ainsi liée à l'aspect physique des fonctions urbaines. De grandes compagnies de matériel et de logiciels, comme ABB, Siemens, IBM ou Cisco, promettent avec leurs systèmes d'importants gains d'échelle et d'efficacité. Des villes comme Vienne regroupent leurs technologies urbaines au sein d'agences de services[5]. Il en ressort que la ville ne peut fonctionner 24 heures sur 24 qu'en étant équipée d'une colonne vertébrale performante constituée d'infrastructures fortement connectées. « Construire la ville » devient ainsi l'aventure du XXI[e] siècle.

« Construire la ville » devient l'aventure du XXI[e] siècle.

« Smart city » 2.0 : la domination de Google & Co.

« Smart city » 2.0 ajoute une nouvelle couche à la précédente : les gestionnaires de systèmes, dont la compétence clé consiste à collecter des données et à les transformer en informations. Google, Apple, Facebook, Amazon et autres laissent entrevoir leur capacité à être « maîtres de cérémonie ». Pour eux, le destin des villes du futur sera lié à l'Internet des objets. Ces grands groupes unissent progressivement les ramifications croissantes de l'Internet des objets pour en faire de nouvelles chaînes de valeur ajoutée territoriales.

Pour les utilisateurs de « Smart city » 2.0, ces nouvelles offres semblent d'abord attrayantes, car les services connectés voient leurs coûts marginaux baisser. Or, ce sont eux qui centralisent et livrent les données concernant les déplacements et les caractéristiques de ces mêmes utilisateurs, d'où l'euphémisme « fournisseurs de données ».

« Smart city » 3.0 : le savoir, fruit de l'interaction

Il existe des chances pour que « smart city » 3.0 se développe à son tour. De nouvelles coopérations entre utilisateurs voient le jour sous la forme de partenariats public-privé qui permettent des échanges de données et de prestations plus équilibrés qu'aujourd'hui. « Smart city » 3.0 nécessite que nous voyions la ville comme un théâtre d'interactions sociales.

Vue sous l'angle relationnel, la ville a toujours profité des échanges entre acteurs et lieux pour se transformer[6]. Il en a résulté un cumul de connaissances, qui sous-tend l'« économie du savoir ». Nous entendons celle-ci comme la partie croissante de l'économie qui combine stratégiquement le savoir très spécialisé et des capacités provenant de différents segments de la chaîne de valeur ajoutée pour permettre l'innovation et garantir des avantages compétitifs[7].

L'économie est un processus de renouvellement constant, résumé dans la notion d'innovation. Le savoir, qui est un facteur stratégique de concurrence, joue ici un rôle décisif, par exemple pour piloter des chaînes entrepreneuriales de valeur ajoutée. La production de savoir est aujourd'hui un moteur crucial de la transformation spatiale des villes. Pour reprendre les termes de

5 Voir Tinavienna.at.
6 Taylor (2013).
7 Bentlage, Thierstein et Lüthi (2014).
8 Malecki (2000), p. 110.

L'Allemagne pionnière

Le gouvernement allemand fait avancer la ville « intelligente ». D'après l'Institut fédéral de recherches sur la construction, l'urbanisme et le territoire, les technologies de l'information et de la communication (TIC) ont pour but de « mettre en œuvre des modèles d'urbanisme en matière de climat, de transports et mobilité, de modernisation de l'administration, de prévoyance et de sécurité publique[a] ».

Autre exemple, cette fois à Munich. Dans le cadre du projet *Ludification, marchés de la prédiction, Wikis & Co. : de nouveaux savoirs pour la ville ?*[b], une équipe de chercheurs étudie des approches novatrices en matière de connaissance et de prise de décision, afin d'aider les communes à affronter les défis et les tâches du futur. Les auteurs en attendent de nouvelles pistes notamment pour assurer les services communaux de base, renforcer la transparence des procédures administratives, innover dans la surveillance de la dynamique du territoire et améliorer la gestion des risques dans les systèmes urbains complexes. Il faudra également développer de nouveaux moyens de regrouper au niveau communal le savoir décentralisé en matière de pilotage, approfondir les connaissances concernant des groupes cibles importants pour le développement communal et trouver des possibilités de les interpeller et de les activer spécifiquement.

a BBSR (2015b), p. 7.
b BBSR (2015a).

l'économiste américain Edward Malecki, « si la connaissance ne se trouve pas n'importe où, alors l'endroit où elle se trouve devient une question particulièrement importante[8] ».

Les technologies de l'information et de la communication permettent de plus en plus de générer du savoir et de l'utiliser. Ces technologies indépendantes de la distance ne signifient, toutefois, en rien la fin des atouts locaux. Au contraire, des études empiriques mettent en lumière le paradoxe apparent que constitue la complémentarité de la proximité spatiale et relationnelle pour la genèse du savoir. Ce dernier naît, en effet, uniquement des échanges entre personnes qui sont proches et familières aussi bien dans l'espace qu'à travers les réseaux, mais qui sont aussi

Songdo en Corée du Sud. Pour certains, la ville a été conçue sans la population.

SHUTTERSTOCK

prêtes et équipées pour ces échanges. Recourir systématiquement au savoir – conçu comme une combinaison de connaissances scientifiques et empiriques – lors de la création de valeurs économiques contribue aujourd'hui de façon décisive à la transformation du territoire[9].

Caractère décisif de l'autoréflexion des citadins

Le besoin humain d'échanges, de rencontres et de reconnaissance peut être considéré comme une constante de l'utilisation du territoire. Telle est du moins mon hypothèse. On n'obtiendra donc les effets espérés que si l'« intelligence » n'est pas la seule qualité des personnes ou des institutions. Elle doit concerner, en premier lieu, l'ordonnancement et l'organisation de l'environnement bâti et non bâti. Le but devrait être de comprendre la « smart city » 3.0 comme un outil de pilotage de la transformation urbaine.

Cette notion débouche sur deux constats. Premièrement, sa mise en œuvre améliorera la performance des villes. Deuxièmement, il faut que les utilisateurs puissent y réfléchir et que cela aboutisse à une nouvelle connaissance de la dynamique territoriale, par exemple des conséquences que le comportement de chaque utilisateur peut avoir au niveau de l'espace. Ce savoir sera à son tour la base d'un pilotage plus efficace des villes.

Si l'action politique vise en fin de compte une transformation urbaine axée sur les impacts[10] – soit la connaissance intentionnelle des liens de cause à effet –, la ville intelligente aidera la société à s'entendre sur ce qu'est le savoir. La « smart city » 3.0 pourrait être une voie royale en ce domaine. Elle offre au citadin la possibilité de redevenir « maître de cérémonie » dans la construction de sa ville. Nous devons considérer que nous sommes à la fois les producteurs et les utilisateurs de nos propres informations en matière de pilotage. En créant publiquement la science du pilotage, nous permettrons une discussion plus efficace du développement urbain.

9 Bentlage, Thierstein et Lüthi (2014).
10 Förster (2014).

Alain Thierstein
Professeur de développement territorial, faculté d'architecture, université technique de Munich, partenaire de l'entreprise de conseil Ernst Basler et partenaires, Zurich

Bibliographie

Batty Michael, *The New Science of Cities*, Boston, 2013, MIT Press.
Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR), *Smart Cities – Gamification, Prognosemärkte, Wikis & Co: Neues Wissen für die Stadt*, 2015a (le lien se trouve sur le site de *La Vie économique)*.
Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR), *Smart Cities International – Strategien, Strukturen und Pilotvorhaben*, Bonn, 2015b, Eva Schweitzer (éd.), BBSR.
Bentlage Michael, Thierstein Alain et Lüthi Stefan, « Knowledge Hubs: Poles of Physical Accessibility and Non-physical Connectivity », dans Ben Derudder, Sven Conventz, Alain Thierstein et Frank Witlox (éds), *Hub Cities in the Knowledge Economy. Seaports, Airports, Brainports*, Londres, Ashgate, 2014, pp.31–54.
Bettencourt Luis M. A., « The Origins of Scaling in Cities », *Science*, 340 (1438), 2013, pp. 1438–1441.
Commission européenne, *Smart Cities and Communities. The European Innovation Partnership on Smart Cities and Communities. Operational Implementation Plan: First Public Draft*, s.d. (le lien se trouve sur le site de *La Vie économique)*.
Förster Agnes, *Planungsprozesse wirkungsvoller gestalten – Wirkungen, Bausteine und Stellgrößen kommunikativer planerischer Methoden*, faculté d'architecture, chaire de développement territorial, université technique de Munich, 2014 (le lien se trouve sur le site de *La Vie économique)*.
Glaeser Edward, *Triumph of the City: How our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*, New York, 2011, The Penguin Press.
Kitchin Rob, « Big Data, new epistemologies and paradigm shifts », *Big Data & Society* 1(1), 2014.
Malecki Edward J., « Creating and Sustaining Competitiveness. Local Knowledge and Economic Geography », dans John Bryson, Peter Daniels, Nick Henry et Jane Pollard (éd.), *Knowledge, Space, Economy*, Londres, New York, 2000, pp. 103–119.
Storper Michael, *Keys to the City. How economics, institutions, social interactions, and politics shape development*, Oxford, 2013, Princeton University Press.
Taylor Peter J., *Extraordinary Cities: Millennia of Moral Syndrome, World-Systems and City/State Relations*, 2013, Edward Elgar Pub.

The answer is “Smart” – But what was the question?

Why smart cities have no discourse, but are in need of a Utopian patroness, which protects them from digital feudalism 500 years after Utopia.

Many cities around the globe are in the process of realizing the real utopia of the smart city. The paper claims that contemporary arguments communicated by the information technology industries and renown scholars alike are stressing elements of the utopian tradition. Hence the paper is discussing arguments about the smart city by the IBM representatives, Saskia Sassen, Rem Koolhaas and Adam Greenfield, among others. This discussion is challenged by arguments raised by Thomas Morus in his Utopia.

The so called smart city discourse is unveiled to be mere propaganda by the IT industries in successfully constructing a powerful floating signifier with the term smart city. The major drive underlying Mores argumentation is read as the claim for common wealth.

While propaganda is claiming to deliver neutral technological optimization for cities, smart city technology is affecting civil rights issues without any discussion. This development might contribute to the reestablishment of social structures comparable with feudalism, but enhanced by digital means.

In concluding the paper is raising two claims, one against using techno-centric values as premises for socio-spatial processes, the other, for the development of an involved and democratic narration to accompany the phenomena of urbanization while focusing on common wealth.

Regional Studies, Vol. 36.9, pp. 957–975, 2002

Carfax Publishing
Taylor & Francis Group

A New Map of Hollywood: The Production and Distribution of American Motion Pictures

ALLEN J. SCOTT

Center for Globalization and Policy Research, School of Public Policy and Social Research, UCLA, Los Angeles, CA 90095, USA. Email: ajscott@ucla.edu

(Received September 2001; in revised form December 2001)

SCOTT A. J. (2002) A new map of Hollywood: the production and distribution of American motion pictures, *Reg. Studies* **36**, 957–975. In this paper, I offer a reinterpretation of the economic geography of the so-called new Hollywood. The argument proceeds in six main stages. First, I briefly examine the debate on industrial organization in Hollywood that has gone on in the literature since the mid-1980s, and I conclude that the debate has become unnecessarily polarized. Second, I attempt to show how an approach that invokes both flexible specialization and systems-house forms of production is necessary to any reasonably complete analysis of the organization of production in the new Hollywood. Third, and on this basis, I argue that the Hollywood production system is deeply bifurcated into two segments comprising: (1) the majors and their cohorts of allied firms on the one hand; and (2) the mass of independent production companies on the other. Fourth, I reaffirm the continuing tremendous agglomerative attraction of Hollywood as a locale for motion-picture production, but I also describe in analytical and empirical terms how selected kinds of activities seek out satellite production locations in other parts of the world. Fifth, I show how the majors continue to extend their global reach by means of their ever more aggressive marketing and distribution divisions, and I discuss how this state of affairs depends on and amplifies the competitive advantages of Hollywood. Sixth and finally, I reflect upon some of the challenges that Hollywood must face up to as new cultural-products agglomerations arise all over the globe, offering potential challenges to its hegemony.

Motion-picture industry Cultural economy Hollywood Agglomeration Regional development Globalization

SCOTT A. J. (2002) Une nouvelle carte de Hollywood: la production et la distribution des films américains, *Reg. Studies* **36**, 957–975. Cet article cherche à remettre en question la géographie économique du soi-disant nouvel Hollywood. Le raisonnement se déroule en six étapes. Premièrement, on examine le débat sur l'organisation industrielle à Hollywood qui a eu lieu dans la litérature depuis le milieu des années 80, et on affirme que le débat s'est polarisé inutilement. Deuxièmement, on essaie de démontrer comment une approche qui invoque à la fois la spécialisation flexible et des formes de production dites 'system houses' est necessaire à l'analyse de l'organisation de la production dans le nouvel Hollywood. Troisièmement, on affirme que le système de production à Hollywood se divise nettement en deux parties; à savoir, (1) d'un côté les 'majors' et leurs cohortes d'entreprises alliées et (2) de l'autre côté la masse de firmes indépendantes. Quatrièmement, on affirme de nouveau la grande tendance continuelle à l'agglomération de Hollywood comme un endroit ou on product des films. On décrit aussi, en termes analytiques et empiriques, comment certaines activités recherchent des lieux de production délocalisés sur le plan mondial. Cinquièmement, on montre comment les 'majors' continuent d'étendre leur portée mondiale par moyen de leurs services de marketing et de distribution de plus en plus aggressifs. On examine aussi comment cette situation dépend de et amplifie les avantages compétitifs de Hollywood. Sixièmement, on réfléchit à quelques-uns des défis auxquels Hollywood devra faire face au fur et à mesure

SCOTT A. J. (2002) Eine neue Hollywoodkarte: Herstellung und Verteilung amerikanischer Spielfilme, *Reg. Studies* **36**, 957–975. Dieser Aufsatz stellt eine Neuinterpretation der Wirtschaftsgeographie des sogenannten neuen Hollywood vor. Zuerst wird die Debatte um die industrielle Organisation, die die Literatur seit Mitte der achtziger Jahre beschäftigt, kurz untersucht, mit der Schlußfolgerung, daß sie unnötig polarisiert worden ist. Danach wird versucht, zu zeigen, inwiefern eine einigermaßen vollständige Analyse der Produktionsorganisation im neuen Hollywood eine Einstellung verlangt, die sowohl an flexible Spezialisierung als auch an 'Systems-house' Produktionsformen appelliert. Drittens, und auf dieser Basis, wird der Standpunkt vertreten, daß das Produktionssystem von Hollywood stark gegabelt ist, wobei die beiden Zinken a) einerseits 'die Großen' und ihre Kohorten verbündeter Firmen, und b) andrerseits die breite Masse unabhängiger Produktionsgesellschaften in sich vereinigen. Viertens wird bestätigt, daß die enorme agglomerative Anziehungskraft Hollywoods als Ort for Spielfilmproduktion weiter anhält, aber auch beschrieben, wie, analytisch und empirisch gesehen, bestimmte Arten von Tätigkeiten Satellitenproduktionsorte in anderen Teilen der Welt aufsuchen. Fünftens wird gezeigt, wie 'die Großen' ihren globalen Einflußbereich durch immer aggressiveres Marketing und Verteilerabteilungen weiterausdehnen, und erörtert, inwiefern diese Zustände zugleich von den Wettbewerbsvorteilen Hollywoods abhängen und sie verstärken. Sechstens, schließlich, folgen Überlegungen zu

0034-3404 print/1360-0591 online/02/090957-19 ©2002 Regional Studies Association DOI: 10.1080/0034340022000022215
http://www.regional-studies-assoc.ac.uk

que de nouvelles agglomérations de produits culturels se multiplient partout dans le monde, ce qui représente un défi à son hégémonie.

Industrie cinématographique Economie culturelle Hollywood Agglomération Aménagement du territoire Mondialisation

Herausforderungen, mit denen Hollywood konfrontiert wird, da auf der ganzen Welt neue Kulturproduktballungen entstehen, und seine Vorherrschaft durchaus herausfordern könnten.

Spielfilmindustrie Kulturwirtschaft Hollywood Ballung Regionale Entwicklung Globalisierung

INTRODUCTION

Some time in the 1980s, entertainment industry analysts began to refer more and more insistently to a so-called 'new Hollywood', in contradistinction to the old Hollywood that had thrived over the pre-war decades on the basis of the classical studio system of production (see SCHATZ, 1983; LITWAK, 1986; GOMERY, 1998; SMITH, 1998). This new Hollywood emerged slowly and painfully out of the profound restructuring of the old studios that occurred from the 1950s to the 1970s, and that finally resulted not only in a new business model but also in a new aesthetics of popular cinema. Over the last two decades of the 20th century and on into the 21st century, there has been a complex deepening and widening of the trends first recognized in terms of the new Hollywood, and the present paper is a modest effort to shed some light on the way these processes are currently working themselves out.

Hollywood has always been identified, in one of its principal representations, as a disembodied bundle of images. But it is, too, a distinctive geographic phenomenon, which, right from its historical beginnings a century ago, has assumed the form of a dense agglomeration of motion-picture production companies and ancillary services, together with a peculiar local labour market, within the wider context of Los Angeles (or, more generally, Southern California).[1] This persisting geographic base has been the arena of many and perplexing transformations over the last few decades. In addition to the break-up of the old studio system, five principal changes have had particularly strong impacts. These are:

1. The penetration of new computerized technologies into all stages of the motion-picture production and distribution process
2. The steady bifurcation (as I shall argue) of the Hollywood production system into makers of high-concept blockbuster films on the one side, and more modest independent filmmakers on the other
3. The intensifying geographic decentralization of film-shooting activities away from the core complex of Hollywood
4. The proliferation of new markets based on the packaging and repackaging of intellectual property rights
5. The merging of the major studios (or 'majors') into giant media conglomerates whose scale of operation is nothing less than global.

Not all of these changes are dealt with in equal depth in what follows, but they represent important background to any general examination of the recent growth and development of the Hollywood motion-picture industry. Three overarching themes structure the discussion of these changes here. The first is concerned with the shifting logic and dynamics of Hollywood's competitive advantages as a dense agglomeration of firms and workers and associated institutions. The second is focused on the segmentation of production activities in Hollywood into two distinctive, though overlapping segments. The third deals with the ways in which Hollywood projects its outputs onto wider markets, and with the ways in which it is now greatly intensifying its global reach and hold.

ECONOMIC GEOGRAPHY AND THE NEW HOLLYWOOD DEBATE

The key research on the economic geography of the new Hollywood was carried out by Susan Christopherson and Michael Storper in the mid to late 1980s (CHRISTOPHERSON and STORPER, 1986; STORPER and CHRISTOPHERSON, 1987; and STORPER, 1989). The basic argument set forth by these two authors revolves around the transformation of the classical vertically-integrated studio system of Hollywood into the much more vertically-disintegrated production complex that it has become today.

Christopherson and Storper describe the old studio system in terms of a dominant group of seven majors, each of them vertically integrated across production, distribution and exhibition. They also characterize the actual work of making films under the studio system as a mass production process (see also BORDWELL *et al.*, 1985). They then go on to claim that the restructuring of this system was induced by two main factors; the *Paramount* antitrust decision of 1948; and the advent of television in the 1950s. The *Paramount* decision forced the majors to divest themselves of their extensive theatre (cinema) chains[2] (see CASSADY, 1958), and television drained off the audiences that had previously flocked to motion-picture theatres. The net effect, according to Christopherson and Storper, was a dramatic rise in competitiveness, uncertainty and instability in the motion-picture industry, followed by the break-up of studio-based mass production, whose peculiar process and product configurations could no longer sustain profitable operations. Instead, the system was succeeded by a new order in which the majors divested themselves of much of their former productive capacity and contractual engagements, and became the nerve centres of vertically-disintegrated production networks. In this process, many kinds of skilled employees who had previously been on studio payrolls (producers, directors, writers, actors, musicians, camera operators, and so on) became freelance agents of their own labour (ANDERSON, 1994). Equally, large numbers of small flexibly-specialized firms sprang up in a wide range of subsectors in the motion-picture industry, providing both direct and indirect inputs of all kinds to the majors. This turn of events allowed the majors to cut their overheads, to pursue ever more diversified forms of production, and eventually to flourish in the new high-risk Hollywood (see KRANTON and MINEHART, 2000). Recently, CAVES, 2000, has described this same kind of development in the creative industries generally as a contractual model of business activity. In Christopherson and Storper's account, the break-up of the studio system and the emergence of a new flexibly-specialized Hollywood was associated with a 'loss of control by the majors over production' (STORPER, 1993, p. 482), though the authors also note that the majors continued to play important roles in Hollywood as centres of financing, deal-making and distribution. With the reconstitution of the production system as a transactions-intensive congeries of small and specialized but complementary firms, the agglomerative forces holding the entire complex together in geographic space were reinforced and its regional competitive advantages secured.

The Christopherson–Storper story represents the first really serious attempt to understand the organizational and locational foundations of Hollywood as a productive agglomeration, and it must be given high marks for its pioneering analysis, especially in view of the fact that the shifts the authors were trying to understand were far from having fully emerged and were still very much subject to confusing cross-currents. Their basic characterization of the new Hollywood in terms of shifting networks of small flexibly-specialized firms provides us with eminently useful insights, notwithstanding the criticism to which this idea has been subject of late. Analysts such as AKSOY and ROBINS, 1992; WASKO, 1994; SMITH, 1998; VÉRON, 1999; and BLAIR and RENNIE, 2000; have all questioned the emphasis on flexible specialization, and have instead averred that contemporary patterns of production in Hollywood can only be understood as an expression of the economic power and leverage of the majors. Other recent work that has laid stress on the continuing muscle of the majors has been presented by SCHATZ, 1997; BALIO, 1998; GOMERY, 1998; LITMAN, 1998; and MALTBY, 1998; among others.

In actuality, it is no doubt fair to argue that Christopherson and Storper paid insufficient attention to the durability and importance of the majors in the Hollywood production system, though they certainly did not overlook this aspect altogether. For their part, a number of the critics – notably, Aksoy and Robins – can be faulted for their radical depreciation of the role played by small producers, amounting virtually to an exercise in writing them out of any meaningful analysis of contemporary Hollywood. Aksoy and Robins are right to maintain that oligopoly never ceased to exist in Hollywood (though its foundations were greatly shaken over the 1960s and 1970s). However, they fail seriously to identify the sources of the majors' market power, which at least since the Second World War have resided mainly in the internal economies of scale that characterize their distribution systems (HUETTIG, 1944). The market power created in this manner is by no means necessarily incompatible with the existence of an efficient and dynamic production system based on flexible specialization *à la* Christopherson and Storper. In fact, in Aksoy and Robins' account, we lose sight entirely of the production system itself as a dense regional complex made up of thousands of intricately interdependent firms and, by the same token, of the independent segment of the industry as a substantial locus in its own right of innovation, skilled work, and many and varied final products.

I shall attempt to demonstrate in what follows that a more accurate portrayal of Hollywood today involves acknowledgment of the important roles played by both large and small firms, i.e. by the majors, by independent production companies and by the firms that supply them with different specialized service inputs. The argument interweaves with a further refrain focusing on the globalization of Hollywood's market range (BALIO, 1996), and this phenomenon actually appears – for the moment at least – to be reinforcing the centripetal locational attraction of Southern California for motion-picture production activities of all kinds.

AN ANALYTICAL TAXONOMY OF FIRMS

At the outset, we need to clarify some of the conceptual language that has already made its entry in the previous section, and to use this exercise as a platform for a more disciplined description of the motion-picture industry. This will help us, in addition, to overcome some of the more disabling elements of the one-sided debates about the extent of flexible specialization versus

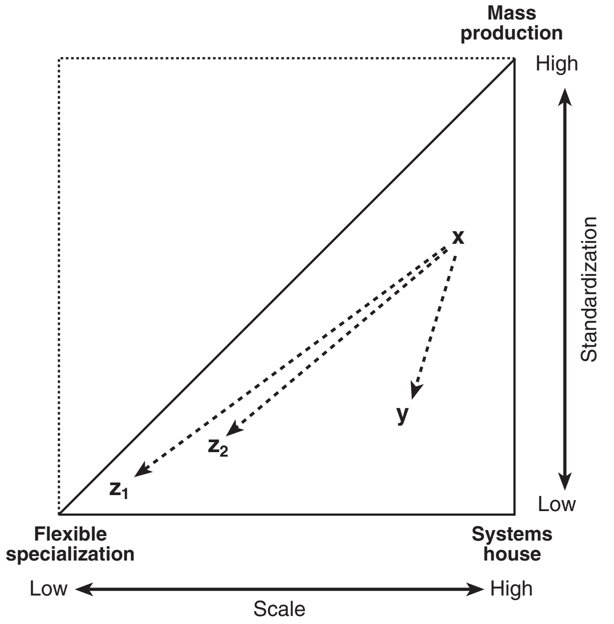


Fig. 1. *Schema of basic organizational possibilities in industrial systems:* x *and* y *represent old and new style studios, respectively;* z_1 *and* z_2 *represent common kinds of independent production companies or service suppliers*

large-scale oligopolistic production and distribution in Hollywood today.

In a very schematic way, any modern production system may be represented in terms of the size of its component units and the standardization (or variability) of their outputs. This idea is codified in Fig. 1, where scale and standardization represent orthogonal axes defining a space within which any given unit of production can then be situated. Three paradigmatic outcomes relative to these axes are identified in the figure. One of these is designated *mass production*, which is exemplified by plants or establishments that produce standardized outputs in large quantities. Another is represented by *systems houses,*[3] which can be defined as large-scale production units turning out limited numbers of extremely variable and complex products (like space satellites or blockbuster films). The third case is labelled *flexible specialization*, which refers to small production units that focus on a relatively narrow line of business (for example, talent agencies, costume designers or film editing services), but where the design specifications of each particular job, or batch of jobs, are different from all preceding jobs. A fourth paradigmatic case (corresponding to the northwest corner of Fig. 1) can conceivably be identified in the guise of small-scale process industries (Adam Smith's pin factory comes to mind), though this case would seem to be of limited empirical interest today. In reality, we rarely observe pure examples of these paradigmatic cases. Rather, actual production systems are usually composed of more or less hybrid production units, representable by points at intermediate positions relative to the vertices in Fig. 1.

In spite of the notion developed by several analysts to the effect that the Hollywood studios had moved into something like paradigmatic mass production by the late 1930s, a brief scrutiny of their actual working operations reveals that this was not quite the case. Even less is it plausible to claim, as some have done (see CRETON, 1994, 1997), that the system approximated to *Fordist* mass production. Notwithstanding the efficiency gains that flowed from finely-grained technical divisions of labour, the use of continuity scripts, the constant re-utilization of formulaic plot structures, and the search for regular production schedules, film-making in the classical studio era was never standardized in any ultimate sense. We might argue, rather, that the classical studio can be represented by a location somewhere in the vicinity of point x in Fig. 1, which is to say that its technical and organizational configuration was marked by quite high levels of scale and a degree of routinization, but nothing equivalent, say, to the typical Detroit automobile assembly plant churning out identical models by the thousands.

The old studios were nonetheless very much characterized by vertical integration over virtually all segments of the industry. STORPER and CHRISTOPHERSON, 1987, are probably correct to invoke the *Paramount* decision and the advent of television as being major factors in the destabilization of the industry's markets and its consequent restructuring, though the fact that very similar processes of break-up were proceeding at the same time in the music-recording industry and in television broadcasting (see SCOTT, 1999a; HIRSCH, 2000) suggests that there may well have been more general trends at work. In addition to the rupture between distribution and exhibition, two other main organizational effects flowed from vertical disintegration in the motion-picture industry. The first was the transformation of the studios themselves into something closer to systems houses, i.e. large-scale (though comparatively downsized) establishments now focusing on the production of many fewer and increasingly grandiose films. The point labelled y in Fig. 1 represents a typical case. The second was the emergence of masses of small independent production companies and service providers as exemplified by points z_1 and z_2 in Fig. 1. This second category of firms comprises flexibly-specialized producers and near relations in the sense that they concentrate on making a narrow range of outputs in comparatively limited quantities, and in ever-changing shapes and forms. With the disintegration of the old studios after the 1940s, the latter types of firms colonized an enormous number of different market niches, and they have continued subsequently to push out the organizational boundaries of Hollywood, a notable recent instance being the formation of a vigorous special-effects sector in the 1980s and 1990s. To be sure, even in the age of the classical Hollywood

Table 1. Feature films released in the US by majors and independents[1]

Year	Releases		
	Majors	Independents	Total
1980	134	57	191
1985	138	251	389
1990	158	227	385
1995	212	158	370
2000	191	270	461

Note: 1. In this table, the term majors refers to both the majors proper and their subsidiary releasing companies.
Source: Motion Picture Association of America, *2000 US Economic Review* (http://www.mpaa.org/useconomicreview/).

studios, small specialized firms were not uncommon, but in no way did they achieve the significance, either as independent film producers or as specialized suppliers, that they have now.[4]

The Hollywood production system today can hence be described in terms of a prevailing pattern of major and independent film production companies (see Table 1), intertwined with ever-widening circles of direct and indirect input suppliers. These firms interact with one another in complicated ways as any given motion-picture production project moves through its three main stages of development, namely; (1) pre-production, involving elaboration of the initial idea, scenario preparation, raising finances, set design, casting, and so on; (2) production proper, an intense period in which large numbers of workers are mobilized in directing, acting, camera-operating, and numerous allied functions from set construction to lighting and make-up (DeFillippi and Arthur, 1998); and (3) post-production, namely, photographic processing, film editing, sound editing, and so on. In practice, as the discussion will now show, the production companies engaged in this sort of work can be segregated into two distinctive functional tiers, represented on the one side by the majors and associated firms (both subsidiaries and independents), and on the other side by a mass of independent production companies whose sphere of operations rarely or never intersects with that of the majors.

A BIFURCATED PRODUCTION SYSTEM

The world of the majors

At the present time, there are eight major studios in Hollywood: Metro-Goldwyn-Mayer; Paramount Pictures; Sony Pictures Entertainment (Columbia-Tristar); Twentieth Century Fox; Universal Studios; Walt Disney Company; and Warner Brothers; together with newcomer Dreamworks (see Fig. 2). The first seven of these units are joined together in the Motion Picture Association of America (MPAA), which functions as an exclusive cartel promoting their interests.

The majors have traditionally concentrated on the financing, production and theatrical distribution of motion pictures, but over the last few decades they have actively diversified their operations, and they now earn as much, if not more, of their revenues through their specialized divisions in such fields as television programming, home video, multimedia, theme parks and merchandising. Most of the Hollywood majors today constitute operating units within even larger multinational media and entertainment conglomerates (Litman, 2001). Three of these – the News Corporation (which owns Fox), Sony (Columbia Tristar) and Vivendi (Universal) – are foreign-owned. As Acheson and Maule, 1994; Wasko, 1994; Balio, 1998; Gomery, 1998; Puttnam and Watson 1998; Prince, 2000; and others have suggested, the growing complexity of these conglomerates can be ascribed in large degree to attempts to internalize the synergies that are frequently found at intersections between different segments of the media and entertainment (and hardware) industries. The modern media-entertainment conglomerate accordingly functions as a sort of parallel in economic space to industrial clusters in geographic space, i.e. as an economic collective, with the difference that if, in the one case, the relevant synergies are activated under the umbrella of common ownership, in the other they owe their genesis to geographic proximity. Fig. 2 sketches out the ownership relations between the Hollywood majors and their parent companies, as well as between the majors and their most important subsidiary film-production and distribution companies. However, the figure refers only to the feature-film operations of the majors and makes no reference to other divisions, e.g. in television programming or home video production.

The majors as currently constituted engage in feature-film production with varying degrees of vertical integration and disintegration of the relevant tasks. One way in which they proceed entails integrated in-house development, shooting and editing using their own creative staffs and equipment as basic resources. It is important to note in this context that whereas the majors today are certainly more vertically-disintegrated than their pre-war forerunners, they never fully gave up all of their capacity to produce motion pictures in-house, and in most instances, their continuing level of vertical integration is quite considerable. Many of them, for example, still own large-scale sound stages, and maintain significant pre- and post-production facilities, all of which are also available for lease by outside companies.[5] That said, as any given production project by the majors moves forward, other firms and individuals (such as producers, directors, set designers,

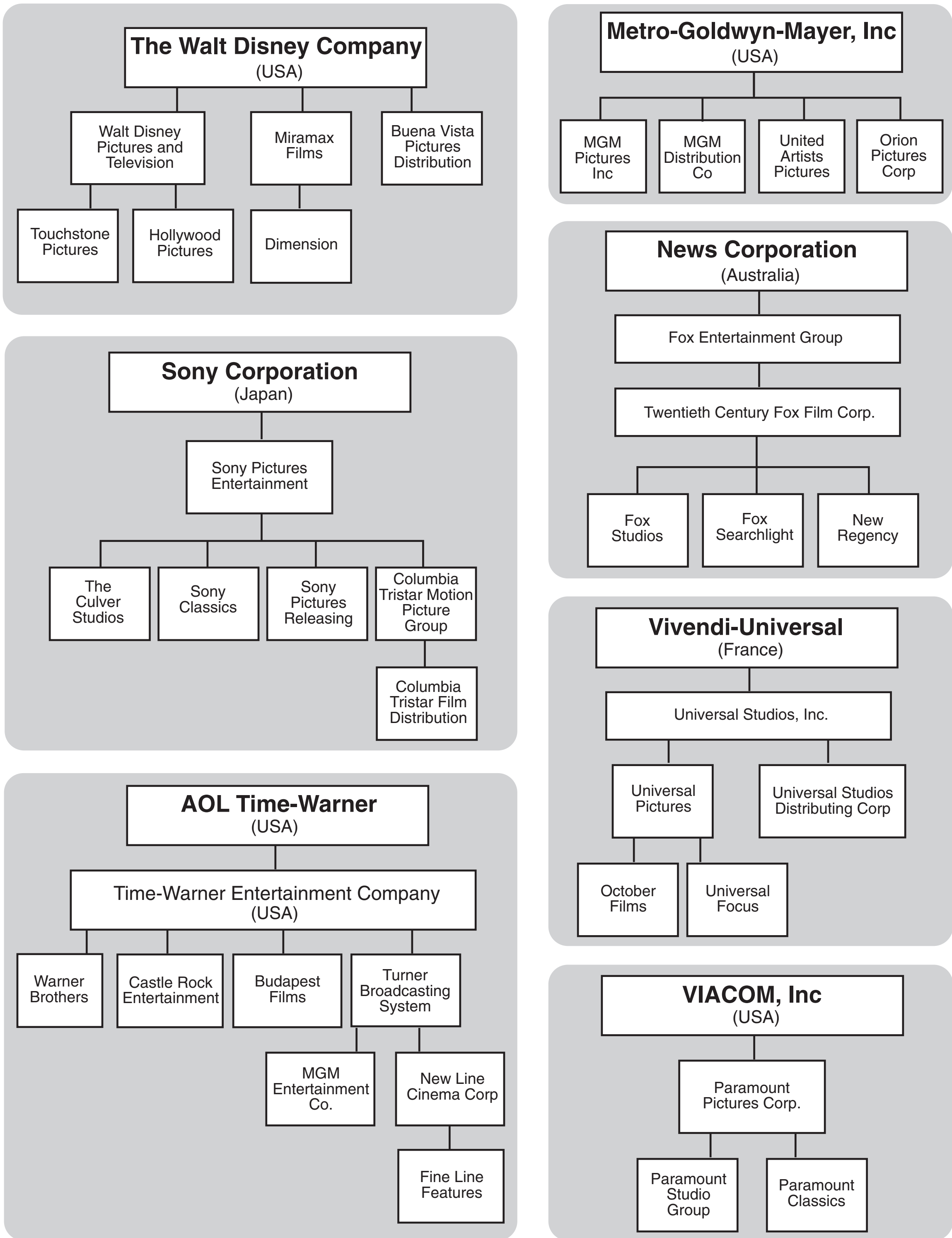


Fig. 2. The Hollywood majors: corporate ownership relations

Sources: Various directories, reports and web sites.

and so on) are commonly brought in on a subcontract or limited-term contract basis to perform specific tasks.

Another way in which the majors proceed is to work with smaller production companies, where the latter assume primary responsibility for organizing overall production tasks. The smaller companies involved in these ventures comprise both the majors' own subsidiaries and selected independent producers in projects that may range anywhere from a niche-oriented film to a high-budget blockbuster. In these collaborative

Table 2. Films released by majors and their subsidiaries, 1980–2000

Year	Releases by majors less releases by subsidiaries	Releases by majors' subsidiaries	Subsidiaries as a percentage of majors
1980	94	0	0
1985	103	12	11·7
1990	109	27	24·8
1995	127	85	66·9
2000	104	75	72·1

Source: Calculated from information in *Annual Index to Motion Picture Credits*, Academy of Motion Picture Arts and Sciences. Subsidiaries involved in these counts are: Castle Rock Entertainment; Dimension Films; Fine Line Features; Fox Searchlight; Miramax; New Line Distribution; October Films; Orion Pictures; Orion Film Classics; Paramount Classics; Sony Classics; Tristar Pictures; Triumph Releasing; Twentieth Century Fox International Classics; United Artists Films; Universal Focus; and Warner Classics. Note that some of these entities have operated as independents at various times, and some are no longer in existence. They are included in the present counts only for years when they actually functioned as subsidiaries of majors. Buena Vista Pictures is the principal distributing arm of Disney and is treated as a major. The discrepancies observable between the data given in Table 1 and Table 2 stem from the different sources used.

ventures, the majors work in a range of protocols, though in probably the majority of cases these grant significant control to the majors over production and editing decisions. Typical procedures include financing, production and distribution deals, co-production pacts, joint ventures, split rights agreements, 'first look' contracts, and any and all combinations of these arrangements. The majors also enter into negative pick-up contracts with independents, that is, agreements to distribute films that have already been completed before being brought to the attention of any given studio. Many independents also unilaterally assemble packages of scripts, actors, directors and other assets that they then present to the studios in the hope of securing a production or distribution agreement, though few are ever successful (ACHESON and MAULE, 1994).

The data given in Table 2 shed important light on the evolving world of the majors. The table indicates the number of releases by majors and their subsidiary companies at five-year intervals since 1980. Recall that the release or distribution of films is to be distinguished from the actual production of films, and that the majors release films made by themselves or their subsidiaries as well as films in which independent production companies participate in different degrees. By contrast, releases by subsidiaries owned by the majors are overwhelmingly produced by smaller independent companies. Many of these subsidiaries (e.g. Castle Rock, Miramax, New Line, Orion Pictures) began their existence as independents; others (like Fox Searchlight or Universal Focus were set up as auxiliary units from the start. Table 2 reveals that, since 1980, the number of films released by the majors proper has remained more or less constant at close to a hundred a year, whereas the subsidiaries have greatly increased their releasing activities, in parallel with production in the independent sector (see below). Thus, although the majors continue to dominate the entire industry, and continue to maintain a significant degree of in-house production capacity, they also rely more and more on smaller subsidiaries and independent production companies in order to spread their risks, to diversify their market offerings, and to sound out emerging market opportunities. In this connection, the business strategies of the motion-picture industry majors strongly resemble those of majors in the music business (DALE, 1997; NEGUS, 1998; SCOTT, 1999a; HIRSCH, 2000).

The world of the independents

The independent segment of the industry represents an important and flourishing element of the Hollywood complex. As the information set forth in Table 1 suggests, independent film production has increased greatly over the last two decades, with the period of most intense growth being the early to mid-1980s when a boom in independent film production occurred, fuelled by the growth of ancillary markets (PRINCE, 2000). Independent production companies make films for both domestic and foreign markets and for presentation in any and all formats (theatrical exhibition, television or home video). They cater to a great variety of market niches, and their outputs include art films for specialized audiences, genre movies of all kinds, documentaries, television commercials and direct-to-video films (among which a fair proportion is composed of the numerous pornographic films made by firms clustered in the San Fernando Valley). The distribution of films made by independent producers is handled for the most part by independent distribution companies, many of them highly specialized with respect to market niche (DONAHUE, 1987; ROSEN and HAMILTON, 1987).

According to *County Business Patterns* some 3,500 establishments were engaged in motion-picture and video production (NAICS 51211) in the five counties of Southern California in 1999, and the median size of these establishments was just two employees. The vast majority of these establishments comprise a cohort of independent production companies, some of which are allied with majors but most of which operate in an entirely separate sphere. Thus, in face-to-face interviews with representatives of many different firms it was found that significant numbers – perhaps the majority – of Hollywood independents rarely or never come into contact with a major, and work in an entirely separate sphere of commercial and creative activity. This observation is confirmed by results from a postal survey of independent production companies in Hollywood carried out in the summer of 2001. Out of a total of 115 respondents who were asked if they had engaged in any production deals with majors over the previous 12 months, 83 (72·2%) responded in the negative, and it was evident that some of those who responded positively were actually making a very liberal interpretation of the term 'major'. Additionally, the average response to a question asking firms to rate the dependence of their business upon the production and distribution divisions of majors on a scale ranging from 1 (no dependence) to 5 (high dependence) was a rather low 2·8.

A bipartite or tripartite system?

It is to be stressed that the two tiers of productive activity identified above are far from being hermetically sealed off from one another. First of all, there are obvious symbioses between the two in the sense that each generates externalities that are of value to the other, including important flows of new talent from the lower to the upper tier. Second of all, some independent production companies work partially in the one tier and partially in the other, and others move erratically in and out of the sphere of operation of the majors. Indeed, we might well want to qualify any description of the Hollywood production complex as a bifurcated system with the idea that the two tiers described above are actually complemented by a more indistinct circle of companies as represented by independents strongly allied to the majors together with the majors' own subsidiaries. This middle tier provides a shifting but evidently widening bridge between the two more clearly definable segments as represented by the majors proper and the pure independents.

THE GEOGRAPHY AND DYNAMICS OF THE HOLLYWOOD PRODUCTION COMPLEX

A schematic overview

Hollywood is neither just a metaphor nor just a business model; it is also a unique *place*, with a very distinctive structure as a production locale. As such, one important approach to understanding its character and evolution is offered by the contemporary theory of industrial districts and regional development. Since there already exists a large general body of literature on this issue (see, for example, SCOTT, 1993; STORPER and SCOTT, 1995; COOKE and MORGAN, 1998; PORTER, 2001), I shall be brief in what follows.

The key elements of the Hollywood production complex today can be described by reference to four

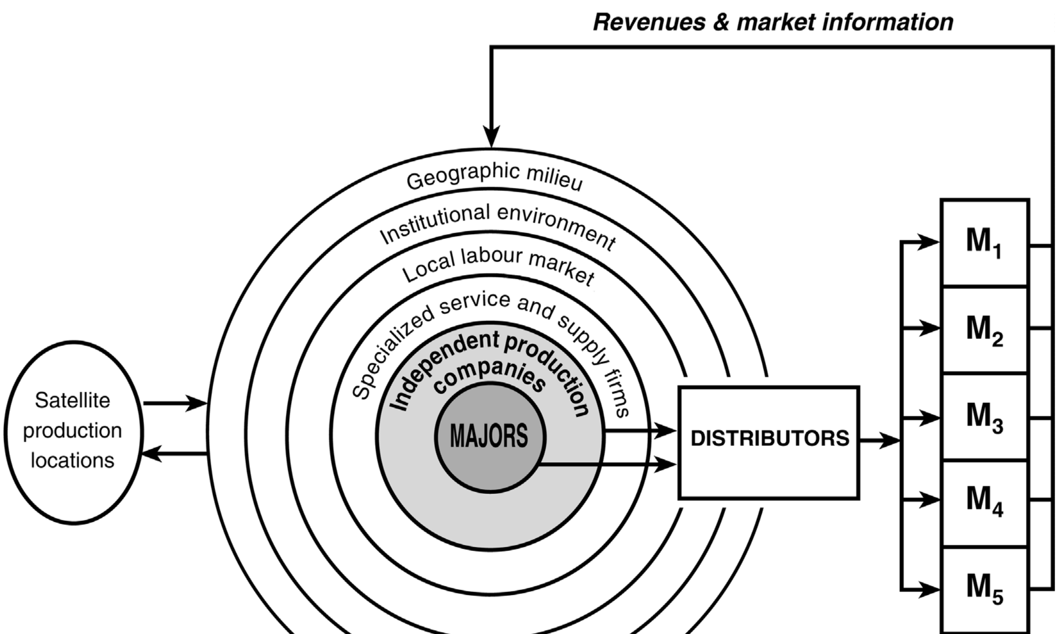


Fig. 3. Schema of the Hollywood motion-picture production complex and its external spatial relations
Note: M_1, $M_{2,\ldots,}M_5$ represent markets differentiated by niche and by geography.

main functional and organizational features (see Fig. 3). These are:

1. A series of overlapping production networks in various states of vertical disintegration. The nodes of these networks are composed of majors, independents and providers of specialized services from script writing to film editing.
2. A local labour market comprising a large number of individuals differentiated according to skills, sensibilities and forms of habituation. This labour market is constantly being replenished by new talent from all over the rest of North America and the world.
3. An institutional environment made up of many organizations and associations representing firms, workers and governmental agencies.[6] Some of these organizations exert considerable influence over the developmental trajectory of the industry.
4. A regional milieu whose peculiar geographic and historical features emerge in part in relation to the phenomena identified in points 1, 2, and 3, and which is a repository of crucial resources for the industry. These range from the cinematic traditions that are embedded, as it were, in the very fabric of Hollywood as a production locale, through the conventionalized background landscapes of Southern California, to the synergy-laden potentials offered by proximity to the region's many other cultural-products industries (Molotch, 1996).

These four points all allude to important positive externalities underlying the Hollywood production complex, endowing it with strong competitive advantages in the form of increasing returns to scale and scope and positive agglomeration economies. Such advantages are fundamental in maintaining the status of the region as the leading centre of motion-picture production in the world today. They are also major elements of an organizational-geographic framework that functions as a seedbed of creativity and innovation for the industry (see Scott, 1999b). Like many other regional complexes, this framework evinces a periodic tendency to lock-in to relatively fixed configurations over time; yet so far in its long history, the industry has always in the end managed to overcome the many crises of adjustment that have been sparked off by periodic shifts in basic technological and market conditions like the invention of talking movies or the development of new digital technologies.

Two other brief remarks complement the discussion of Fig. 3, and will be picked up again in more detail later. First, in spite of the centripetal locational pull of Hollywood, expanding streams of production activities have been moving to distant satellite locations since the 1980s. Second, distribution represents an especially critical adjunct to production. Without effective distribution, the production system could attain neither the scale nor the scope that help to make it such a formidable source of competitive advantages today.

Development and growth, 1980–2000

All of this productive activity calls for an enormous variety of worker skills, service inputs and entrepreneurial effort, and Southern California offers an extraordinarily dense concentration of these assets. Most of the industry is clustered in a relatively small geographic area centred on Hollywood itself, but also spilling over into other parts of the region. The detailed geographic outlines of the complex are represented by Fig. 4, which shows the locations of individual production companies in the region. Observe the dense swath of firms sweeping from Burbank in the east through the central pivot of Hollywood to Beverly Hills and Santa Monica in the west. Remarkably few production companies are located outside this dense primary cluster.

The motion-picture industry has grown greatly in Los Angeles County over the last few decades, as indicated by Figs. 5 and 6, which trace out changes in employment and number of establishments in motion-picture production and in allied services since 1980. The information presented in these two figures is defined in terms of the old standard industrial classification (SIC), as opposed to the new North American industrial classification system (NAICS) that succeeded it in 1997. As it happens, SIC 7812 (motion picture and video production) is perfectly matched by NAICS 51211 and thus we can extend any data series under the former rubric beyond 1997. SIC 7819 (services allied to motion pictures) has no corresponding NAICS codes, so that data series defined under this rubric cannot be continued after 1997. Over the period 1980 to 1999, employment in SIC 7812 in Los Angeles County actually declined from 39,318 to 29,262; by contrast, the number of establishments in the same sector increased massively from 983 to 3,237. Employment in SIC 7819 grew from 10,946 to 120,000 from 1980 to 1997, and the number of establishments in the same sector expanded from 509 to 2,326, which clearly reflects the great rise in demand for intermediate inputs to the industry, including special effects and other digital services (Scott, 1998; Hozic, 1999, 2001). Thus, taken as a whole, motion-picture production and service activities (SIC 7812 plus SIC 7819) in Los Angeles County grew at a rate of 194·0% for employment and 248% for establishments between 1980 and 1997. This trend runs parallel to the considerable increase in the total number of films produced in the US over the same period (from 214 in 1980 to 684 in 2000, according to MPAA records). The global deregulation of television in the 1980s and 1990s no doubt also helped to stimulate this expansion. In the same period, a significant downsizing of establishments occurred in SIC 7812, in association with a

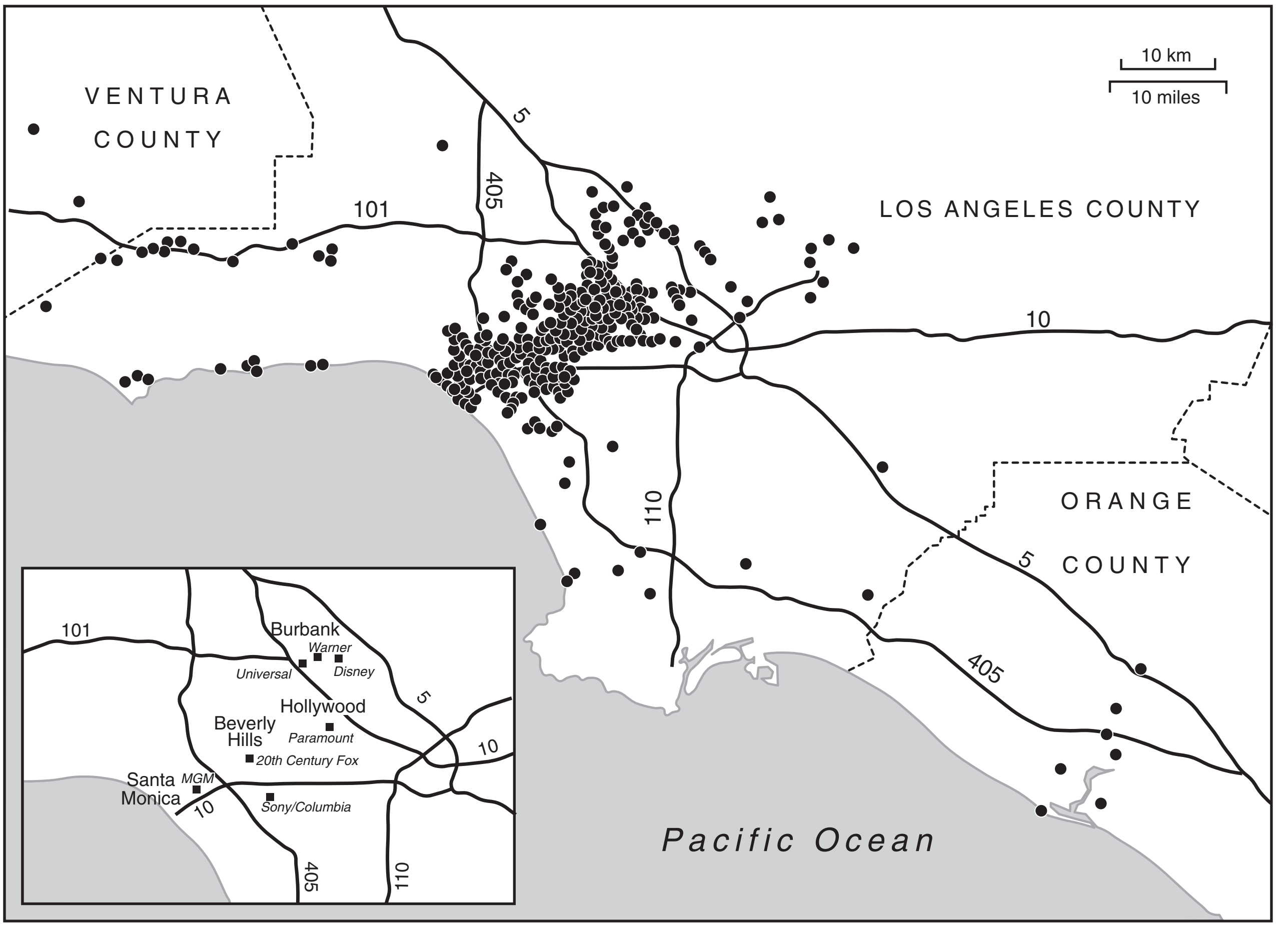


Fig. 4. Motion-picture production companies in Southern California

Note: The inset shows locations of the majors and selected place-names.

Sources: Blu-Book, 2001 (Los Angeles: Hollywood Reporter), and *Producers* (Los Angeles: Ifilm).

corresponding increase in average establishment size in SIC 7819. In the former case, establishment size in Los Angeles County fell from 39·9 to 9·0 employees on average; in the latter, it rose from 21·5 to 51·6. It is tempting to interpret these data in terms of continued vertical disintegration in production and increasing internal economies of scale in associated service providers, but in the absence of suitable statistics at the individual firm level, analysis of the precise mechanisms at work here must await further research.

Figs. 5 and 6 indicate that the growth of the motion-picture industry in Southern California has been accompanied by parallel expansion of the industry (and most notably by an increase in the number of small establishments) in the rest of the US. Even so, Southern California remains the primary agglomeration in the country, followed in distant second place by New York. In 1980, combined employment in SICs 7812 and 7819 in Los Angeles County represented 63·3% of the US total. In 1997, the figure was 61·4%. Hence, the industry not only continued to grow in absolute terms in Los Angeles over the 1980s and 1990s, but maintained its high level of relative geographic concentration as well.

Satellite production locations

Despite this outstanding historical performance, there has been much decentralization of production activities from Hollywood to satellite locations in recent years. Decentralization occurs for two main reasons, one being the search for realistic outdoor film locations (which has always been a feature of the industry's operations), the other being the search for reduced production costs (which is a more recent phenomenon). In the vernacular of contemporary Hollywood, firms engaging in these two types of decentralization are referred to as 'creative runaways' and 'economic runaways' respectively (MONITOR, 1999).

A number of studies have shown that the latter kind of decentralization has been increasing rapidly for film-shooting activities since the late 1980s (MONITOR, 1999; COE, 2001; ENTERTAINMENT INDUSTRY

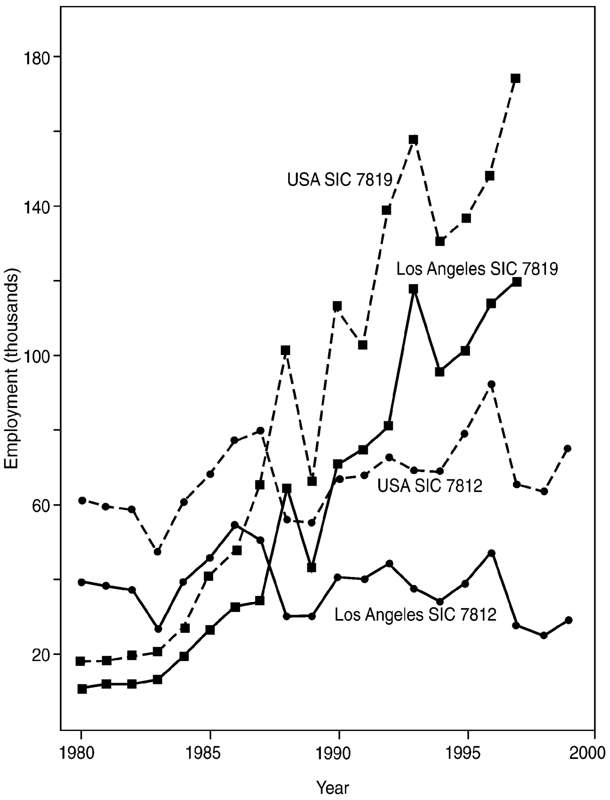


Fig. 5. Employment in the motion-picture industry, Los Angeles County and the US

Notes: SIC 7812 = Motion-picture and video production; SIC 7819 = services allied to motion pictures.

Source: County Business Patterns, Bureau of the Census, US Department of Commerce.

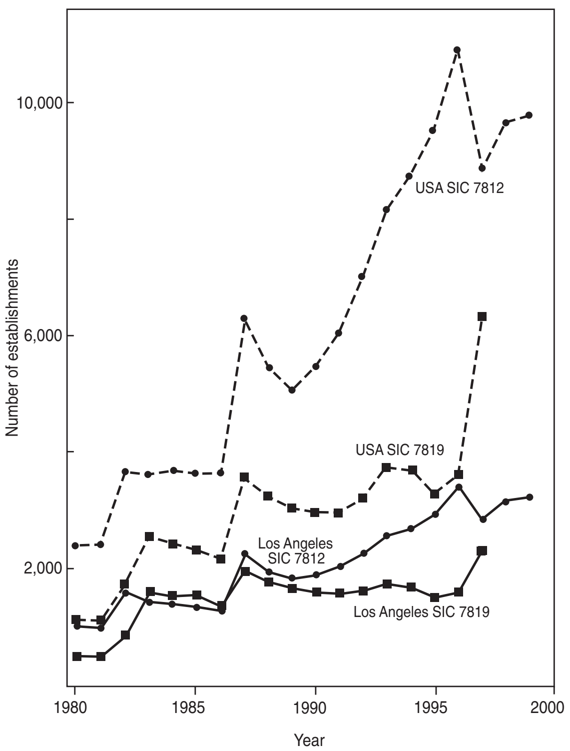


Fig. 6. Number of establishments in the motion-picture industry, Los Angeles County and the US

Notes: SIC 7812 = Motion-picture and video production; SIC 7819 = services allied to motion pictures.

Source: County Business Patterns, Bureau of the Census, US Department of Commerce

Development Corporation (EIDC), 2001; International Trade Administration (ITA), 2001). Most of it is directed to Canada, Australia, Britain and Mexico, with Canada receiving 81% of the total. The Monitor Company (1999) estimates that the total dollar loss to the US as a result of economic runaways was \$2·80 billion in 1998. The presumed Canadian share of this total is \$2·27 billion. By contrast, the Canadian Film and Television Production Association estimates that total revenues from foreign film shooting in Canada in the 1999/2000 season was just \$1·00 billion[7] (Canadian Film and Television Production Association (CFTPA), 2001). Even given the discrepancy between these two estimates, the total loss to Hollywood is evidently of major proportions, and the point is underlined by the fact that of 1,075 US film and television projects surveyed in 1998, 285 involved runaway production to foreign countries (Monitor, 1999). According to EIDC, 2001, most of this activity is currently accounted for by television productions, with movie-of-the-week programmes being the staple item.

Fig. 7 provides a simple analytical language for thinking about this issue. The central elements of the figure are average cost curves for a particular package

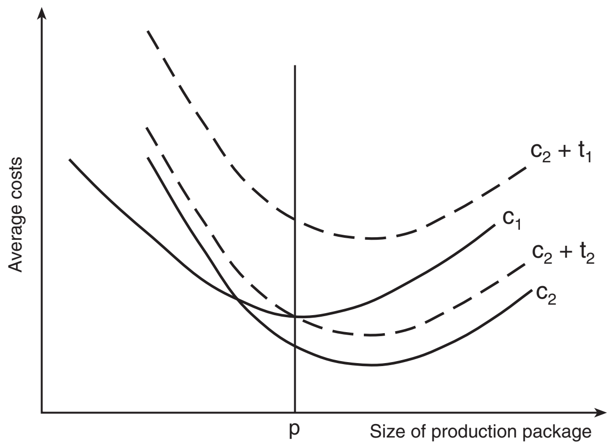


Fig. 7. Analysis of runaway production

Notes: In the graph, c_1 is the average cost curve for a given package of production tasks in Hollywood; c_2 is the average cost curve for the same tasks at a satellite location; t_1 and t_2 are unit costs of transacting between the two locations.

of production tasks (in the present instance, film shooting) at two different locations. One of these, Hollywood, is the home base, the other is a satellite location. The cost curves depict typical increasing and decreasing returns effects as a function of the size of any given package, with c_1 representing production in Hollywood, and c_2 production in the satellite location. We may assume that there are fixed set-up costs at the satellite location so that c_2 is greater than c_1 at relatively low levels of production; but because of various advantages at the satellite location, c_2 falls below c_1 at higher levels. Among these advantages we may count relatively low wages, low rental rates for sound stages and equipment, advantageous foreign exchange rates, governmental tax credits and subsidies, and so on (ITA, 2001). Any shift of production from the home base to a satellite location also entails transactions costs, including expenses for transport of personnel and equipment, communications charges between the home base and the satellite as production is proceeding, and perhaps most importantly of all, implicit costs due to diminished managerial and creative control over day-to-day work activities and hence over the quality of the final product. These transactions costs (which may, like production costs, be subject to increasing and decreasing returns) are apt to be particularly onerous where a high-budget feature film is concerned, but relatively low in the case of a more routine television programme with limited production values and frequent repetition of the basic package specifications. The interplay between average production costs and transactions costs at different levels of scale then determines whether runaway production is economically feasible in any given case. As shown in Fig. 7, a strong production cost advantage at the satellite location can be completely eliminated where transactions costs are high; then, as transactions costs are lowered there will be a greater and greater incentive to shift production to the satellite location. Eventually, if average transactions costs fall from t_1 to t_2, as in Fig. 7, production at the satellite is likely to occur for task packages larger than *p*.

This exercise clarifies the interactions between all of the various costs brought into play in runaway production relative to the size and complexity of the tasks to be completed. In view of this analysis, we can obtain a clearer grasp of just why (relatively standardized) television films are more susceptible to runaway production than feature films. And we can extend the analysis by increasing the number of possible satellite locations. For example, a significant appreciation of the Canadian dollar from its current exchange rate of Can$1.54 per US$1.00 to Can$1.15 per US$1.00, would just offset the reputed 25% cost advantage of Canada for film production, therefore making substitute satellite locations more attractive. Given this logic, there is every likelihood that Hollywood will continue indefinitely to lose certain kinds of production to one country or another, subject to the availability of adequate sound stage facilities and crews at alternative locations. A dramatic parallel case can be found in the Los Angeles clothing industry where a steady increase in offshore 'full-package' contracting has been occurring since the late 1980s (KESSLER, 1999; SCOTT, 2002). So far, runaway production has not seriously undermined the vitality of the Hollywood film industry, and it may well never become life threatening, at least in the more creative segments of the industry. This inference is based on two presumptions: (1) that the towering competitive advantages of Hollywood in pre- and post-production work will continue to prevail; and (2) that films requiring close supervisory control and complex customized inputs at all stages of production will continue to constitute a significant core of the industry's product range. Accordingly, and even though the great flow of shooting activities to Canada has unquestionably given a developmental boost to the motion-picture industries of Toronto and Vancouver where most of the work takes place (see COE, 2000a, 2000b, 2001), there seems little reason to suppose that the locational attractions of Hollywood are on the point of dissipation. In the same way, it is surely implausible to claim, along with CLOUGH, 2000, that as a consequence of the increasing use of special effects and digital technologies in the industry, its centre of gravity in California may be shifting toward the Bay Area. The claim is yet more implausible (despite the defection of Industrial Light and Magic to Marin County in the late 1970s) in view of the extensive development of precisely a robust cluster of digital media and special effects firms in Southern California over the last decade (SCOTT, 1998).

In brief, Hollywood's competitive advantages – deriving from its overlapping transactional networks, the skills and creativity of its workers, its dense institutional underpinnings (including the many guilds, labour unions and producers' associations), its roots in a supportive regional milieu (one of whose attributes is the diverse and striking visual imagery of Southern California) and its proximity to related cultural-products industries – would appear to afford it some durability as a going concern. Its current vibrancy is all the more assured when we add to these advantages the benefits that it derives from its unparalleled distribution system (WILDMAN and SIWEK, 1988). Accordingly, the pronouncements of AKSOY and ROBINS, 1992, p. 19, to the effect that: 'Hollywood is now everywhere . . . production now moves almost at will to find its most ideal conditions, and with it go skills, technicians, and support services', and of HOZIC, 2001, p. 153, who talks about 'Hollywood's exodus into worldwide locations', are both exaggerated and premature.

DISTRIBUTION, MARKETS AND COMPETITION

Hollywood today is a large-scale, many-sided, cultural-production and franchising complex, disgorging an endless variety of products designed for many different market niches. The linchpin of the entire system is the high-concept, mass-appeal blockbuster, that is, a big-budget film with a simple but climactic central narrative, an uplifting finale, a major star presence and possessing many marketable assets (GARVIN, 1981; WASKO, 1994; WYATT, 1994; BRANSTON, 2000). The origins of this type of film are usually traced back to *Jaws* in the mid-1970s, with *Titanic* as its ultimate expression to date.[8] The market for all films is risky, and the high-concept blockbuster faces especially hazardous prospects. Only a few such films actually recoup their costs directly from theatrical exhibition, but the ones that do generally compensate for the ones that do not (DE VANY and WALLS, 1997). In addition, the studios now also reap large revenues from repackaging films for home video, broadcast and cable television licensing, product placements and spin-off products such as recorded music, games, toys, fashions, books, theme park rides, and so on.

Table 3. Top 10 film distributors in the US, 2000

Distribution company	Number of films released	Domestic box-office revenue (US$ millions)	Average per film
Buena Vista (Disney)	19	1,089	57·3
Universal	15	1,053	70·2
Warner Brothers	22	863	39·2
Twentieth-Century Fox	14	849	60·6
Paramount	12	792	66·0
Dreamworks	10	668	66·8
Sony	22	664	30·2
Miramax (Disney)	25	507	20·3
New Line (Warner)	14	388	27·7
USA Films	17	202	11·9
Totals	170	7,075	41·7

Source: Hollywood Reporter, Film 500, August, 2001.

The distribution system

Distribution has always been a vital element of the motion-picture industry. The distribution system disseminates the industry's products on wider markets, pumps revenues and information back into Hollywood, and hence is a basic condition of the sustained economic well-being of the central agglomeration (see Fig. 3).

Employment in the distribution branch of the business is densely developed in Los Angeles alongside the production activities that it serves. In 1999, Los Angeles County could claim 22,399 employees and 299 establishments in NAICS 51212 (motion-picture and video distribution), compared with 27,669 employees and 706 establishments in the entire US. Distribution is the segment of the industry where oligopoly is most in evidence. For the US as a whole in 1997, the four-firm concentration ratios in NAICS 51212 was 74·6%, as compared with equivalent ratios of 33·5% and 16·4% in NAICS 51211 (motion picture and video production) and NAICS 51219 (post-production and other motion picture and video industries), respectively.[9] This high level of concentration derives from the internal economies of scale that are inherent to distribution activities, especially where, as in the film industry, they assume the form of extensive networks with strong central management and widely-diffused regional offices. These networks can then be organized on the basis of repetitive operating rules in which the transmission of the variable product itself becomes relatively routinized. The economics of blockbuster production, with its associated logic of high-intensity, saturation marketing and distribution, greatly intensifies this tendency to concentration, especially given that the marketing and distribution costs of many blockbusters today are equal to or even greater than their actual production costs (CONES, 1997).

For any given blockbuster, prevailing marketing/distribution practices typically entail intense publicity campaigns over a short period of time, and exhibition in many different theatres simultaneously.[10] These practices, combined with the huge sums of money at stake, make it imperative for the majors to engage in close relational contracting with owners of theatre chains to secure assured and regular bookings well in advance of the publicized release date of their films. By the same token, there are certainly strong incentives to vertical reintegration of the entire production-distribution-exhibition chain in the motion-picture industry (see WATERMAN, 1982; BLACKSTONE and BOWMAN, 1999), and vertical integration has indeed been on the increase of late. This development can be traced back to the Reagan era when the Antitrust Division of the Justice Department began to take a more tolerant attitude toward infringements of the *Paramount* decision (PRINCE, 2000), and by the early 1990s, according to PRINDLE, 1993, the majors owned over 10% of all theatres in the US.

In light of these remarks, it is not surprising to observe that the same bifurcation characteristic of production activities in the Hollywood motion-picture industry is also – and even more – characteristic of distribution. As Table 3 indicates, nine of the top 10 film distributors in the US are either majors or subsidiaries of majors, and the one independent shown in the table (USA Films) has only recently displaced MGM from the top 10. The point is brought further home by an examination of detailed box-office statistics for films distributed in the US. Consider Fig. 8, which shows frequency counts of domestic box-office returns for major and independent producers. The figure is based on 142 films released by majors, and 304 films released by independents in 1999. A strikingly bimodal structure characterizes the pattern of frequencies displayed in Fig. 8, and the two counts overlap only in a small intermediate zone. For independent distributors, the average domestic box-office per film is $2·3 million, and for majors it is $46·1 million. So great is the discrepancy between the two, there might well be a

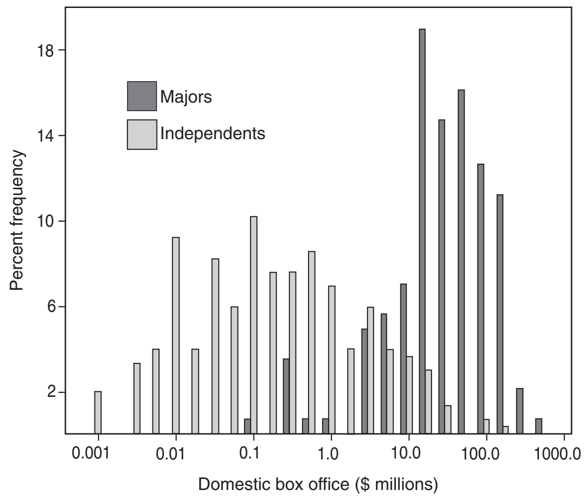


Fig. 8. Frequency distributions of domestic box-office returns for films released by majors and independents, 1999

Note: The *x*-axis is defined on a logarithmic scale.
Source: Encyclopedia of Exhibition, 2000–2001, National Association of Theatre Owners.

prima facie case for inferring that the majors are in some sense crowding the independents out of more lucrative markets (DALY, 1980). The business concentration of the majors is magnified by what CONES, 1997, refers to as their 'creative accounting' practices, where revenues are creamed off at the distribution phase, thus reducing the flow-back to production and by the same token diminishing claims for payment by outside contractors based on a percentage of producers' revenues.

In a study of motion-picture distribution in Canada, GLOBERMAN and VINING, 1987, have claimed that because there is rotating leadership among the larger distributors, and low barriers to entry at the bottom end of the system, the market is 'workably competitive'. Whatever the situation in Canada may be, the very marked concentration in the distribution sector in the US and the difficulties of outsider penetration into the top tier of the sector mean that significant impediments to competition exist. This situation can scarcely be qualified as being anything but oligopolistic.

The globalization of Hollywood

Astute marketing and distribution are crucial not only to the majors' domination of domestic markets, but to their ever-growing incursions on foreign markets too (HOSKINS *et al.*, 1997; JARVIE, 1998; SCOTT, 2000).

Table 4 shows the main pattern of US film and tape rentals to other countries over the period from 1986 to 1999. The first thing to notice from the data displayed is the tremendous growth in the total volume of exports since 1986 (364·7% growth in terms of constant dollars compared with 28·9% growth in domestic box office returns). Europe is the main destination with 65·5% of all exports in 1999, followed by Asia and the Pacific region with 17·4%. Britain, Germany, the Netherlands, France and Japan are the top individual importers. Two of the majors provided unpublished information to the author indicating that they own distribution facilities in all of the main countries mentioned in Table 4, as well as in a number of lesser markets such as Austria, Finland, Israel, Hong Kong, Singapore, Panama and Peru. In yet other markets the majors engage in joint ventures and contractual agreements with local distributors. It is not uncommon for the majors to have a main office in the largest city in their principal foreign markets (almost always in geographical association with an agglomeration of local audiovisual firms), together with a number of field offices in the provinces (see BONNELL, 1996; NACHUM and KEEBLE, 2000). Since 1995 the costs of Hollywood feature films released by the majors have on average consistently exceeded their domestic box office returns, so that foreign box-office earnings are now critical to overall profitability (see VOGEL, 1998). Thus, contrary to views expressed from time to time by European critics about the 'dumping' of Hollywood films on foreign markets, this is not strictly the case in economic terms, even if the charge may ring a sympathetic chord on a more cultural register.

Table 4. US exports in the form of film and tape rentals; percentage values by destination

Destination	1986	1991	1996	1999
France	10·2	8·6	8·6	7·8
Germany	7·5	9·6	10·5	13·2
Italy	10·0	7·3	4·7	5·3
Netherlands	15·2	17·5	17·4	11·5
Spain	—	5·1	5·9	6·6
Sweden	—	1·8	1·4	1·2
UK	10·6	11·0	9·8	13·4
Europe	*60·3*	*66·5*	*64·9*	*65·5*
Australia	10·3	3·4	4·8	4·6
Japan	8·3	11·5	8·7	7·7
Republic of Korea	—	0·8	1·8	1·1
Taiwan	—	0·5	0·7	1·1
Asia and Pacific	*22·1*	*18·3*	*19·3*	*17·4*
Brazil	—	0·8	2·2	2·9
Canada	10·4	8·7	6·8	5·2
Mexico	1·3	0·9	1·3	1·8
Americas	*17·9*	*12·5*	*13·0*	*13·0*
South Africa	—	—	1·1	1·0
Africa	—	*1·0*	*1·2*	*1·1*
Middle East	—	*0·5*	*0·8*	*1·2*
World ($ millions, current)	1,071	1,962	4,982	7,556
World ($ millions, constant)	1,628	2,400	5,290	7,566

Source: Survey of Current Business, US Bureau of Economic Analysis.

Strategic trade rather than dumping is Hollywood's trump card in international commerce. In contradistinction to classical atomized competition between individual free-wheeling firms, strategic trade is an outcome of imperfect competition in the context of increasing-returns effects, and is hence a source of rents over and above normal profits. In the case of the motion-picture industry, these effects derive in important ways from the potent agglomeration economies of Hollywood itself and from efficiencies of size in distribution, and it is almost certainly these endowments that constitute the primary source of the strategic prowess that has pushed American films so firmly to the fore in so many different foreign markets. Concomitantly, the relatively less well-developed character of these endowments in the film industries of other countries severely impedes them from making stronger inroads into American markets (CHASE, 2000; WATERMAN and JAYAKAR, 2000). There are no doubt also marketing difficulties that foreign films in the US face as a result of peculiarities of language and culture, though it is difficult to understand why these should operate necessarily in one direction but not in the other, all else being equal. In fact, all else is far from being equal, because over the entire post-war period, large US-based multinational corporations have honed their competencies to a fine point in matters of commercial propaganda and far-flung product distribution, perhaps most especially in media, entertainment and other culturally-charged products. Even if the multinationals of other countries are rapidly catching up in this respect, the pioneering efforts of US firms have more or less naturalized American cinematic idioms on many foreign markets, making Hollywood films highly competitive with purely local products (FINNEY, 1996; WATERMAN and JAYAKAR, 2000). Furthermore, under the provisions of the Webb-Pomerane Act of 1918, monopolistic practices on the part of American firms are explicitly permitted on foreign markets, enabling them to penetrate and dominate those markets more effectively. Thus, block-booking by US-owned film distributors is prevalent in foreign markets, even though it is illegal in the US.

The rents generated by strategic trade can almost always be much enhanced by agencies of collective action, such as industry associations and governmental bodies. A standard manoeuvre in this regard is to work on clearing away obstacles that limit access to foreign markets, thus releasing new rounds of growth based on self-perpetuating increasing-returns effects. This is certainly a principal objective of the powerful MPAA, which has offices in Washington DC and Los Angeles, as well as in several foreign countries. The MPAA is a highly-financed cartel representing the combined voice of the majors, and it has proven itself to be extraordinarily aggressive and successful in shaping trade agendas in audiovisual products, as well as in many other political tasks of concern to the industry. Thanks to the lobbying efforts of the MPAA, Hollywood has always received untold help from the US Commerce Department and the US State Department (SEGRAVE, 1997). Under the banner of free trade and fair competition, it has long carried out intensive lobbying campaigns with different government agencies in the effort to pressure foreign countries to lower barriers to the strategic trading activities of its members. Independent distributors, too, have a collective representative in the guise of the American Film Marketing Association (AFMA), which is based in Los Angeles and counts over 170 different companies as members. Besides vigorous defence of its members' interests in general, AFMA also holds the annual American Film Market in Santa Monica, which has grown over the last two decades to become the world's largest motion-picture fair, attended by more than 7,000 people from 70 countries.

The very success of American motion pictures on foreign markets has, of course, given rise to a worldwide debate not just about the economics of trade in audiovisual products, but also about the cultural predicaments that follow in its train (FEIGENBAUM, 1999; CHASE, 2000). The issue came to a head in the GATT negotiations of 1993 when, under the prompting of the European Union backed up strongly by France, audiovisual products were exempted from the trade liberalization provisions contained in the final agreement. It is, however, an open question as to how long the exemption will survive in its present form under the newly-constituted World Trade Organization. In any case, stubborn cultural and political resistances to the globalization of Hollywood are evident all over the world, from Canada to China and from France to South Korea. And since culture is always, and in profound ways, about identity, ideology and power, as much as it is about profits and cash flow, the current situation poses predicaments that call for some more imaginative framework of supra-national regulation than approaches based on the erroneous proposition that cultural products are essentially just inert commodities like steel or car parts.

CONCLUSION

With the steady improvement of electronic methods of distribution and information diffusion, the predicaments alluded to in the previous paragraph are liable to intensify greatly. This remark reflects in part the speed with which new communications technologies are currently demolishing international borders; it also is based on an expectation that the majors are just as likely to dominate content supply in the new order as they have done in the old. More accurately, we should say that if, in theory, new electronic means of communications allow small producers to tap readily into global markets, the massive resources of the majors will still in all likelihood enable them to gain a decisive edge in publicity and marketing, and hence in sales.

Over a more distant time horizon, the situation becomes increasingly murky. For one thing, as I have argued elsewhere (SCOTT, 2001), new and revivified cultural-products agglomerations are on the rise in many different parts of the world today. Notwithstanding the current hegemony of Hollywood, the ingredients of its success are not in principle forever locked in at one place, and it is entirely conceivable that other regions may eventually mount credible challenges to it on global markets, even granted the enormous hurdles that exist. For another thing, policy makers in other countries are now turning their attention to the tasks of building indigenous cultural-products industries with much greater capacities for market contestation than in the past. In the European Union, for example, the Media Plus Programme initiated in January 2001 (in succession to the earlier Media I and Media II programmes) is now engaged in a many-sided effort to improve the international competitiveness of the European audiovisual industries, including a push to put more effective distribution systems into place. The increasing trend to international co-productions lends further complexity to these matters. Moreover, expansionist European media corporations such as Vivendi, Bertelsmann or Polygram are vigorously scouring the world for new production and marketing opportunities while at the same time strengthening their roots in their home territories. Quite apart from these developments, the notorious unpredictability of consumer tastes in matters of popular culture means that Hollywood production companies can never rest on their laurels. They are always potentially subject to devastating competition from unexpected quarters, and despite Hollywood's long domination of worldwide film markets, its ascendance can never be absolute or final. Indeed, there have been numerous instances in the past when it has faltered even on its home terrain, including the notable period in the late 1960s when imports grew to the point where they represented fully two-thirds of all the films released in the US (SCHATZ, 1997).

If these comments point to potential perils ahead, Hollywood in its current incarnation is nonetheless one of the most remarkable examples of a successful industrial agglomeration anywhere in the world. Its size and complexity, its longevity, its global impact and the mystique that surrounds its products, all combine to bring it into sharp relief. It is all the more fascinating because unlike many other case study industrial districts (Silicon Valley, Orange County, or Boston's Route 128, for example), its outputs trade on a purely cognitive register. For this reason alone, Hollywood is one of the most arresting examples of the burgeoning cultural-products agglomerations that are on the rise all over the world today, no matter whether their stock-in-trade is film, multimedia, music, fashion, or any other vehicle of aesthetic and semiotic expression.

In the present paper, I have described the mainsprings of the Hollywood production complex, with special reference to its status as both a local and a global system of relationships. I have attempted, in particular, to lay out a new map of Hollywood and the world that is attentive to the ways in which this local/global system reflects the industry's peculiar tendency to structural and functional bifurcation. Much more research, of course, is needed on particular aspects of Hollywood's operations, including many questions about new digital technologies, creativity and innovation, local labour markets, the institutional fabric of the industry, agglomeration and decentralization processes, corporate organization, marketing, the dynamics of demand, and so on. The discussion presented here offers a conceptual and empirical context that eases the task of approaching these and allied questions. The discussion also points firmly, if laconically, to the steady convergence that appears to be occurring between the economic and the cultural in contemporary global capitalism, and to a few of the important analytical problems raised by this turn of events.

Acknowledgements – This research was supported by the National Science Foundation under award number BCS-0091921, with supplementary funding from the Haynes Foundation. I am grateful to Kathleen Lee and Michael Storper for their comments on an early draft of this paper.

NOTES

1. Nowadays the industry spills over into other parts of the region lying well beyond its original confines in Hollywood proper; the term thus now has a synecdochic rather than literal meaning as a geographic designation.
2. The *Paramount* decision did not, however, sever the link between production and distribution (ROBINS, 1993). Had it done so, the entire subsequent history of Hollywood would almost certainly have turned out quite differently. Among other possible outcomes, the degree of concentration in the distribution segment of the industry might well have been reduced, thus opening up a wider market space for independent films of all kinds, and possibly inhibiting the majors' shift into blockbuster productions.
3. The term is taken from the jargon of the aerospace industry; see SCOTT, 1993.
4. MEZIAS and MEZIAS, 2000, suggest that, in 1929, vertically-integrated firms controlled about 80% of the market.
5. MGM, by contrast, owns no studio facilities whatever, preferring to rent these as and when they may be needed.
6. Some of the more important of these institutions include: the International Alliance of Theatrical Stage Employees; the Directors' Guild; the Producers' Guild; the Screen Actors' Guild; the Writers' Guild; the Academy of Motion Picture Arts and Sciences (which organizes the annual Academy Awards); the Alliance of Motion Picture and Television Producers; the American Federation of Television and Radio Artists; the American Film Marketing Association; the Motion Picture Association of

America; and the Entertainment Industry Development Corporation.
7. 'Runaways' resurface in Canada as 'export value'.
8. These examples draw attention to the misnomer involved in the term 'high concept' and clearly distinguish the type from the *film d'auteur*.
9. Data from US Department of Commerce, Bureau of the Census, *Economic Census*, 1997.
10. A strategy that has now more or less displaced earlier distribution methods based on exclusive booking of films in a few selected theatres, and reliance on word-of-mouth as a means of garnering audiences for subsequent rounds of release.

REFERENCES

ACHESON K. and MAULE C. J. (1994) Understanding Hollywood's organization and continuing success, *J. Cultural Econ.* **18**, 271–300.

AKSOY A. and ROBINS K. (1992) Hollywood for the 21st century: global competition for critical mass in image markets, *Camb. J. Econ.* **16**, 1–22.

ANDERSON C. (1994) *HollywoodTV: The Studio System in the Fifties*. University of Texas Press, Austin, TX.

BALIO T. (1996) Adjusting to the new global economy: Hollywood in the 1990s, in MORAN A. (Ed) *Film Policy: International, National and Regional Perspectives*, pp. 23–38. Routledge, London.

BALIO T. (1998) A major presence in all the world's major markets: the globalization of Hollywood in the 1990s, in NEALE S. and SMITH M. (Eds) *Contemporary Hollywood Cinema*, pp. 58–73. Routledge, London.

BLACKSTONE E. A. and BOWMAN G. W. (1999) Vertical integration in motion pictures, *J. Communication* **49**, 123–39.

BLAIR H. and RAINNIE A. (2000) Flexible films?, *Media, Culture & Society* **22**, 187–204.

BONNELL R. (1996) *La Vingt-Cinquième Image: Une Economie de l'Audiovisuel*. Gallimard, Paris.

BORDWELL. D., STAIGER J. and THOMPSON K. (1985) *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*. Columbia University Press, New York.

BRANSTON G. (2000) *Cinema and Cultural Modernity*. Open University Press, Buckingham.

CASSADY R. (1958) Impact of the Paramount decision on motion-picture distribution and price making, *Southern Calif. Law Rev.* **31**, 150–80.

CAVES R. E. (2000) *Creative Industries: Contacts between Art and Commerce*. Harvard University Press, Cambridge, MA.

CANADIAN FILM AND TELEVISION PRODUCTION ASSOCIATION (CFTPA) (2001) *Profile 2001: Canadian Independent Production: Growth Opportunities in a Period of Consolidation*. Ottawa.

CHASE A. (2000) Globalization versus localization: cultural protection and trade conflict in the world entertainment industry, paper presented at the annual meeting of the International Studies Association, Los Angeles.

CHISTOPHERSON S. and STORPER M. (1986) The city as studio, the world as back-lot: the impact of vertical disintegration on the location of the motion-picture industry, *Environ. Plann. D* **4**, 305–20.

COE N. M. (2000a) On location: American capital and the local labor market in the Vancouver film industry, *Int. J. Urban & Reg. Research* **24**, 79–91.

COE N. M. (2000b) The view from out West: embeddedness, inter-personal relations and the development of an indigenous film industry in Vancouver, *Geoforum* **31**, 391–407.

COE N. M. (2001) A hybrid agglomeration? The development of a satellite-marshallian industrial district in Vancouver's film industry, *Urban Studies* **38**, 1,753–75.

CONES J. W. (1997) *The Feature Film Distribution Deal*. Southern Illinois University Press, Carbondale and Edwardsville, IL.

COOKE P. and MORGAN K. (1998) *The Associational Economy: Firms, Regions and Innovation*. Oxford University Press, Oxford.

CRETON L. (1994) *Economie du Cinéma: Perspectives Stratégiques*. Nathan, Paris.

CRETON L. (1997) *Cinéma et Marché*. Armand Colin, Paris.

DALE M. (1997) *The Movie Game: The Film Business in Britain, Europe, and America*. Cassell, London.

DALY D. A. (1980) *A Comparison of Exhibition and Distribution Patterns in Three Recent Feature Motion Pictures*. Arno Press, New York.

DE VANY A. and WALLS W. D. (1997) The market for motion pictures: rank, revenue, and survival, *Econ. Inquiry* **35**, 783–97.

DEFILLIPPI R. J. and ARTHUR M. B. (1998) Paradox in project-based enterprise: the case of film-making, *Calif. Mgt. Rev.* **40**, 125–39.

DONAHUE D. (1987) *American Film Distribution: The Changing Marketplace*. UMI Research Press, Ann Arbor, MI.

ENTERTAINMENT INDUSTRY DEVELOPMENT CORPORATION (EIDC) (2001) *MOWs – A Three-Year Study: An Analysis of Television Movies of the Week, 1997–1998, 1998–1999, and 1999–2000*. EIDC, Los Angeles.

FEIGENBAUM H. B. (1999) The production of culture in the postimperialist era: the world versus Hollywood?, in BECKER D. G. and SKLAR R. L. (Eds) *Postimperialism and World Politics*. Praeger, Westport, CN.

FINNEY A. (1996) *The State of European Cinema: A New Dose of Reality*. Cassell, London.

GARVIN D. A. (1981) Blockbusters: the economics of mass entertainment, *J. Cultural Econ.* **5**, 1–20.

GLOBERMAN S. and VINING A. (1987) *Foreign Ownership and Canada's Feature Film Distribution Sector: An Economic Analysis*. The Fraser Institute, Vancouver, BC.

GOMERY D. (1998) Hollywood corporate business practice and periodizing contemporary film history, in NEALE S. and SMITH M. (Eds) *Contemporary Hollywood Cinema*, pp. 47–57. Routledge, London.

Hirsch P. M. (2000) Cultural industries revisited, *Organization Science* **11**, 356–61.
Hoskins C., McFadyen S. and Finn A. (1997) *Global Television and Film: An Introduction to the Economics of the Business*. Clarendon Press, Oxford.
Hozic A. A. (1999) Uncle Sam goes to Siliwood: of landscapes, Spielberg, and hegemony, *Rev. Int. Pol. Econ.* **6**, 289–312.
Hozic A. A. (2001) *Hollyworld: Space, Power, and Fantasy in the American Economy.* Cornell University Press, Ithaca, NY.
Huettig M. (1944) *Economic Control of the Motion Picture Industry.* University of Pennsylvania Press, Philadelphia.
International Trade Administration (ITA) (2001) *Impact of the Migration of US Film and Television Production*. ITA, US Department of Commerce, Washington, DC.
Jarvie I. (1998) Free trade as cultural threat: American film and TV exports in the post-war period, in Nowell-Smith G. and Ricci S. (Eds) *Hollywood and Europe: Economics, Culture and National Identity, 1945–95*. British Film Institute, London.
Kessler J. A. (1999) The North American Free Trade Agreement, emerging apparel production networks and industrial upgrading: the Southern California/Mexico connection, *Rev. Int. Pol. Econ.* **6**, 565–608.
Kranton R. E. and Minehart D. F. (2000) Networks versus vertical integration, *RAND J. Econ.* **31**, 570–601.
Litman B. R. (1998) *The Motion Picture Mega-Industry.* Allyn & Bacon, Boston.
Litman B. R. (2001) Motion picture entertainment, in Adams W. and Brock J. (Eds) *The Structure of American Industry,* 10th edition, pp. 171–98. Prentice Hall, Upper Saddle River, NJ.
Litwak M. (1986) *Reel Power: The Struggle for Influence and Success in the New Hollywood*. Morrow, New York.
Maltby R. (1998) Nobody knows everything: post-classical historiographies and consolidated entertainment, in Neale S. and Smith M. (Eds) *Contemporary Hollywood Cinema*, pp. 21–44. Routledge, London.
Mezias J. M. and Mezias S. J. (2000) Resource partitioning, the founding of specialist firms, and innovation: the American feature film industry, 1912–1929, *Organization Sci.* **11**, 306–22.
Molotch H. (1996) LA as design product: how art works in a regional economy, in Scott A. J. and Soja E. W. (Eds) *The City: Los Angeles and Urban Theory at the End of the Twentieth Century,* pp. 225–75. University of California Press, Berkeley/Los Angeles, CA.
Monitor (1999) *US Runaway Film and Television Production Study Report*. Monitor Company, Santa Monica, CA.
Nachum L. and Keeble D. (2000) Localized clusters and the eclectic paradigm of FDI: film TNCs in central London, *Transnational Corporations* **9**, 1–38.
Negus K. (1998) Cultural production and the corporation: musical genres and the strategic management of creativity in the US recording industry, *Media, Culture & Society* **20**, 359–79.
Porter M. E. (2001) Regions and the new economics of competition, in Scott A. J. (Ed) *Global City-regions: Trends, Theory, Policy,* pp. 139–57. Oxford University Press, Oxford.
Prince S. (2000) *A New Pot of Gold: Hollywood under the Electronic Rainbow, 1980–1989*. Charles Scribner's Sons, New York.
Prindle D. F. (1993) *Risky Business: The Political Economy of Hollywood*. Westview Press, Boulder, CO.
Puttnam D. and Watson N. (1998) *Movies and Money.* Knopf, New York.
Robins J. A. (1993) Organization as strategy: restructuring production in the film industry, *Strat. Mgt. J.* **14**, 103–18.
Rosen D. and Hamilton P. (1987) *Off-Hollywood: The Making and Marketing of Independent Films*. Grove Weidenfeld, New York.
Schatz T. (1983) *Old Hollywood/New Hollywood Ritual, Art and Industry.* UMI Research Press, Ann Arbor, MI.
Schatz T. (1997) The return of the Hollywood system, in Aufderheide P. (Ed) *Conglomerates and the Media*, pp. 73–106. The New Press, New York.
Scott A. J. (1993) *Technopolis: High-technology Industry and Regional Development in Southern California*. University of California Press, Berkeley/Los Angeles, CA.
Scott A. J. (1998) From Silicon Valley to Hollywood: growth and development of the multimedia industry in California, in Braczyck H. J., Cooke P. and Heidenreich M. (Eds) *Regional Innovation Systems*, pp. 136–62. UCL Press, London.
Scott A. J. (1999a) The US recorded music industry: on the relations between organization, location, and creativity in the cultural economy, *Environ. Plann. A* **31**, 1,965–84.
Scott A. J. (1999b) The cultural economy: geography and the creative field, *Culture, Media, & Society* **21**, 807–17.
Scott A. J. (2000) French cinema: economy, policy and place in the making of a cultural products industry, *Theory, Culture & Society* **17**, 1–38.
Scott A. J. (2001) Capitalism, cities and the production of symbolic forms, *Trans. Inst. Brit. Geogr.* **26**, 11–23.
Scott A. J. (2002) Competitive dynamics of Southern California's clothing industry: the widening global connection and its local ramifications, *Urban Studies* **39**, 1,287–306.
Segrave K. (1997) *American Films Abroad: Hollywood's Domination of the World's Movie Screens*. McFarland, Jefferson, NC.
Smith M. (1998) Theses on the philosophy of Hollywood history, in Neale S. and Smith S. (Eds) *Contemporary Hollywood Cinema*, pp. 3–20. Routledge, London.
Storper M. (1989) The transition to flexible specialization in the US film industry: external economies, the division of labour, and the crossing of industrial divides, *Camb. J. Econ.* **13**, 273–305.
Storper M. (1993) Flexible specialization in Hollywood: a reply to Aksoy and Robins, *Camb. J. Econ.* **17**, 479–84.
Storper M. and Christopherson S. (1987) Flexible specialization and regional industrial agglomerations: the case of the US motion-picture industry, *Ann. Ass. Am. Geogr.* **77**, 260–82.
Storper M. and Scott A. J. (1995) The wealth of regions: market forces and policy imperatives in local and global context, *Futures* **27**, 505–26.

VÉRON L. (1999) The competitive advantage of Hollywood industry, *Columbia International Affairs on Line*, https//www.cc.columbia.edu/sec/dlc/ciao/wps/ve101

VOGEL H. L. (1998) *Entertainment Industry Economics: A Guide for Financial Analysis*. Cambridge University Press, Cambridge.

WASKO J. (1994) *Hollywood in the Information Age: Beyond the Silver Screen*. Polity Press, Oxford.

WATERMAN D. (1982) The structural development of the motion-picture industry, *American Econ*. **26**, 16–27.

WATERMAN D. and JAYAKAR K. P. (2000) The competitive balance of the Italian and American film industries, *Europ. J. Communication* **15**, 501–28.

WILDMAN S. S. and SIWEK S. E. (1988) *International Trade in Films and Television*. Ballinger, Cambridge, MA.

WYATT J. (1994) *High Concept: Movies and Marketing in Hollywood*. University of Texas Press, Austin, TX.

WYATT J. (1998) The formation of the major independent: Miramax, New Line, and the new Hollywood, in NEALE S. and SMITH M. (Eds) *Contemporary Hollywood Cinema*, pp. 74–90. Routledge, London.

Regional Studies
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cres20

Beyond the Creative City: Cognitive-Cultural Capitalism and the New Urbanism

Allen John Scott[a]

[a] Geography, UCLA, Hilgard Avenue, Los Angeles, CA 90095, USA. Email:

Published online: 17 Mar 2014.

Regional Studies, 2014
http://dx.doi.org/10.1080/00343404.2014.891010

Beyond the Creative City: Cognitive–Cultural Capitalism and the New Urbanism

ALLEN JOHN SCOTT
Geography, UCLA, Hilgard Avenue, Los Angeles, CA 90095, USA. Email: ajscott@ucla.edu

(Received January 2013: in revised form January 2014)

SCOTT A. J. Beyond the creative city: cognitive–cultural capitalism and the new urbanism, *Regional Studies*. Creativity is a concept whose time has come in economic and urban geography. It is also a concept that calls for enormous circumspection. An attempt is made to show that the interdependent processes of learning, creativity and innovation are situated within concrete fields of social relationships. Because much existing research on creative cities fails adequately to grasp this point, it tends to offer a flawed representation of urban dynamics and leads in many instances to essentially regressive policy advocacies. Cognitive–cultural capitalism is a more robust theoretical framework through which contemporary urbanization processes can be described. The framework of cognitive–cultural capitalism shapes the peculiar logic of learning, creativity and innovation that are observed in cities today but also has many wider and deeper impacts on urban outcomes. It has important policy implications so a critique of current policy stances derived from creative city ideas is also provided.

Cognitive–cultural economy   Creative city   Creativity   Service underclass   Symbolic analysts   Urbanization   Urban policy

SCOTT A. J. 超越创意城市：认知—文化资本主义与新的城市主义，*区域研究*。创造力之概念，在经济与城市地理学中已发展成熟，但它同时也必须慎以待之。本文试图显示，学习、创造力和创新之间相互依赖的过程，坐落于社会关係的坚实领域中。但诸多既有的创意城市文献无法充分地捕捉此一特点，因而倾向对城市的动态提出错误的再现，并导致在诸多案例中倡议本质上是倒退的政策。认知—文化资本主义，则是一个较为强健的理论架构，当代的城市化过程得以藉此描绘之。认知—文化资本主义的架构，形塑着今日城市中可观察到的特定学习、创造力与创新逻辑，却也同时对城市结果有着诸多广泛且深层的影响。它具有重要的政策意涵，也因此对由创意城市概念衍生出的当前政策立场提出批评。

认知—文化经济   创意城市   创造力   提供服务的底层阶级   符号分析人员   城市化   城市政策

SCOTT A. J. Au-delà de la ville créative: le capitalisme cognitif-culturel et le nouvel urbanisme, *Regional Studies*. La créativité est une notion qui arrive à point dans la géographie économique et urbaine. C'est une notion qui nécessite aussi beaucoup de circonspection. On essaye de démontrer que les processus interdépendants d'apprentissage, de créativité et d'innovation se situent au sein des champs concrets des rapports sociaux. Parce que beaucoup de la recherche actuelle à propos des villes créatives ne réussit pas à bien saisir ce point, elle a tendance à fournir une image imparfaite de la dynamique urbaine et, par la suite, prône dans beaucoup des cas des politiques qui sont essentiellement régressives. Le capitalisme cognitif-culturel constitue un cadre théorique plus solide par lequel on peut comprendre des processus d'urbanisation contemporains. Le cadre du capitalisme cognitif-culturel influence la logique d'apprentissage, de créativité et d'innovation qui est observable dans les villes d'aujourd'hui mais qui a des effets plus larges et plus approfondis sur le développement du milieu urbain. Les conséquences pour la politique s'avèrent importantes, ainsi on fournit aussi une critique des positions de principe remontant aux idées relatives aux villes créatives.

Économie cognitive et culturelle   Grande ville créative   Créativité   Classe marginale de service   Analystes symboliques   Urbanisation   Politique urbaine

SCOTT A. J. Jenseits der kreativen Stadt: kognitiv-kultureller Kapitalismus und der neue Urbanismus, *Regional Studies*. Die Kreativität ist ein Konzept, dessen Zeit in der Wirtschafts- und Stadtgeografie gekommen ist. Ebenso ist sie ein Konzept, das enorme Umsichtigkeit erfordert. In diesem Beitrag wird versucht zu verdeutlichen, dass die voneinander abhängigen Prozesse des Lernens, der Kreativität und der Innovation innerhalb konkreter Felder von sozialen Beziehungen angesiedelt sind. Da dieser Punkt in der bisherigen Forschung über kreative Städte meist nicht richtig erkannt wird, wird die urbane Dynamik oft fehlerhaft dargestellt, was in zahlreichen Fällen zu im Wesentlichen regressiven politischen Empfehlungen führt. Der kognitiv-kulturelle Kapitalismus bietet einen robusteren theoretischen Rahmen zur Beschreibung von modernen Urbanisationsprozessen. Der Rahmen des kognitiv-kulturellen Kapitalismus prägt die besondere Logik des Lernens, der Kreativität und der Innovation, die sich heute in Städten beobachten lassen, hat aber auch zahlreiche breitere und tiefere Auswirkungen auf die Ergebnisse in den Städten. Ebenso ist er mit wichtigen politischen Auswirkungen verbunden, weshalb auch eine Kritik der derzeitigen politischen Standpunkte aufgrund der Vorstellungen von der kreativen Stadt geliefert wird.

Kognitiv-kulturelle Wirtschaft Kreative Stadt Kreativität Dienstleistungs-Unterklasse Symbolanalysten Urbanisierung Stadtpolitik

SCOTT A. J. Más allá de la ciudad creativa: el capitalismo cognitivo–cultural y el nuevo urbanismo, *Regional Studies*. La creatividad es un concepto a la que le ha llegado su hora en la geografía económica y urbana. Es también un concepto que necesita una enorme circunspección. Aquí se intenta mostrar que los procesos interdependientes de aprendizaje, creatividad e innovación están situados en campos concretos de las relaciones sociales. Debido a que este aspecto no se entiende bien en muchos estudios sobre las ciudades creativas, se tiende a ofrecer una representación errónea de las dinámicas urbanas y esto conduce en muchos casos a defender políticas básicamente regresivas. El capitalismo cognitivo–cultural es un marco teórico más sólido mediante el que se pueden describir procesos de urbanización modernos. El marco del capitalismo cognitivo–cultural forma la lógica peculiar del aprendizaje, la creatividad y la innovación que se observan en las ciudades de hoy día, pero también tiene muchos efectos más amplios y profundos en los resultados urbanos. Debido a sus importantes repercusiones políticas, en este artículo se ofrece una crítica de las actuales posturas políticas que proceden de las ideas de la ciudad creativa.

Economía cognitiva-cultural Ciudad creativa Creatividad Clase inferior de servicios Analistas de símbolos Urbanización Política urbana

JEL classifications: O18, R10

INTRODUCTION

'Creativity' is a concept whose time has come in economic and urban geography. The withering away of fordist capitalism and the steady rise of a new cognitive–cultural economy (SCOTT, 2008) have put a premium on academic research focused on the interplay of new digital technologies, advanced forms of human capital, the logic of process and product innovation, and intellectual property; that is, on greatly intensified creative performance over a wide range of economic and social relata. As will be shown, these phenomena are in turn associated with a number of important shifts in patterns of urbanization and the character of the urban environment. But 'creativity' is also a concept that calls for enormous circumspection, not least in view of its deeply positive but also problematical resonances suggestive of inspired and avant-garde accomplishment. This paper examines one especially contentious expression of this concept in current geographical thinking, namely, its resurgence in creative city discourse. It is argued that while this discourse offers a number of quite useful insights into urbanization processes today, it also has significant blind spots and leads in many instances to essentially regressive policy advocacies. The overall approach adopted here consists in an attempt to reassemble the diverse phenomena that it describes, and to evaluate the normative actions that it advocates, within a more encompassing theoretical framework focused on the concrete realities of contemporary capitalism.

The idea of the creative city as both a descriptive figure and a policy desideratum has taken especially firm hold over the last decade or so (BAYCAN, 2011). The precise substantive meaning of the idea varies widely from one author to another, but it might be said that a kind of composite ideal vision of the creative city as it emerges in the literature includes ingredients such as an employment base comprising successful new-economy industries, a vibrant pool of talented and qualified labour, high levels of environmental quality, a dynamic cultural milieu including artists, bohemians and gays, a glamorous nightlife, recurrent festivals and spectacles, iconic architecture, and a unifying symbolic identity in the guise of a striking global brand. This description is, of course, a caricature, but one that nonetheless captures some of the main themes that have now entered into the ever-broadening international discussion on the creative city. It should be added that there are aspects of the description that certainly reflect current urban realities, and at least some of the research that proceeds under the rubric of the creative city has considerable merit. The purpose of this paper is not to deny that cities are often endowed with certain kinds of creative potentials as it is to propose a theoretical formulation that resituates these potentials in the context of a more widely ranging portrayal of urbanization dynamics in the current conjuncture. In pursuit of this goal, an attempt will be made to evaluate the extant body of creative city research while simultaneously pointing to a number of alternative horizons of investigation, and perhaps most importantly of all extending a warning to policy-makers that the quest for the creative city, at least in the terms of many current formulations, is as likely as not to be attended by heavy social costs and disappointments as it is by some sort of urban efflorescence.

BRIEF ARCHAEOLOGY OF AN IDEA

Psychologists have long been interested in the question of individual creativity and its general aetiology, but it was Jane Jacobs (JACOBS, 1984) who first alluded to the creative city, as such, in a discussion of innovative small-scale craft industries, inspired by the inquiries of SABEL (1982) into industrial development in the Third Italy.[1] Sabel himself went on to write at length (partly in partnership with Piore) about the peculiar model of agglomeration-specific innovation that seemed to be developing in the Third Italy in the late 1970s and early 1980s (PIORE and SABEL, 1984). A number of parallel research thrusts by authors such as AYDALOT (1986), OAKEY (1985) and STÖHR (1986) also raised questions about the broadly innovative character of industrial agglomerations. At about the same time, ANDERSSON (1985) came up with the related proposition that structural instabilities in particular localities might give rise to 'creative regions'. Quite independently of these efforts, the idea of integrating arts and culture into urban planning was put forward by YENCKEN (1988) who also suggested that this manner of proceeding pointed toward a new kind of creative city. It was only after about 1990, however, that the notion of the creative city really began to gather momentum. This was the year in which Glasgow was proclaimed as the European Capital of Culture, leading directly to a detailed report by COMEDIA (1991)[2] in which the creative city was portrayed prescriptively as a vortex of innovation in all spheres of life and most especially in the arts, design and new media. The Comedia report was succeeded in 1993 by Vancouver's decisive policy push to develop its artistic and cultural assets (DUXBURY, 2004). Shortly thereafter, a number of other cities (e.g. Toronto and Cologne) declared that they, too, would henceforth move in much the same direction.

Over the 1980s and the first half of the 1990s, then, various initial research and policy thrusts started to raise questions about the genetics of creativity and innovation in spatial agglomerations, even though there was a tendency, still discernible in the literature today, to restrict the term 'creativity' to activities in culture and the arts while the somewhat more sober notion of 'innovation' was more consistently invoked in analyses of the manufacturing economy and especially technology-intensive industry. Virtually all this effort resonated explicitly or implicitly with an emergent theoretical framework focused on postfordism and its expression in flexible production (cf. AMIN, 1994; ESSER and HIRSCH, 1989; SABEL and ZEITLIN, 1985; SAYER, 1989). The intellectual terrain was thus already well prepared when, in the mid-1990s, two extended statements, one by LANDRY and BIANCHINI (1995), the other by LANDRY *et al.* (1996), helped further to bring the notion of creative cities to the fore. LANDRY (2000) subsequently published a landmark manifesto that issued an all-azimuths call for the investment of creative energies in virtually every aspect of urban existence and especially for major efforts to be made in promoting the cultural life of the city. A further gloss on these propositions was added by HALL (1996, 1998) who argued that density, human interaction and synergy were essential foundations of the creativity of individual places. Alongside these lines of enquiry, a number of geographers and sociologists were starting to develop ideas about the cultural economy of cities (e.g. LASH and URRY, 1994; MOLOTCH 1996; PRATT, 1997; SCOTT, 1996), and these ideas came rapidly to be intertwined with work on the creative city.

Despite this evident ferment, research on the urban and regional foundations of creativity was still confined to a relatively small number of scholars up to the end of the 1990s. The year 2002 marks a major turning point. This was the year in which Richard Florida published his influential book on the rise of the 'creative class' (FLORIDA, 2002), an event that almost immediately sparked off an extended debate about creativity as a force in urban development. Florida restated his ideas two years later in a book that advanced the additional claim that cities could attract large numbers of creative class workers if they offered high-order amenities (among which social diversity and tolerance were said to play a major part), thereby stimulating local economic growth (FLORIDA, 2004). Florida's identification of the creative class as a significant new social stratum in contemporary capitalism was unquestionably an important insight (though anticipated in different ways by analysts such as BELL (1973), GOULDNER (1979) and REICH (1992), but his further asseverations to the effect that urban growth flowed spontaneously from the presence of this class were met – appropriately, as will be shown – with considerable scepticism (e.g. MARKUSEN, 2006; PECK, 2005).

In spite of this scepticism, Florida's work marked the beginning of a widening stream of academic work on the topic of the creative city that has continued to expand down to the present day (Fig. 1). Moreover, Florida's writings found instant echoes in the practical initiatives that were building up around the creative city idea so that his message (consolidated by his frenetic consulting efforts and extensive press coverage) found a sympathetic audience among policy-makers all around the world. This enthusiastic reception is reflected in the proliferation of cities over the last decade or so that claim to have been touched in one way or another by the viaticum of creativity. By one account, there are now over 60 self-professed creative cities worldwide (KARVOUNIS, 2010), and even such palpably improbable places – on the face of it, at least – as Sudbury, Canada (PAQUETTE, 2009), Milwaukee, USA (ZIMMERMAN, 2008), Huddersfield, UK (CHATTERTON, 2000), and Darwin, Australia (LUCKMAN *et al.*, 2009) have now jumped into the fray. As KONG and O'CONNOR (2010) indicate, the idea has caught on with special tenacity in Asian policy circles, and is notably strong in China where the cities of Beijing, Shanghai, Guangzhou, Chongquin and Wuhan (not to forget Hong Kong and Macau) are all now asserting

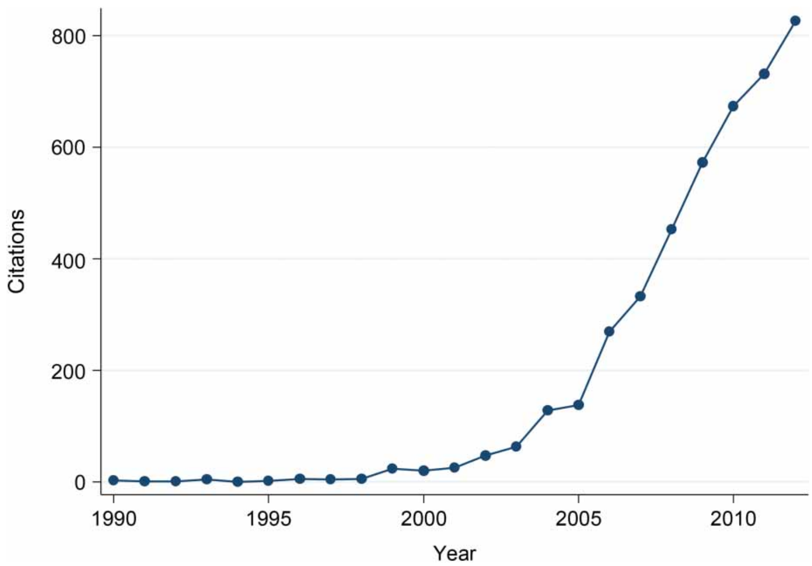


Fig. 1. Annual number of citations of the term 'creative cities' as recorded by Google Scholar, 1990–2012

their creative accomplishments and potentials. In addition, recent research on the creative city has steadily widened the terms of reference of this idea. Some of the major themes running through the current literature focus on aspects of creative urbanism revolving around the institutions of the new economy (the cultural economy above all, but other forms of enterprise as well, including high-technology industry), human capital formation, cultural policy, and innovative idioms of urban design and architecture (e.g. BONTJE *et al.*, 2011; COHENDET *et al.*, 2010; COMMUNIAN, 2011; COSTA *et al.*, 2008; CUNNINGHAM, 2012; EVANS, 2009; GREFFE, 2011; GRODACH, 2012; KAGAN and HAHN, 2011; KONG and O'CONNOR, 2010; KRÄTKE, 2011; MCCANN, 2007; PONZINI and ROSSI, 2010; PRATT, 2008; VIVANT, 2009). This research has also pointed to some of the more naïve and/or cynical appropriations of the creative city idea in the world of policy. A further emerging theme refers not only to major urban centres but also to the creative countryside and small towns in rural areas (BELL and JAYNE, 2010; LEWIS and DONALD, 2010; SCOTT, 2010b; WAITT and GIBSON, 2009).

There is much that can be defended in this overall body of work, including the deep reservations in the recent literature about any approach that starts with the so-called creative class as a sort of privileged independent variable. There is also much, however, that needs to be questioned with regard to its theoretical bases and orientation as well as its broader political implications. The discussion that follows is an attempt to rethink various aspects of all of this work by putting it in the context of 21st-century capitalism and related overarching shifts in the economic and social characteristics of cities. The attempted reconstruction here seeks to move away from 'creativity' as a foundational concept for approaching urban issues, not only because creativity itself needs to be problematized and its essential social origins revealed, but also because the term so clearly fails to capture so much that is at stake in the urbanization process today, even in those cities that have been identified as being on the cutting edge of creative performance. It will be argued that even when urban outcomes are most intimately connected to creativity, as such, one still needs to go well beyond the confines of this concept in order to achieve an adequate understanding of its motions – and limits – in the city.

DECIPHERING CREATIVITY

Creativity is an extraordinarily difficult word whose meaning is bedevilled by its oft-presumed connection with exalted states of mind, and notably with the 'mysterious' workings of artistic and scientific genius. Right at the outset, therefore, one needs to be wary of any metaphysical connotations of the term, especially its not-infrequent association with the transcendent. Instead, one can usefully begin by situating creativity between two concrete polarities, one psychological or internal to the individual, the other sociological or external. On the one side, then, creativity resides in the mental capacities and personal endowments of individual subjects. Some individuals have the native talent and acquired know-how enabling them to accomplish certain kinds of creative acts; some have little or none. Much useful analysis has been published by psychologists on these aspects of the problem (for a summary, see STERNBERG and LUBART 1999). On the other side, creativity is also embedded in concrete social contexts that shape its character and objectives in many different ways (cf. CSIKSZENTMIHALYI, 1990; HEMLIN *et al.*, 2008; SEITZ, 2003). It is this second facet of the problem that is most pertinent to the present discussion, though assuredly one must not fall into an untenable dualism that separates these two moments of creativity into watertight compartments.

In this context, one also needs to distinguish between three interconnected but rather distinctive processes, namely, learning, creativity and innovation. *Learning*, the essential complement of the 'prepared mind', is a preliminary to creativity (cf. PASTEUR, 1854); *creativity* itself is the act of producing meaningful new ideas, where the qualifiers 'meaningful' and 'new' both need to be stressed (BASTIDE, 1977); and *innovation* entails the translation of those ideas into concrete, effective outcomes. Thus, creative thinking is always in important respects moulded by the knowledge and skills of individuals. These assets are acquired to a significant degree through education, practice and informal socialization, that is, from external sources that are themselves permeated with definite historical and geographical character. Equally, knowledge and skills are bound by all manner of checks and limits (e.g. by theoretical closure, by prescriptive ideologies, by historical tradition, by habit, and so on) though the degree of rigidity of these constraints will vary greatly depending on specific circumstances (AMIN and ROBERTS, 2008; BAUTÈS and VALETTE, 2004; BROWN and DUGUID, 1991). In brief, individuals typically internalize elements of their daily environment and these are then reflected back – via further mental processing – in more or less socially conditioned creative efforts. At the same time, individuals caught up in dense transactional networks of various kinds are obviously in a more favoured position to acquire useful information and to explore its wider potentialities than those who are more socially isolated. In fact, much of the labour process in the new economy is organized specifically in ways that seek to capture and optimize these transactional aspects of creativity. As GRABHER (2001) has written, this manner of organizing work is especially evident in the case of project-oriented teams in which selected individuals are brought together for a period of time in order to pool their know-how and to cross-fertilize each other's thinking in a context of close collaboration directed to problem-solving exercises. This remark exemplifies another important dynamic of creative activity, namely, that certain kinds of disruptive communications or situations can enhance creative efforts. In this respect, NOTEBOOM (1999) suggests that interpersonal cognitive distance (further refined in terms of novelty and communicability) is an important intermediate variable in the way such disruption works. In particular, too little novelty in any given transaction is apt to be unproductive because it merely reinforces what is already known; and so is too much, because it may not be decodable at the point of reception. Intermediate doses are calculated to push creativity forward most effectively.

The point of all this is that creativity is in deeply meaningful ways a social phenomenon. What this signifies in the present account is that whatever it is that might be identified as the essential nexus of the creativity of cities, its motions can only be deciphered in the last instance, by reference to the communal, context-dependent and purpose-driving conditions that are found in the urban milieu. As will be shown, there is also always a dimension of political choice that shapes these conditions and their effects. In any case, creativity is most certainly not a purely self-acting *primum mobile* of urban development. An alternative way of referring to these matters is to say that something like an energized *creative field* extends across the city in the guise of overlapping physical and social infrastructures and that the dense, polarized, multifaceted mesh of transactions generated within this field is a major factor in moulding locally distinctive patterns of ingenuity and imagination (SCOTT, 1999, 2006, 2010a). It should be added that creative fields can be identified at many other levels of scale, including the national and the global, but for present purposes attention is focused solely on the urban.[3] What is the specific concrete character of the creative field of the city? How does it reflect the logic of urbanization at the beginning of the 21st century? What are its potentialities and limits? And once these questions have been dealt with, what remains of a creative city problematic? These issues can now be dealt with by an extended discussion of urban processes in the emerging, transformative world of cognitive–cultural capitalism. The discussion will underline the legitimacy and urgency of concerns about creativity in contemporary cities, but will, at the same time, seek to dissolve these concerns into what is taken to be a considerably more robust theoretical framework.

CITIES IN COGNITIVE–CULTURAL CAPITALISM

Over the history of capitalism, a distinction was frequently made between cities of industry and commerce, on the one hand, and cities of art and culture, on the other, and, indeed, these two forms of urban development were widely seen as being quite incompatible with one another. Today, this distinction is disappearing in favour of a more syncretic view of cities that is in some degree captured under the rubric of the 'postfordist city', one of whose declinations is the 'creative city', i.e. a city where production, work, leisure, the arts and the physical milieu exist in varying degrees of mutual harmony.

This trend is in significant ways related to a number of important transformations that have come about in the shift from fordism to postfordism that occurred over the 1980s and 1990s and that have a strong bearing on learning, creativity and innovation. Four main points follow immediately from this remark. First, the main capitalist economies have come to focus increasingly on unstandardized products in sectors like technology-intensive production, business, financial and personal services, and a wide array of cultural industries ranging from the media to fashion-intensive crafts. Second, the same sectors display definite tendencies to extensive horizontal and vertical disintegration with selected groups of firms then being recomposed into networks of specialized but complementary producers with strong proclivities to agglomeration, especially in large cities. Third, increasing shares of the output of these sectors are marked by firm- and place-specific product specifications. Fourth, the conspicuous growth of the new economy is echoed in the expansion of a labour force (roughly equivalent to Florida's creative class, or what Reich, 1992, has called 'symbolic analysts') that is called upon increasingly to deploy high-level cognitive and cultural skills such as deductive reasoning capacities, technical insight, leadership, communication abilities, cultural awareness and visual imagination – in other words, more or less creative capacities – in the workplace. As these transformations have come about, economy and culture have fused together in important ways, in the sense that economic outputs are subject to ever-increasing injections of aesthetic and semiotic meaning, while the culture that is consumed is produced more and more by profit-seeking firms in the commodity form.

In line with these remarks, and as advocated elsewhere, the term 'postfordism' should doubtless be abandoned and replaced by the more affirmative designation 'cognitive–cultural capitalism' (Scott 2011, 2012). The former label has the disadvantage of expressing itself only by what it is not, whereas the latter has the advantage of positively reflecting the foundations of much contemporary economic activity – above all in the more advanced centres of capitalism – in the cerebral and affective capacities of the labour force. It can be argued that we are now entering a period marked by a distinctive third wave of urbanization based on cognitive–cultural capitalism, in contradistinction to a first wave associated with the 19th-century factory and workshop system and a second wave associated with 20th-century fordism (Scott, 2011). This remark is not intended to suggest that all cities have entered this phase as equal participants, but it is certainly the case that more and more large cities in North America and Western Europe are taking part in this trend, as well as cities in the Asia-Pacific Region and elsewhere. A very partial illustrative list of such cities would include New York, Los Angeles, London, Paris, Amsterdam, Barcelona, Sydney, Tokyo, Seoul, Singapore, Hong Kong and Bangkok. At the same time, many smaller cities and even rural areas are increasingly subject to transformation within this cognitive–cultural order (Scott, 2012).

The contention here is that these new socio-economic arrangements in third-wave cities generate distinctive creative fields while simultaneously providing terms of reference that point well beyond a narrow focus on creativity as such. A core element of the creative fields within these cities consists of the clusters of technology-intensive, service and cultural producers that nowadays constitute so much of their economic base. These clusters are sites of productive labour in which know-how is accumulated, skills are honed, and firm- and place-specific forms of product configuration (the source of Chamberlinian monopolistic competition) are worked out. The same clusters are shot through with dense transactional networks that in part hold them together as geographical entities, and that function as important channels of information and mutually cross-fertilizing creative signals (Asheim and Coenen, 2005; Bathelt *et al.*, 2004; Desrochers and Leppälä, 2011). The labour markets that form around these clusters are also important sources of creative energies as workers rotate through different jobs and interact with one another in project-oriented teams. At the same time, the wider urban system plays a role in helping to sustain (or impede) individual creative drives via the forms of social reproduction that it sustains. In this regard, the cultural and social spaces of the city as well as local institutions, like schools and universities, are of major consequence (Scott, 2010a). In favoured cases, the physical fabric and cultural institutions of the city help to underpin these functions by providing a milieu that is supportive of the overall social and economic vocation of the city, as well as exuding images that function as branding devices and as advertisements for its capabilities, attractions, and ambitions (Okano and Samson, 2010; Vanolo, 2008).

The most visible expression of the latter phenomenon can be found in the central business districts of major global cities where peculiar new forms of aestheticization, as represented above all by markedly idiosyncratic buildings signed by star architects, amplify the individuality and visibility of the urban milieu. The historical and cultural patrimony of cities serves much the same purpose, and often forms the basis of lucrative heritage and place marketing efforts (Graham *et al.*, 2000; Philo and Kearns, 1993). Even many former manufacturing centres are attempting to upgrade parts of their ageing built stock in efforts to construct some sort of 'creative' future for themselves, above all one that emphatically involves a break with the old economy and that reaches out to more knowledge- and culture-intensive forms of production. Initiatives like these are evident in the widespread recycling of derelict industrial and warehouse properties in cities that prospered in earlier phases of urban economic development, and their utilization for art centres and galleries, music venues, boutique retail outlets, and facilities such as small design, media and fashion firms (cf. Andres and Grésillon, 2011; Brown *et al.*, 2000; Communian, 2011; Gnad, 2000;

Wynne, 1992). In brief, and to repeat an earlier refrain, forms of creative expression in the contemporary city are not simply sovereign emanations from the minds of the citizenry but are also mobilized and moulded by the complex interweaving between relations of production, social life and the urban milieu at large.

These trends have been associated with a notable if selective resurgence of urban growth over the last few decades, and this in turn has sparked off a debate – deeply interwoven with the creative city idea – about the causalities underlying this growth. One fashionable view about this matter can be summarized in Florida's claim that cities with abundant amenities are apt to grow because the creative class will preferentially migrate to such cities, and their presence will then be reflected in bursts of local economic dynamism (see also Florida, 2008; and Mellander *et al.*, 2013, for extensions). One variation on this theme can be found in Clark (2004) who views successful modern cities as 'entertainment machines' that attract and maintain highly qualified workers by virtue of their amusement value. Another is offered by Glaeser (2011) for whom cities have evolved from being centres of production in the 20th century to consumer cities in the 21st, and he avers that the latter cities grow when they have abundant amenities (sunshine, low crime rates, 'playground' effects, etc.) that draw in workers endowed with high levels of human capital. In opposition to these amenity-based and sumptuary views on urban dynamism, Storper and Scott (2009) have argued that cities develop primarily on the basis of their job-generating capacities. This argument does not dismiss the auxiliary role of human capital in the recursive, path-dependent process of urban growth. However, it does claim that even today it is jobs not amenities that attract highly qualified workers to particular cities and that keep them durably in place. The complexities of this debate are much too great to be summarized in the present context, except perhaps to say that one of the fatal flaws of the amenities-based view is that it is quite incapable of explaining how it is that the production systems of cities, and their corresponding stocks of human capital, are so frequently specialized. If the view as articulated by Storper and Scott is correct, it means that even so-called creative cities (or, in what is taken to be a more theoretically informed vocabulary, cities with vibrant cognitive–cultural economies) remain directly dependent for their growth and prosperity on the health of their internal productive arrangements (and, of course, export markets). This statement holds as much for cities dominated by sectors like tourism, the heritage industry, theatre and casino gambling, whose outputs are largely immobile, as it does for cities dominated by technology-intensive sectors, business and financial services, and cultural industries, producing outputs that can normally be readily transported to distant markets.

CHIAROSCURO: INTRA-URBAN SOCIAL AND SPATIAL TRANSFORMATIONS

Human capital and social restratification. Much, though by no means all, of the research accomplished on the creative city to date celebrates the ostensibly benign dimensions of this concept while neglecting or deemphasizing its more malignant elements. But cities in advanced capitalism, even those that are well-endowed with advanced cognitive and cultural employment opportunities, are in practice awash in crisis conditions – exacerbated by fiscal austerity in these last few years – not in a purely contingent way, but as an organic outgrowth of their role as cynosures of the new economy. Among the problems and predicaments created in this manner, those that revolve around the widening economic and social gap between the upper and lower halves of the labour force are undoubtedly the most explosive in political terms.

In classical fordist society the dominant division of labour in production was represented by the white- and blue-collar fractions of the workforce. This division of labour was projected out into intra-urban space where it reappeared in the form of a pervasive but never fully accomplished division of neighborhoods. With the advent of the new economy, an alternative bipartite division of labour is now overriding the old white/blue-collar social split (Levy and Murnane, 2004) leading to a significant restratification of urban society and thence to significant readjustments in the geography of urban neighborhoods. The new division of labour is represented on the one side by the creative class or symbolic analysts, and on the other by a low-wage service underclass. In more polemical terms, the latter group might be said to constitute a new servile class. Standing (2011) and others refer to a 'precariat' that is more or less equivalent to this class. The steadily widening gap in the fortunes of these two strata is in large part a function of their differential command of skills and formal qualifications, but is greatly exacerbated by the increasing atomization and competitiveness that have invaded labour markets as neoliberal fundamentalism and globalization have continued to run their course. The deepening divide in advanced capitalist cities is all the more evident given the steady retreat of manufacturing employment (by both attrition and movement offshore) from major urban centres and the concomitant shrinkage of the traditional blue-collar labour force.

The human capital assets of the top half of the urban labour force comprise advanced technical knowledge, analytical prowess and relevant forms of socio-cultural know-how. These assets are typically valorized within a system of formal credentialing. By contrast, the abilities of the service underclass are very much more informal and undervalued. These workers are deployed in jobs whose subaltern status is underlined by the high proportion of politically marginal social groups (such as immigrants from poor countries) who carry them out. These jobs are focused above all on sustaining the facilities and infrastructure of the urban system, and on providing diverse kinds of domestic and personal help. This remark implies at once that service underclass work is geared in large degree to providing for the direct and indirect demands of the upper echelon of workers in the cognitive–cultural economy. It would be a grave mistake, however, to suppose that the labour of the service underclass is devoid of cognitive and cultural skills. Consider, for example, the kinds of flexibility and discretionary decision-making required for janitors, motor-vehicle operators, and crossing guards to accomplish their work, or the communicative and resourceful capacities that must be mobilized by child care workers, home health aides, and beauty salon personnel. Moreover, because the members of the service underclass must always be available in proximity to the point of service, the tasks that they perform, unlike much manufacturing activity, cannot be repackaged and sent offshore (GATTA *et al.*, 2009). Indeed, the incidence of these workers has been rising in both absolute and relative terms in American cities of late. Thus, among the fastest growing jobs in large US cities in the first decade of the 21st century are found such low-wage service occupations as couriers and messengers, dining room attendants, dishwashers, food preparation workers, grounds maintenance workers, highway maintenance workers, hotel clerks, miscellaneous motor vehicle operators, painters, parking lot attendants, and service station attendants (SCOTT, 2009, 2010c).

Land use changes. As the social shifts noted above have occurred, significant rearrangements of intra-urban space have also come about. Among the more dramatic of these changes is the revitalization of selected areas in the city, most especially in and around the urban core. This form of revitalization comprises two related but distinctive phases, each frequently equated with creative city dynamics. One involves the upgrading of deteriorated residential areas, notably but not exclusively, in inner-city areas; the other is focused on the redevelopment of commercial and business properties within the central business district. Both phases are commonly referred to in the literature by reference to ‘gentrification’, though the term is unsatisfactory in many ways. What follows will continue to refer to residential upgrading as gentrification, but the more evocative phrase ‘aestheticized land use intensification’ will be used to refer to commercial and business land-use redevelopment in central business districts.

Gentrification, as such, was first identified by GLASS (1964) who observed that a number of poorer neighborhoods in Central London were undergoing social transformation as upper middle-class families started to take over much of the local housing stock. The historical origins of gentrification thus precede the rise of the cognitive–cultural economy, though it might well be maintained that Central London was a harbinger of things to come with its already declining manufacturing industries and its shifting occupational structure reflecting the rapid growth of its financial, media, fashion and entertainment sectors. In many instances, the incipient stages of gentrification are signalled by incursions of artists and bohemians into run-down working-class neighborhoods (LLOYD, 2002; ZUKIN 1982). A pioneer fringe of middle class gentrifiers then frequently starts to move in and renovate local properties, followed by successive waves of further gentrification and rapidly increasing property values. These events are accompanied by steady displacement of the original low-income residents as rental rates rise and as what remains of their local employment base in manufacturing is demolished. The pace of change is frequently boosted by unscrupulous landlords eager to reap the benefits of higher property values and by overzealous city councils anxious to enhance the image of the city (LEES *et al.*, 2008; SLATER, 2006; WACQUANT, 2008). As this happens, inner-city residential areas become increasingly dominated by ‘creative’ workers with demographic profiles like young professional families, cohabiting couples, people in same-sex unions, apartment sharers, metrosexual singles, and so on (HAASE *et al.*, 2010; HAMNETT and WHITELEGG, 2007; HARRIS, 2008).

The allied process of aestheticized land-use intensification can be described as the quest for increased productivity per unit of urban land in central business districts (as well as more outlying business clusters), particularly with regard to cognitively and culturally inflected sectors. Today, this quest is typically associated with forms of embellishment coinciding with strikingly new idioms of architecture and urban design that are very different from the ageing modernist style that formerly dominated downtown areas in different parts of the world. These idioms in many ways reflect something of the character of the cognitive–cultural activities that now dominate in downtown areas as well the ideology and tastes of the new transnational capitalist class whose members move with ease between the various global foci of new economy (SKLAIR, 2005). They are also a reflection of an aggressive global urbanism that thrives on dramatized city branding strategies not only in the interests of self-assertion but also as a means of attracting inflows of capital investment and highly qualified labour.

A widely accepted theoretical explanation of these processes of land-use redevelopment and change was proposed by SMITH (1982, 1986) in terms of what he called a ‘rent gap’. Smith argues that inner-city properties often seem to command potential rents far above the actual rent earned and that this gap provides the incentive for land-use upgrading and socio-economic succession. A phenomenon analogous to the rent gap can certainly be observed at different times in different cities. However, it is suggested that it should more properly be seen not as a cause but as an effect of gentrification (in all of its senses), or perhaps, better yet, an endogenous element of the gentrification process relative to the wider economic and social changes currently going on in cities. Above all, these changes need to be considered in the light of two basic and mutually reinforcing trends. On the one hand, these peculiar forms of redevelopment owe much to the collapse of manufacturing employment in inner-city areas and the erosion of adjacent working-class neighborhoods. On the other hand, they also reflect the enormous recent expansion of the new economy in central business districts, leading in turn to revalorization of surrounding residential areas and their steady colonization by cognitive–cultural workers. Central business district redevelopment has also typically been accompanied by major investments in cultural and entertainment facilities such as museums, art galleries, music venues and sports arenas, and this has added to the attractions of nearby neighborhoods for well-paid, well-qualified, cognitive and cultural workers.

As these developments have moved ahead, patterns of socio-spatial segmentation in cities have been significantly reshaped by comparison with the situation under fordism. In some respects, socio-spatial segmentation has actually become more strongly indurated than it was in the fordist era as the incomes of different occupational strata diverge and as members of the upper tier of the labour force increasingly secure their own residential seclusion by means of gated communities and stringent zoning regulations, not only in central cities but also in the suburbs (Blakely and Snyder, 1997; Le Goix, 2005). The incongruity of cities of the third wave, despite frequent claims about their creative vitality, is the glaring contrast between the shimmer of their more affluent landmark areas and the squalor of their darker underside. Increasing numbers of cities in both the Global North and Global South today are marked by a yawning void between their internal islands of prosperity linked to the global economy, and widely ranging tracts where social and political marginalization is the order of the day. In these circumstances, the right to the city that Lefebvre (1968) saw, correctly, as one of the basic conditions for a renewal of democratic values, social solidarity and the capacity for *réjouissance* seems as distant as ever. Some of the most creative cities in today's world are deeply caught up in this stubborn predicament.

THE URBAN POLICY DIMENSION

In its unreconstructed form, creative city theory offers a seductively glowing vision of urban potentialities. The call to creativity as a centrepiece of urban development goals is certainly compelling by comparison with politically fraught and often ineffectual measures focused on more traditional approaches entailing attempts to improve the local business climate and efforts to lure inward investors by means of fiscal incentives. Small wonder, in view of this vision, that large numbers of self-declared creative cities are asserting themselves on every continent at the present moment in history.

To be sure, cities have always been centres of creativity, even if in different ways at different times (Andersson, 2011; Hall, 1998; Hessler and Zimmerman, 2008). Today, forms of urban creativity reside substantially in the specific socio-economic relationships built into the new cognitive and cultural order. As already noted, however, the descriptive and normative resonances of 'creative city' discourse systematically occlude much of what is most fundamental in the contemporary urban process and are hence prone both to overlook many crucial questions and to generate dysfunctional policy advocacies. It is undeniably the case that a great deal of what goes on in contemporary cities is symptomatic of diverse creative impulses, especially in the guise of the continual destruction and reconstruction of ideas, images, styles, routines, organizational arrangements, and all the rest. These impulses are embedded in a new socio-economic regime of volatility, flexibility, intensified competition and a type of consumerism that is increasingly focused on the aesthetic, semiotic and libidinal content of the products that circulate through global markets. These circumstances, by and large, remain beyond the purview of creative city discourse, which in its uncritical optimism, expresses a sort of credulity about the urban future compounded by policy recommendations that in practice, as Pratt (2011) has argued, lead frequently to regressive policy outcomes. For one thing, as noted above, creative city policies help to turbo-charge gentrification processes thus exacerbating the exclusion of low-income families from central city areas and underwriting the takeover of those areas by the new bourgeoisie (Bayliss, 2007; McCann, 2007; Smith, 2002). The irony here is that although creative city theory puts much emphasis on diversity and tolerance, the policy advocacies that have been constructed in its name actually make few gestures in the direction of social inclusion and even less in the direction of income redistribution. For another thing, creative city discourse all too often licenses deeply flawed programmes that proceed on the faith that investments in amenities will function as a magnet for creative class migrants and that this will then presumably foster rising urban prosperity for all. It has already been suggested that this faith is misplaced, and several authors have shown that in many instances the expected returns to major investment in urban amenities far exceed the actual returns, notably when policy-makers have adopted exaggerated hopes based on mediatized but dubious models like Bilbao's attempted renaissance by means of its Guggenheim Museum (cf. Evans, 2005; Jayne, 2004; Paquette, 2009; Sasaki, 2010). Moreover, investments in amenities calculated to appeal primarily to the creative class are *ipso facto* liable to involve regressive subsidies to privileged groups at the expense of other social fractions. Such investments are particularly likely to provide disproportional benefits to property owners (O'Connor and Gu, 2013).

It is no surprise, therefore, to observe that popular political movements in large cities everywhere, have rather consistently turned their backs on these kinds of policy advocacies in favour of goals that address the specific needs of low-wage workers, the unemployed and the destitute (Borén and Young, 2012; Holm, 2010). Alternative movements like these are well exemplified by the Bus Riders Union and the Janitors for Justice campaign in Los Angeles (cf. Milkman, 2006; Soja, 2010), or by the community-based workers' centres that offer help in cities across the United States to low-income and especially immigrant workers in their quest for improved wages and working conditions (Fine, 2005). Even groups of artists (who, on the basis of narrow self-interest, would seem to have much to gain from creative city policies) have started to campaign for socially inclusive approaches in cities as far apart as Hamburg and Toronto, and to call for reconsideration of some of the more overtly regressive and philistine public initiatives intended to bring creative city ideas into concrete realization (Catungal *et al.*, 2009; Holm, 2010).

More generally, the tasks of harnessing and regulating urban realities in the interests of a progressive future require initiatives that recognize the positive, forward-looking energies of cities, but that go radically beyond the advocacies of creative city enthusiasts. Three imperatives, responding to core economic and social breakdowns in the large city today, are of particular importance and urgency. The first is to build institutional frameworks that can effectively manage the common-pool resources that abound within the cognitive–cultural economy at the scale of the individual city and that are otherwise susceptible to gross inefficiencies. The second is to rectify the huge discrepancies of incomes and life chances that currently distort the social landscape of large cities all over the world. The third is to secure the wider democratization of urban space and to promote the rehabilitation of communal life. These desiderata are essential for achieving the full developmental possibilities of third wave cities, and for countering some of the more frankly rapacious and narcissistic qualities of existence in the world of the new cognitive–cultural economy.

TAKING STOCK

Despite the critical comments made here on creative city theory and policy-making, it should be re-emphasized that much of the substantive content of this approach echoes some very real currents in contemporary society, albeit in a wayward and distorted manner. These currents stem from a number of important dimensions of the creative field in contemporary cities. Hence, cities, large and small, in many different parts of the world are most assuredly being transformed in economic terms as the new cognitive–cultural economy deepens and widens its hold; even rural areas are participating in this shift. A distinctive stratum of highly paid workers with well-honed cerebral and affective human capital is also coming to the fore in large cities as these changes occur, though this ostensibly cheerful outcome is counterbalanced by the emergence of a low-wage service underclass and all that this implies in terms of the socio-spatial segmentation of urban life. The resurgence and rising wealth of cities worldwide over the last few decades are further reflected in significant upgrading of selected parts of the urban environment together with expanding stocks of high-grade amenities like museums, concert halls, libraries, recreational facilities and public art. And in the context of the intensifying race for competitive advantage and the quest for inward flows of investment and human capital, there is continually intensifying pressure on cities to assert their global presence and ambitions by means of vibrant visual images and branding campaigns emphasizing local attractions such as lifestyle, cultural facilities and historical heritage.

Then again, as this paper has tried to show, these trends and the creative impulses that they promote are subjacent to a much wider set of social and economic forces rooted in the dynamics of cognitive–cultural capitalism. The primary theoretical challenge therefore is to reveal how these dynamics undergird the spatial and temporal logic of urbanization today. An exclusive focus on the creativity-engendering capacities of the city, as such, misses much of what is most crucial in this challenge, namely, the social and economic forces that bring specific modes of urban life into being in the first instance. By the same token, what Peck (2005) calls the 'creative cities script' assuredly contains an unwarranted dose of wishful thinking not to mention its encouragement of top-down, leadership-style political recuperation and regressive policy-making (see also Campbell, 2013). So even if its time has come, the concept of creativity in economic and urban geography needs to be approached with all due caution.

Two further observations can be appended to this injunction. First, to the degree that creativity emerges out of the complex physical and social infrastructures of the city it is an epiphenomenon; and second, to the degree that it is invoked – as is so often the case – simply as a synonym for upgrading the building stock and cultural facilities of the city, it is a misnomer. Still, many striking new prospects for urban development are now coming to the fore by reason of the powerful and intensifying interplay between cognition, culture and economy in the capitalism of the 21st century. A basic condition for the full flowering of these prospects is the reining in of the neoliberal frameworks that regulate so much of contemporary city governance, and a dramatic enlargement of the sphere of urban democratic order.

Acknowledgements – This paper is based on the Regional Studies Association Annual Lecture given by

Professor Scott at the Association of American Geographers Annual Conference in Los Angeles, CA, USA, April 2013.

NOTES

1. The Third Italy coincides with the regions of Emilia-Romagna, Friuli-Venezia Giulia, Marche, Trentino-Alto Adige, Tuscany, Umbria, and Veneto. Sabel himself was greatly influenced by Italian theorists like Arnaldo Bagnasco, Sebastiano Brusco and Giacomo Becattini.
2. Comedia was founded by Charles Landry in 1978.
3. Of course, a fuller treatment of these fields would also pay attention to their interaction, and, in the end, to their essential fusion.

REFERENCES

Amin A. (Ed.) (1994) *Post-Fordism: A Reader*. Blackwell, Oxford.

Amin A. and Roberts J. (2008) Knowing in action: beyond communities of practice, *Research Policy* **37**, 353–369.

Andersson Å. (1985) Creativity and regional development, *Papers in Regional Science* **56**, 5–20.

Andersson Å. (2011) Creative people need creative cities, in Andersson D. E., Andersson Å. E. and Mellander C. (Eds) *Handbook of Creative Cities*, pp. 14–55. Edward Elgar, Cheltenham.

Andres L. and Grésillon B. (2011) Les figures de la friche dans les villes culturelles et créatives: regards croisés européens, *L'Espace Géographique* **1**, 15–30.

Asheim B. T. and Coenen L. (2005) Knowledge bases and regional innovation systems: comparing Nordic clusters, *Research Policy* **34**, 1173–1190.

Aydalot P. (1986) Trajectoires technologiques et milieux innovateurs, in Aydalot P. (Ed.) *Milieux Innovateurs en Europe*, pp. 345–361. Groupe de Recherche Européen sur les Milieux Innovateurs (GREMI), Paris.

Bastide R. (1977) *Sociologie de l'Art*. Payot, Paris.

Bathelt H., Malmberg A. and Maskell P. (2004) Clusters and knowledge: local buzz, global pipelines, and the process of knowledge creation, *Progress in Human Geography* **28**, 31–56.

Bautès N. and Valette E. (2004) Miniature painting, cultural economy and territorial dynamics in Rajasthan, India, in Scott A. J. and Power D. (Eds) *Cultural Industries and the Production of Culture*, pp. 207–223. Routledge, London.

Baycan T. (2011) Creative cities: context and perspectives, in Fusco Girard L., Baycan T. and Nijkamp P. (Eds) *Sustainable City and Creativity*, pp. 15–54. Ashgate, London.

Bayliss D. (2007) The rise of the creative city: culture and creativity in Copenhagen, *European Planning Studies* **15**, 889–903.

Bell D. (1973) *The Coming of Post-Industrial Society: A Venture in Social Forecasting*. Basic Books, New York, NY.

Bell D. and Jayne M. (2010) The creative countryside: policy and practice in the UK rural cultural economy, *Journal of Rural Studies* **26(3)**, 209–218.

Blakely E. J. and Snyder M. G. (1997) *Fortress America: Gated Communities in the United States*. Brookings Institution Press, Washington, DC.

Bontje M., Musterd S. and Pelzer P. (2011) *Inventive City-Regions: Path Dependence and Creative Knowledge Strategies*. Ashgate, Farnham.

Borén T. and Young C. (2012) Getting creative with the creative city? Towards new perspectives on creativity in urban policy, *International Journal of Urban and Regional Research* **37**, 1799–1815. DOI:10.1111/j.1468–2427.2012.01132.x.

Brown J. S. and Duguid P. (1991) Organizational learning and communities of practice: toward a unified view of working, learning, and innovation, *Organization Science* **2**, 40–57.

Brown A., O'Connor J. and Cohen S. (2000) Local music policies within a global music industry: cultural quarters in Manchester and Sheffield, *Geoforum* **31**, 437–451.

Campbell P. (2013) Imaginary success? The contentious ascendance of creativity, *European Planning Studies* DOI:10.1080/09654313.2012.753993.

Catungal J. P., Leslie D, and Lii Y. (2009) Geographies of displacement in the creative city: the case of Liberty Village, Toronto, *Urban Studies* **46**, 1095–1114.

Chatterton P. (2000) Will the real creative city please stand up?, *City: Analysis of Urban Trends, Culture, Theory, Policy, Action* **4**, 390–397.

Clark T. N. (2004) Introduction: taking entertainment seriously, in Clark T. N. (Ed.) *The City as an Entertainment Machine*, pp. 1–18. Elsevier, Amsterdam.

Cohendet P., Grandadam D. and Simon L. (2010) The anatomy of the creative city, *Industry and Innovation* **17**, 91–111.

Comedia (1991) *Making the Most of Glasgow's Cultural Assets: The Creative City and its Cultural Economy: Final Report*. Glasgow Development Agency, Glasgow.

Communian R. (2011) Rethinking the creative city: the role of complexity, networks, and interactions in the urban creative economy, *Urban Studies* **48**, 1157–1179.

Costa P., Magalhães M., Vasconcelos B. and Sugahara G. (2008) On 'creative cities' governance models: a comparative approach, *Service Industries Journal* **28**, 393–413.

Csikszentmihalyi M. (1990) The domain of creativity, in Runco M. A. and Albert R. S. (Eds) *Theories of Creativity*, pp. 190–212. Sage, Newbury Park, CA.

Cunningham S. (2012) The creative cities discourse: production and/or consumption?, in Anheier H. and Isar Y. R. (Eds) *Cities, Cultural Policy and Governance*, pp. 111–121. Sage, Los Angeles, CA.

Desrochers P. and Leppälä S. (2011) Creative cities and regions: the case for local economic diversity, *Creativity and Innovation Management* **20**, 59–69.

Duxbury N. (2004) *Creative Cities: Principles and Practices*. Canadian Policy Research Networks, Ottawa, ON.

Esser J. and Hirsch J. (1989) The crisis of fordism and the dimensions of a 'postfordist' regional and urban structure, *International Journal of Urban and Regional Research* **13**, 417–437.

Evans G. (2005) Measure for measure: evaluating the evidence of culture's contribution to regeneration, *Urban Studies* **42**, 959–983.

Evans G. (2009) Creative cities, creative spaces, and urban policy, *Urban Studies* **46**, 1003–1040.

Fine J. (2005) *Worker Centers: Organizing Communities at the Edge of the Dream*. Briefing Paper. Economic Policy Institute (available at: http://www.epi.org/publications/entry/bp159).

Florida R. (2002) *The Rise of the Creative Class*. Basic Books, New York, NY.

Florida R. (2004) *Cities and the Creative Class*. Routledge, London.

Florida R. (2008) *Who's Your City? How the Creative Economy is Making Where You Live the Most Important Decision of Your Life*. Basic Books, New York, NY.

Gatta M., Boushey H. and Appelbaum E. (2009) High-touch and here-to-stay: future skills demands in US low wage service occupations, *Sociology – Journal of the British Sociological Society* **43**, 968–989.

Glaeser E. L. (2011) *The Triumph of the City*. Penguin, New York, NY.

Glass R. (1964) *Introduction to London: Aspects of Change*. Centre for Urban Studies, London.

Gnad F. (2000) Regional promotion strategies for the culture industries in the Ruhr area, in Gnad F. and Siegmann J. (Eds) *Culture Industries in Europe: Regional Development Concepts for Private-Sector Cultural Production and Services*, pp. 172–177. Ministry for Economics and Business, Technology and Transport of the State of North Rhine-Westphalia, and Ministry for Employment, Social Affairs and Urban Development, Culture and Sports of the State of North Rhine-Westphalia, Dusseldorf.

Gouldner A. (1979) *The Future of Intellectuals and the Rise of the New Class*. Seabury, New York, NY.

Grabher G. (2001) Ecologies of creativity: the village, the group, and the heterarchic organization of the British advertising industry, *Environment and Planning A* **33**, 351–374.

Graham B., Ashworth G. J. and Tunbridge J. E. (2000) *A Geography of Heritage: Power, Culture and Economy*. Arnold, London.

Greffe X. (2011) La ciudad creative, in Manito F. (Ed.) *Ciudades Creativas: Economía Creativa, Desarollo Urbano y Políticas Públicas*, pp. 26–51. Fundación Kreativa, Barcelona.

Grodach C. (2012) Before and after the creative city: the politics of urban cultural policy in Austin, Texas, *Journal of Urban Affairs* **34**, 81–97.

Haase A., Kabisch S., Steinführer A., Bouzarovski S., Hall R. and Ogden P. (2010) Emergent spaces of reurbanisation: exploring the demographic dimension of inner-city residential change in a European setting, *Population Space and Place* **16**, 443–463.

Hall P. (1996) High-technology industry in the New York Metropolitan area: a view from history, *Annals of the New York Academy of Sciences* **787**, 42–66.

Hall P. (1998) *Cities in Civilization*. Pantheon, New York, NY.

Hamnett C. and Whitelegg D. (2007) Loft conversion and gentrification in London: from industrial to postindustrial land use, *Environment and Planning A* **39**, 106–124.

Harris A. (2008) From London to Mumbai and back again: gentrification and public policy in comparative perspective, *Urban Studies* **45**, 2407–2428.

Hemlin S., Allwood C. M. and Martin B. R. (2008) Creative knowledge environments, *Creativity Research Journal* **20**, 196–210.

Hessler M. and Zimmerman C. (2008) Introduction: Creative urban milieus – historical perspectives on culture, economy and the city, in Hessler M. and Zimmerman C. (Eds) *Creative Urban Milieus: Historical Perspectives on Culture, Economy and the City*, pp. 11–38. Campus, Frankfurt.

Holm A. (2010) Urbanisme Néolibéral ou droit à la ville, *Multitudes* **43**, 86–91.

Jacobs J. (1984) *Cities and the Wealth of Nations*. Random House, New York, NY.

Jayne M. (2004) Culture that works? Creative industries development in a working-class city, *Capital and class* **28**, 199–210.

Kagan S. and Hahn J. (2011) Creative cities and (un)sustainability: from creative class to sustainable creative cities, *Culture and Local Governance* **3**, 11–27.

Karvounis A. (2010) Urban creativity: the creative city paradigm, *AthensID* **8**, 53–81.

Kong L. and O'Connor J. (2010) Introduction, in Kong L. and O'Connor J. (Eds) *Creative Economies, Creative Cities: Asian–European Experiences*, pp. 1–5. Springer, Berlin.

Krätke S. (2011) *The Creative Capital of Cities*. Wiley-Blackwell, New York, NY.

Landry C. (2000) *The Creative City: A Toolkit for Urban Innovators*. Earthscan, London.

Landry C. and Bianchini F. (1995) *The Creative City*. Demos, London.

Landry C., Bianchini F., Ebert R., Gnad F. and Kunzmann K. R. (1996) *The Creative City in Britain and Germany*. Anglo-German Foundation for the Study of Industrial Society, Berlin and London.

Lash S. and Urry J. (1994) *Economies of Signs and Space*. Sage, London.

Le Goix R. (2005) Gated communities: sprawl and social segregation in southern California, *Housing Studies* **20**, 323–343.

Lees L., Slater T. and Wyly E. (2008) *Gentrification*. Routledge, London.

Lefebvre H. (1968) *Le Droit à la Ville*. Anthropos, Paris.

Levy F. and Murnane R. J. (2004) *The New Division of Labor: How Computers are Creating the Next Job Market*. Russell Sage Foundation, New York, NY.

Lewis N. M. and Donald B. (2010) A new rubric for 'creative city' potential in Canada's smaller cities, *Urban Studies* **47**, 29–54.
Lloyd R. (2002) Neo-Bohemia: art and neighborhood development in Chicago, *Journal of Urban Affairs* **24**, 517–532.
Luckman S., Gibson C. and Lea T. (2009) Mosquitoes in the mix: how transferable is creative city thinking?, *Singapore Journal of Tropical Geography* **30**, 70–85.
Markusen A. (2006) Urban development and the politics of a creative class: evidence from a study of artists, *Environment and Planning A* **38**, 1921–1940.
McCann E. J. (2007) Inequality and politics in the creative city-region: questions of livability and state strategy, *International Journal of Urban and Regional Research* **31**, 188–196.
Mellander C., Florida R., Asheim B. and Gertler M. (2013) *The Creative Class Goes Global*. Routledge, London.
Milkman R. (2006) *L.A. Story: Immigrant Workers and the Future of the US Labor Movement*. Russell Sage Foundation, New York, NY.
Molotch H. (1996) LA as design product: how art works in a regional economy, in Scott A. J. and Soja E. W. (Eds) *The City: Los Angeles and Urban Theory at the End of the Twentieth Century*, pp. 225–275. University of California Press, Berkeley and Los Angeles, CA.
Noteboom B. (1999) Innovation, learning and industrial organization, *Cambridge Journal of Economics* **23**, 127–150.
Oakey R. P. (1985) *Innovation and Regional Growth in Small High Technology Firms: Evidence from Britain and the USA*. Cambridge University Press, Cambridge, MA.
O'Connor J. and Gu X. (2013) Developing a creative cluster in a postindustrial city: CIDS and Manchester, in Flew T. (Ed.) *Creative Industries and Urban Development: Creative Cities in the 21st Century*, pp. 43–55. Routledge, Abingdon.
Okano H. and Samson D. (2010) Cultural urban branding and creative cities: a theoretical framework for promoting creativity in the public spaces, *Cities* **27**, S10–S15.
Paquette J. (2009) De l'enthousiasme à l'horizontalité: Sudbury, ville créative, *Cahiers de Géographie du Québec* **53**, 47–61.
Pasteur, L. (1854) Dans les champs de l'observation le hasard ne favorise que les esprits préparés. Lecture given at the University of Lille, 7 December.
Peck J. (2005) Struggling with the creative class, *International Journal of Urban and Regional Research* **29**, 740–779.
Philo C. and Kearns G. (Eds) (1993) *Selling Places: The City as Cultural Capital, Past and Present*. Pergamon, Oxford.
Piore M. and Sabel C. (1984) *The Second Industrial Divide: Possibilities for Prosperity*. Basic, New York, NY.
Ponzini D. and Rossi U. (2010) Becoming a creative city: the entrepreneurial mayor, network politics and the promise of an urban renaissance, *Urban Studies* **47**, 1037–1057.
Pratt A. C. (1997) The cultural industries production system: a case study of employment change in Britain, 1984–91, *Environment and Planning A* **29**, 1953–1974.
Pratt A. C. (2008) Creative cities: the cultural industries and the creative class, *Geografiska Annaler Series B – Human Geography* **90B**, 107–117.
Pratt A. C. (2011) The cultural contradictions of the creative city, *City, Culture and Society* **2**, 123–130.
Reich R. (1992) *The Work of Nations*. Vintage, New York, NY.
Sabel C. (1982) Italy's high-technology cottage industry, *Transatlantic Perspectives* **7(December)**.
Sabel C. and Zeitlin J. (1985) Historical alternatives to mass production: politics, markets and technology in nineteenth-century industrialization, *Past and Present* **108**, 133–176.
Sasaki M. (2010) Urban regeneration through cultural creativity and social inclusion: Rethinking creative city theory through a Japanese case study, *Cities* **27**, S3–S9.
Sayer A. (1989) Postfordism in question, *International Journal of Urban and Regional Research* **13**, 666–695.
Scott A. J. (1996) The craft, fashion, and cultural products industries of Los Angeles: competitive dynamics and policy dilemmas in a multi-sectoral image-producing complex, *Annals of the Association of American Geographers* **86**, 306–323.
Scott A. J. (1999) The cultural economy: geography and the creative field, *Culture, Media, and Society* **21**, 807–817.
Scott A. J. (2006) Entrepreneurship, innovation and industrial development: geography and the creative field revisited, *Small Business Economics* **26**, 1–24.
Scott A. J. (2008) *Social Economy of the Metropolis: Cognitive–Cultural Capitalism and the Global Resurgence of Cities*. Oxford University Press, Oxford.
Scott A. J. (2009) Human capital resources and requirements across the metropolitan hierarchy of the United States, *Journal of Economic Geography* **9**, 207–226.
Scott A. J. (2010a) Cultural economy and the creative field of the city, *Geografiska Annaler, Series B – Human Geography* **92**, 115–130.
Scott A. J. (2010b) The cultural economy of landscape and prospects for peripheral development in the twenty-first century: the case of the English Lake District, *European Planning Studies* **18**, 1567–1589.
Scott A. J. (2010c) Space–time variations of human capital assets in the American economy: profiles of abilities and skills across metropolitan areas, 1980 to 2000, *Economic Geography* **86**, 233–249.
Scott A. J. (2011) Emerging cities of the third wave, *City: Analysis of Urban Trends, Culture, Theory, Policy, Action* **15**, 289–321.
Scott A. J. (2012) *A World in Emergence: Cities and Regions in the 21st Century*. Edward Elgar, Cheltenham.
Seitz J. A. (2003) The political economy of creativity, *Creativity Research Journal* **15**, 385–392.
Sklair L. (2005) The transnational capitalist class and contemporary architecture in globalizing cities, *International Journal of Urban and Regional Research* **29**, 485–500.
Slater T. (2006) The eviction of critical perspectives from gentrification research, *International Journal of Urban and Regional Research* **30**, 737–757.

Smith N. (1982) Gentrification and uneven development, *Economic Geography* **58**, 139–155.

Smith N. (1986) Gentrification, the frontier, and the restructuring of urban space, in Smith N. and Williams P. (Eds) *Gentrification of the City*, pp. 15–34. Allen & Unwin, London.

Smith N. (2002) New globalism, new urbanism: gentrification as global urban strategy, *Antipode* **34**, 427–450.

Soja E. W. (2010) *Seeking Spatial Justice*. University of Minnesota Press, Minneapolis, MN.

Standing G. (2011) *The Precariat: The New Dangerous Class*. Bloomsbury, London.

Sternberg R. J. and Lubart T. I. (1999) The concept of creativity: prospects and paradigms, in Sternberg R. J. (Ed.) *Handbook of Creativity*, pp. 3–15. Cambridge University Press, Cambridge.

Stöhr W. B. (1986) Regional innovation complexes, *Papers of the Regional Science Association* **59**, 29–44.

Storper M. and Scott A. J. (2009) Rethinking human capital, creativity and urban growth, *Journal of Economic Geography* **9**, 147–167.

Vanolo A. (2008) The image of the creative city: some reflections on urban branding in Turin, *Cities* **25**, 370–382.

Vivant E. (2009) *Qu'est-ce que la Ville Créative?* Presses Universitaires de France, Paris.

Wacquant L. (2008) Relocating gentrification: the working class, science and the state in recent urban research, *International Journal of Urban and Regional Research* **32**, 198–205.

Waitt G. and Gibson C. (2009) Creative small cities, *Urban Studies* **46**, 1223–1246.

Wynne D. (1992) Cultural quarters, in Wynne D. (Ed.) *The Culture Industry*, pp. 13–23. Avebury, Aldershot.

Yencken D. (1988) The creative city, *Meanjin* **47**, 597–608.

Zimmerman J. (2008) From brew town to cool town: neoliberalism and the creative city development strategy in Milwaukee, *Cities* **25**, 230–242.

Zukin S. (1982) *Loft Living: Culture and Capital in Urban Change*. John Hopkins University Press, Baltimore, MD.

Capitalism and Urbanization in a New Key? The Cognitive-Cultural Dimension

Allen J. Scott, *University of California, Los Angeles*

The cognitive-cultural dimensions of contemporary capitalism are identified by reference to its leading sectors, basic technologies, labor relations systems and market structures. Cognitive-cultural systems of production and work come to ground preeminently in large city regions. This state of affairs is manifest in the diverse clusters of high-technology sectors, service functions, neo-artisanal manufacturing activities and cultural-products industries that are commonly found in these regions. It is also manifest in the formation of a broad stratum of high-skill, cognitive-cultural employees in urban areas. Many of these employees are engaged in distinctive forms of work-based learning, creativity and innovation. At the same time, the cognitive-cultural economy in contemporary cities is invariably complemented by large numbers of low-wage, low-skill jobs, and the individuals drawn into these jobs are often migrants from developing countries. The ideological-cum-political ramifications of this situation are subject to analysis in the context of a critique of the currently fashionable idea of the "creative city." I advance the claim that we need to go beyond advocacies about local economic development that prescribe the deployment of packages of selected amenities as a way of attracting elite workers into given urban areas. Instead, I propose that policy makers should pay more attention to the dynamics of the cognitive-cultural production system as such, and that in the interests of shaping viable urban communities in contemporary capitalism we must be more resolute in attempts to rebuild sociability, solidarity and democratic participation.

Capitalism and Urbanization

Wherever capitalism makes its historical and geographical appearance, peculiar patterns and rising levels of urbanization invariably ensue. This condition follows from pressures in capitalist economic systems that lead persistently to the formation of large aggregations of physical capital and human labor on the landscape. On the one hand, selected groups of profit-seeking firms, especially when they are meshed together in diverse functional interdependencies, have a definite tendency to locate near their common center of gravity. On the other hand, masses of individual workers are typically drawn to centers where employment opportunities are widely available. The developmental trajectory of any given urban node can then be described in terms of a spiral of interdependencies in which capital and labor continually exert an attractive force on one another in round after

The author is grateful to Michael Storper and two anonymous reviewers for their extremely helpful comments on an earlier draft of this paper. *Direct correspondence to Allen J. Scott, Department of Geography, University of California, Los Angeles, P.O. Box 951524, Los Angeles, CA, 90095-1524. Phone: (310) 925-7344. Fax: (310) 206-5976. E-mail: ajscott@ucla.edu.*

round of path-dependent cumulative causation, intensified by the emergence of localized external economies of scale and scope (Scott 2000). To be sure, these processes are intimately dependent on the expansion of final markets, and they are liable to reversal when – among other things – markets collapse.

As capitalism assumes different shapes and substantive content at different times and places, so the urban centers that it breeds reflect a corresponding series of specific substantive outcomes. The 19th century in Britain saw the rise of classical factory towns with their impoverished working-class populations. In 20th century North America, the large industrial metropolis flourished on the basis of Fordist mass production systems. The present moment of history is one in which a so-called post-Fordist (or better yet, a cognitive-cultural) capitalism has entered onto the stage in various parts of the globe, and along with it has come a new urban pattern, one that features a greatly widening gap between the upper and low tiers of the labor force. Much of productive activity today involves digital technologies and flexible organization sustaining the expansion of sectors that thrive on innovation, product diversity and the provision of personalized services. The notion of a cognitive-cultural economy refers above all to the circumstance that labor processes have come to depend more and more on intellectual and affective human assets (at both high and low levels of remuneration), and are increasingly less focused on bluntly routinized mental or manual forms of work. I shall argue that in the context of these developments, we can now pinpoint in some detail the features that mark the particular version of capitalism and urbanization that has been in gestation over the past two decades. One important observation is that the consolidation of the cognitive-cultural economy in many large cities today appears to be sparking new rounds of creative activity and response, not only in the production system, but also in the wider urban environment.

The capitalist system, of course, has always been characterized by a cognitive and cultural dimension, and this has always been a source of creative and innovative forces in cities (Hall 1998). However, the substance and magnitude of these forces typically exhibit a mediated connection to the specificities of the socio-economic order; that is, they are mobilized and assume tangible content by reference to tasks and opportunities that almost always bear a controlled relationship to concrete contemporary realities. I hasten to add, knowing that this last sentence will be viewed askance in certain quarters, that I am not proposing to reinstate some sort of hard-edged structuralism here, but neither am I willing to indulge in the vacuities of a purely voluntaristic conception of social life. Different imaginaries are possible in relation to any given social substrate and - this is the point - can be harnessed in the service of political action directed to social change (cf. Jessop 2004). Today, a very distinctive cognitive-cultural substrate is making great headway in the countries of advanced capitalism, and parallel to this development, a specific and pervasive set of social energies is also coming into play. One of the potent imaginaries that has appeared in the attempt to make sense of and to naturalize this emerging situation is articulated in the work of Florida (2002, 2004) in what he refers to as the "creative class" and the "creative city." In this article, I propose that an alternative way of approaching the issues raised by Florida can be forged on the basis of the more encompassing idea of cognitive-cultural capitalism and its manifestation in a unique and many-sided pattern of urban development.

The Cognitive-Cultural Dimensions of Production and Work in Contemporary Capitalism

Any concrete expression of capitalist economic order can typically be described in the first instance by reference to 1.) its leading sectors, 2.) its technological foundations, 3.) its characteristic forms of labor relations, and 4.) the competitive practices that it unleashes (cf. Boyer 1986). Each of these activity systems is manifest in unique ways in the emerging cognitive-cultural version of capitalism.

First, much of the contemporary economy is driven by key sectors such as technology-intensive manufacturing, services of all varieties (business, financial, personal), fashion-oriented neo-artisanal production, and cultural-products industries (including media). These sectors by no means account for the totality of the capitalist production system at the present time, but they are assuredly at the leading edges of growth and innovation in the most economically advanced countries. Second, and notwithstanding the evident heterogeneity of these sectors, they have all been deeply penetrated by digital technologies that have in turn facilitated the widespread deroutinization of labor processes and the destandardization of outputs. Third, employment relations have been subject to radical flexibilization and destabilization, thereby injecting high levels of precariousness into labor markets for workers at all levels of skill and human capital formation. Fourth, there has been a marked intensification of competition (reinforced by globalization) in all spheres of the economy, though much of this competition occurs in modified Chamberlinian form because products with high quotients of cognitive-cultural content often possess quasi-monopoly features that make them imperfect substitutes for one another and hence susceptible to niche marketing strategies.

As these trends have moved forward, the old white-collar/blue-collar principle of productive organization and labor-market stratification so characteristic of classical Fordism has also been deeply modified. On the one hand as Autor et al. (2003) and Levy and Murnane (2004) have argued, the advent of computerization has meant that many of the routine functions that were integral to the work of both the old white-collar fraction (e.g. accounting, records management, calculating, information sorting and so on) and the old blue-collar fraction (primarily repetitive manual operations) are rapidly being automated. On the other hand, this same trend has been associated with the formation of a new (core) labor-force elite whose work is concentrated primarily on high-level problem-solving tasks, and a new (peripheral) proletarian fraction that is increasingly called upon to function as a source of flexible labor in jobs such as machine operation (driving a vehicle), materials handling (small-batch assembly of variable components), security functions, cleaning and childcare. These jobs involve significant degrees of physical engagement and call for much less in the way of formal qualifications and training than jobs in the upper tier, but even they are imbued with varieties of meaningful cognitive-cultural content.

The upper tier of the labor force of the cognitive-cultural production system can be identified in terms of broad occupational categories such as managers, professional workers, business and financial analysts, scientific researchers, technicians, skilled craftsworkers, designers, artists. These are occupations that require significant levels of human capital, and they are generally well paid, though not invariably so (McRobbie 2004). To begin with, managerial and allied workers carry out the functions of administration, monitoring and control of the production system as a whole. Second, skilled analysts and other professionals are needed to maintain the specialized business and financial operations of modern capitalism. Third, scientific and technical workers are employed in large numbers to supervise the underlying technological infrastructure of the cognitive-cultural economy as well as to satisfy its unquenchable thirst for high levels of innovation. Fourth, many of the most dynamic sectors of the cognitive-cultural economy are characterized by a strong service element requiring human intermediation at the producer-consumer interface, and calling for skilled manipulation of affective-behavioral capabilities on the supply side. Fifth, workers with well-honed artistic and intellectual sensibilities make up an increasingly important part of the labor force because contemporary capitalism is also the site of a remarkable efflorescence of cultural-products industries in the broadest sense (i.e., industries with products that are permeated with some degree of aesthetic and semiotic content, and where such matters as fashion, meaning, look and feel significantly shape consumers' choices). In each of these types of employment, heavy doses of the human touch are required for the purposes of management, research, information gathering and synthesis, communication, inter-personal exchange, design, the infusion of sentiment, feeling and symbolic content into final products. The elite labor force that sustains these functions is expanding rapidly, especially in major metropolitan areas.

Alongside this upper tier of workers there exists a lower tier employed in a thick stratum of manual production activities that are not as well paid and much less gratifying in their psychic rewards. I am referring here both to the workshop and factory operations that underlie much of the cognitive-cultural economy today (as in many high-technology and neo-artisanal sectors), as well as to low-grade jobs in services such as janitorial and custodial work, facilities maintenance, unskilled hotel and restaurant trades. Additionally, a significant informal employment niche is sustained by the demands of more highly paid workers for domestic labor to perform tasks such as house cleaning, repair work, gardening and childcare. This extended underbelly of the cognitive-cultural economy is notorious for its sweatshop operations and frequent brushes with illegality in regard to labor laws. In the more advanced countries, a high proportion of the labor force in this segment is made up of immigrants (many of them undocumented) from developing parts of the world. Large numbers of these immigrants form a polyglot underclass with a marginal social and political presence in their host environments.

The gap between the average incomes of these two strata of the workforce identified in the previous paragraphs has been growing apace in the United

States over the last decade or so (Autor et al. 2006; Morris and Western 1999; Yun 2006). Both, too, are subject to much labor-market instability. Workers of all types face increasingly frequent bouts of unemployment, and are more and more likely to be caught up in temporary, part-time, and freelance modes of labor. Along with these shifts in the structure of the employment relation has gone what some analysts identify as a declining sense of allegiance among workers to any single employer (Beck 2000). To be sure, the capacities of each of these groups for dealing with these predicaments differ dramatically. While social networks are a major source of labor market information for both groups, individuals in the upper stratum usually command resources in terms of contacts and interpersonal know-how that allow them a far greater range of maneuver. In contemporary society, it is not uncommon to come across cognitive-cultural workers who have carried networking to a something like a fine art, or more accurately, perhaps, a semi-routinized habit of life in which they devote considerable amounts of time to socializing with fellow workers and exchanging information with one another about job opportunities and the state of the labor market. Reputation is a key item of currency in these fluid employment conditions, and is a major factor lubricating the progress of upper-stratum workers through the employment system. An essential strategy deployed by many individuals in this stratum involves the accumulation of personal portfolios of employment experiences demonstrating the depth and diversity of their career paths and creative accomplishments hitherto (Neff et al. 2005). For these workers, too, elaborate self-management of careers replaces the bureaucratized personnel functions of the traditional corporation.

The Cognitive-Cultural Economy and the Metropolis

As this new economic order grew over the past couple of decades, it found fertile ground in large metropolitan areas such as New York, Los Angeles, London, Paris, Amsterdam and Tokyo (cf. Sassen 1994). These are the flagship hubs of the new economy, and the primary nerve centers of a cognitive-cultural production system increasingly geared to markets that extend across the entire globe.

Cognitive-cultural production activities, then, are typically concentrated in dense locational clusters, yet their market reach frequently extends to the far corners of the world. Two analytical lines of attack help to clarify this apparently paradoxical state of affairs. In the first place, producers in cognitive-cultural sectors of the economy have a definite proclivity to agglomerate together in geographic space by reason of the external economies of scale and scope (or increasing-returns effects) that flow from selected aspects of their joint operation in particular localities. The role of flexible inter-firm networks, local labor markets and localized learning processes is especially critical here (Cooke and Morgan 1998; Scott 2000; Storper 1997). Groups of producers with strong interdependencies in regard to these variables have a powerful inducement to gravitate toward their common center gravity, thereby reducing the space-time costs of their traded and untraded transactional relations and enhancing the total stock of jointly-generated external economies. Even though it is true that low (and ever falling) transactions costs make it possible for certain kinds of firms to dispense with the advantages of agglomeration and to decentralize to low-cost locations, the same phenomenon also permits many other kinds of producers to enjoy the best of both worlds (to remain anchored within a specific cluster and to continue to appropriate localized competitive advantages while simultaneously contesting global markets). As the market range of producers in any given cluster increases, moreover, local economic growth accelerates, leading to the deepening of localized increasing-returns effects and the intensification of agglomeration. The signs of this developmental dynamic are palpable in the world's great metropolitan areas today, both in the rapidly growing incidence of cognitive-cultural sectors and in the frequent expression of this growth in the formation of intra-urban industrial districts devoted to specialized facets of cognitive-cultural production (Arai et al. 2004; Currid 2006; Pratt 1997; Rantisi 2004; Schoales 2006). Classical examples of such developments are high-technology and software production in the San Francisco Bay Area, the entertainment industry in and around Hollywood, the business and financial centers of New York and London, and the fashion worlds of Paris and Milan.

Along with the widespread growth of cognitive-cultural production systems in the modern city have come numerous parallel transformations of intra-urban space, including significant enhancements of the form and function of privileged parts of the urban fabric. Among the most symptomatic expressions of this trend is a general process of social and economic upgrading in downtown areas and surrounding inner city areas. This process is widely referred to in the literature as "gentrification," (Smith 2002; Zukin 1982) though the concept originally referred to incursions of middle-class households into decaying inner city neighborhoods. What is at stake in this regard nowadays is nothing less than radical transformations of extensive tracts of urban space by a four-fold logic of cognitive-cultural economic development, social transformation, attendant functional changes and the re-imaging of the environment by means of dramatic new symbologies.

An increasingly common manifestation of this process is the recycling and upgrading of old industrial and commercial zones of the city to provide new spaces able to accommodate high-level production and consumption activities. Harbor Front in Baltimore, Docklands in London and the Zürich West development are outstanding examples of this phenomenon. Similar kinds of initiatives can be found in Britain in Manchester's Northern Quarter and Sheffield's Cultural Industries Quarter with their aspirations to develop as dynamic hubs for small creative enterprises such as recording companies, electronic media labs, fashion design studios, and so on. In Los Angeles, a new Fashion District just to the south of the central business district has recently been created in what was originally a dispiriting cluster of grimy clothing factories. This development, with its renovated buildings and colorful street scenes, expresses the rising status of the Los Angeles clothing industry as a global center of designer fashions, and helps to sustain the new-found ambitions of many local producers to compete in high-end markets (Scott 2002). In similar initiatives, local authorities in cities all over the world are engaged in projects that involve the conversion of derelict facilities to serve a diversity of cultural purposes, as in the case of Amsterdam's Westergasfabriek or parts of the Ruhr region of Germany where efforts to rebuild decaying industrial landscapes are aggressively under way.

A related and increasingly spectacular case of the recycling of urban space can be observed in the construction of large-scale architectural set pieces, functioning as iconic expressions of local economic and cultural aspirations in an age of cognitive-cultural capitalism. The grand projects set on foot by President François Mitterand in Paris in the 1980s represent one of the pioneering and certainly one of the most determined examples of this kind of ambition, and have done much to add to the already celebrated reputation of Paris as the city of spectacle and a global cultural reference point. Other illustrative cases of urban re-imaging projects in pursuit of economic and cultural status are the Guggenheim Museum in Bilbao, Toronto's Harbourfront and the Petronas Towers in Kuala Lumpur. These projects register a presence on the global stage while generating prestige and cachet that spill over into the wider urban communities in which they are located. Urban elites in all parts of the world are increasingly committed to the pursuit of projects like these in attempts to assert the visibility of their cities as foci of cultural interest and economic promise in the new global order.

As these changes have occurred, large swaths of low-income neighborhoods in central city areas have been subjected to appropriation and recolonization by the affluent. This process is expressed both in the renovation of old working-class residential properties and derelict slums, and in wholesale land clearances to accommodate new blocks of expensive condominiums. Gentrification in this sense has been going on in American cities for decades, but it has accelerated greatly in recent years as a result of changing structural conditions in the urban environment and changing priorities in residential preferences. In particular, as jobs in traditional manufacturing and wholesaling activities declined in inner urban areas, much of the old working-class population in adjacent neighborhoods migrated to other parts of the city. Correspondingly, job opportunities for cognitive-cultural workers in and around the central business districts of large cities have mushroomed of late years, and many of these workers are assuming residence in nearby neighborhoods to reduce commuting times and gain access to shopping, leisure and cultural facilities in the city. Very often, the first sign that a dilapidated section of the inner city is destined to go through this sort of transition is the irruption of groups of artists and bohemians in the area and the blossoming of studios, cafés, clubs and so on, serving their needs (Zukin 1982). Indeed, some analysts have accorded these groups, along with gays, a special status as key harbingers and tracking molecules of the "creative city" syndrome (Florida 2004; Lloyd 2002; Lloyd and Clark 2001). The overt presence in the urban landscape of such groups is said to symptomatize a state of openness and tolerance in local society, qualities that are thought, in turn, to be essential for the blooming of a creative environment. As such, the presence or absence of these groups in the city is taken by some commentators to represent a sort of litmus test of local prospects for general "creativity."

There are numerous signs, then, of important shifts in the functions and form of the city as the cognitive-cultural foundations of modern capitalism have deepened and widened. These shifts are detectable in the economic patterns, social organization and physical structure of many different cities. Specialized areas of the city dedicated to entertainment, recreation, edification and shopping have also undergone much elaboration and embellishment as individuals with high levels of cognitive and cultural capital – not to mention pecuniary capital – have become a more insistent component of contemporary urban life (Zukin 1995). In these ways, a new kind of balance and integration seems to be emerging at least in privileged sections of modern cities between economy and society, between production and consumption, between work and leisure, and between commerce and culture. A dark shadow is nonetheless cast over this gratifying picture both by the swelling underbelly of low-wage industrial and service functions that are invariably to be found in large metropolitan areas where cognitive-cultural economic functions are most highly developed, and by the often problem-ridden residential areas that are the sources of the labor needed to maintain these functions.

The deepening pall cast by this condition of social and economic inequality almost certainly puts shackles on the potential of the city for creative performance and on its capacity to promote consistently high levels of social learning, economic innovation and human conviviality. Large segments of the urban population face serious impediments to participation as full-blown citizens in daily life and work, a circumstance that generates high costs to the individuals directly concerned and – via the multiple negative externalities that result from this situation – to urban society as a whole. The problems of a divided and unequal citizenry are compounded by the fact that many of the most underprivileged groups in large metropolitan areas today consist of immigrants from poor countries drawn into the orbit of the urban economy by the low-wage employment opportunities that proliferate in these areas. In many cases, these immigrants form polyethnic and polylingual neighborhoods within the social space of the metropolis, thereby exacerbating the social separation and isolation that constantly work against the formation of a wider sense of community. The relentless withdrawal of public services that is occurring in the context of the neoliberal political climate prevailing in many of the more advanced capitalist countries at the present time only serves to intensify the possessive individualism characteristic of so much of modern urban life at the expense of communal values. Equally, as globalization runs its course, extended geographic echoes of these same predicaments become ever louder. On the one hand, new expressions of formal organic solidarity via the division of labor are taking shape on a global scale as increasingly large volumes of low-wage work are transferred from the more economically advanced countries to diverse locations in the developing world. On the other hand, this trend is proceeding without the concomitant inconveniences of propinquity to more privileged social strata in the developed world, so responsibility or accountability by individuals in the upper strata in regard to individuals in lower strata is apt to be further diminished.

Cognitive-Cultural Workers and the Constitution of Urban Life

Over the past few decades, many social scientists have attempted to describe the changing stratification of capitalist society and to typify the shifts that have been occurring in social structure since the heyday of the classical white-collar/blue-collar division that prevailed in American cities over much of the 20th century. In a pioneering statement, Bell (1973) alluded to the advent of what he called post-industrial society, and he suggested that the old social divisions of capitalism were in fact being transcended by a new-found drive for personal fulfillment and self-realization in a service-oriented economy. Gouldner (1979) offers us the idea of a "new class" made up of individuals who have internalized an ideology of critical rationality; for them, reasoned arguments take precedence over hierarchical authority as a basis for belief and action. The modern technocrat is the emblematic figure of this new class. Reich (1992), in turn, refers to "symbolic workers" who constitute, he claims, the elite of an emerging information society. Sklair (2000) broadens the picture with the concept of a "transnational capitalist class" composed of managers, professionals and technicians who are engaged in forms of work that express and promote the historical project of globalization. Most recently of all, Florida (2002) has advanced the argument that a new "creative class," comprising all those workers engaged in one form or another of thought-intensive work, has come into being in American society.

Each of these attempts to say something about the changing organization of society in contemporary capitalism unquestionably has something of interest and significance to convey, though none is entirely satisfactory. The term "class," is perhaps unduly forceful a word to use for some of these rather nebulous social groupings, especially in view of its more orthodox connotation of two opposing strata whose interests clash as a consequence of their structured relations to the means of production and their opposing claims on the economic surplus. Additionally, as Markusen (2006) has argued, Florida's proposed creative class is something of an incoherent concept, for it assembles a wide assortment of very disparately situated individuals – from company executives to software programmers and from international financiers to artists – within its rather elastic boundaries. This assortment does not even look much like the relatively diluted Weberian idea of class with its emphasis on occupation and relative life chances. Still more problematical is the way in which Florida invests those individuals who compose the more privileged segments of capitalist society with a sort of ontological capacity for "creativity," a characterization that carries with it an overload of exhilarating implications, but that is also rather threadbare in terms of its concrete meaning. In reality, the distinctive forms of human capital that these individuals possess – specifically the *cognitive and cultural* tasks they are called on to perform in their work – are, for the most part, wedged in social grooves and infused with very specific substance. Within the framework of contemporary capitalism, these tasks are focused on activities including neoliberal techno-management, innovation-oriented process and product design, the personalized provision of services, the naturalization of socially-useful aptitudes and beliefs (in educational institutions and the media, for example), and the commercialization of experiences, cultural encounters and leisure pursuits. Special mention needs to be made in this context of the enormous recent expansion of cultural-products industries generally and the concomitant emergence of an important segment of the cognitive-cultural labor force dedicated to the conception and fabrication of outputs whose function is to entertain, to instruct, to embellish and to reinforce identity (Bourdieu 1979; Hesmondhalgh 2002; Power and Scott 2004). This is a world, as Lash and Urry (1994) have shown, in which culture is produced increasingly in commodity form, while commodity production itself becomes ever more deeply infused with aesthetic and semiotic meaning. The steady convergence of the economic and the cultural in contemporary capitalism has led some postmodern theorists to claim – no doubt correctly – that the sphere of culture today is endemically subject to a condition of waning symbolic intensity and rising ephemerality (cf. Jameson 1992).

Various intimations of the logic and meaning of the new social forces and alignments that are rising to the fore in capitalist society are now a common feature of journalistic accounts of current economic and urban realities. Among the more prominent of these effusions on the new economy is a stream of managerial theories and advice directed to the personal and affective qualities required to bring order and dynamism into the cognitive-cultural workplace. The normative discourse of management analysts and consultants today is considerably less concerned than it once was with down-to-earth issues of efficiency and control, and much more focused on methods of cultivating human resources including leadership, empathy, self-motivation, adaptability, inventiveness, resourcefulness and ethical consciousness in a fast-moving, high-risk business environment (Boltanski and Chiapello 1999; Thrift 2005). There is incontestably much in this discourse that is helpful to managers and workers trying to find some sort of strategic purchase on the day-to-day problems they face in the new cognitive-cultural economic environment, though it is distinctly less useful as a guide to the formulation of critical insights or as a basis for the construction of sensible and politically plausible imaginaries about alternative possibilities. Various echoes of this discourse resonate in the currently fashionable creed of the creative city, with its upbeat message about the transformation of urban areas by means of programs designed to draw in members of the "creative class" who will then, presumably, express their talents and energies in ways that result in multiple local economic and cultural benefits (Florida 2004; Landry 2000). Once more, we can find elements in the analysis that merit our attention, even if in its raw form it greatly oversimplifies the policy challenges that need to be addressed in any economic development program (nowhere more so than in regard to the construction of employment opportunities), and obscures the historically-specific function and meaning of intellectual and symbolic labor in contemporary capitalism. At the same time, it is perhaps worth reflecting that this same creed tends to display an exaggeratedly optimistic faith in the benign social and political impacts of the so-called creative class, if only by reason of its pregnant silences about the deepening social divide in the cities of advanced capitalism today and its signal failure to call into question any of the more regressive aspects of the contemporary cultural scene wrought by this fraction of the labor force. That said, in Florida's more recent work, 2005, he makes a start on rectifying some of these lacunae by acknowledging the links between the new economy and economic inequality.

The less prepossessing features of the cognitive-cultural economy are amplified by the added problem of rapidly rising levels of social instability and risk, so that all strata – even the urban elite – are subject to an intensification of the general precariousness of life (Beck 2000). Individual members of the labor force exert considerable energy and time in navigating pathways through the reefs and shoals of practical social existence whether by means of very self-conscious social networking on the part of upper-tier workers (Batt et al. 2001; Neff et al. 2005; Ursell 2000) or via diverse ethnic and extended-family ties on the part of the lower tier (Sanders et al. 2002; Waldinger 2001). Many kinds of cognitive-cultural workers – especially in the early stages of their careers – are inveterate joiners of work-related social groups, and they are prone to spend large amounts of time outside their normal working hours in building relationships with allied workers so as to maintain their labor market edge (cf. Scott 2000). In these conditions, human interaction is apt to take on discernible utilitarian undertones. Thus, in a study of workers in the television industry Ursell (2000) has shown how an "economy of favors" has arisen in which information about job opportunities and work-related matters is exchanged on an informal *quid pro quo* basis through extended webs of social contacts. At the same time, the kaleidoscope of shifting opportunities and setbacks that characterize much of the cognitive-cultural economy today is increasingly reflected in careers that unfold across many different employers in many different places, and often -- especially for upper-tier workers -- in many different countries. In this manner, the traditional connection between propinquity and community is subject to further decay, just as a growing ethos of interpersonal engagement without durable commitment becomes a normalized condition of urban existence. The same instability and insecurity provide a strong incentive for members of the upper-tier of the labor force to engage in persistent self-promotion and self-publicity, an incentive that no doubt is magnified the more they are possessed of a portfolio of experiences and qualifications that mark them out as the bearers of a unique package of attributes and talents. In testimony to the above remarks, Sennett (1998) has pointed to an apparent corrosion of traditional forms of affectivity and trust in both the workplace and social life, while Putnam (2000) has written more generally about the weakening of communal ties in America.

It is tempting to attribute at least some of the narcissism that was thought by Lasch (1978) to be on the rise in the American psyche to social forces and predicaments of these types. A less ambitious way of making much the same point is to appeal to the accumulating evidence of the expansion of the sphere of the private and the personal and a corresponding contraction of the public sphere in American cities. Quite apart from the condition of public penury and a broadly decaying sense of community, as already invoked, we can see the immediate effects of this state of affairs in the intense fragmentation of the social space of the contemporary metropolis. The very social diversity that is so often celebrated as one of the main conditions of a creative urban environment today is actually inscribed on the landscape of the metropolis in patterns of separation and detachment, accentuated by the striking marginalization of the ever-expanding immigrant population of the city. For many immigrants, this situation is manifest in relative and absolute poverty as well as in political disenfranchisement. The fact that so many of these denizens of American cities in the early 21st century have curtailed entitlements and restricted channels for the democratic expression of their political aspirations means not only that they are denied full incorporation into urban society, but also that they have limited incentives to make durable commitments to the community at large. The net result is further deterioration of the capacity of the urban system for releasing and mobilizing the creative potential of the citizenry. Perhaps one of the most symptomatic expressions of the inhospitable character of the city of social extremes as found in contemporary America is the proliferation of gated neighborhoods with their transformation of important sections of urban space into zones of explicit exclusion (Blakely and Snyder 1999). This phenomenon represents a direct incursion on the democratic use of urban space and an actual and symbolic violation of the principle of common citizenship.

Beyond the Creative City

As cognitive-cultural forms of production and work penetrate more deeply into contemporary capitalist society, enormously varied bundles of urban responses have been set in motion. A set of privileged intra-metropolitan spaces supporting the work, residence and leisure activities of the new cognitive-cultural elite is now an important ingredient of many world cities. On the other side, and given that large numbers of low-wage, low-skill jobs are a major element of the cognitive-cultural economy, a growing underclass is also a major feature of the very same cities. These trends are embedded in a widening dynamic of economic-*cum*-cultural integration on the global scale, leading to complex forms of urban specialization and interdependence across the global landscape.

Some of the more positive features of this picture have been highlighted in a number of normative commentaries focused on the creative potentials of contemporary cities. Policy makers and planners in many different parts of the world have understandably displayed much enthusiasm in regard to these commentaries, and in numerous cases have actually embarked on attempts to make their cities appealing to the talented and high-skill individuals who are thought, in the more prominent versions of the story, to be the *primum mobile* of the creative city. The idea of the creative city is all the more irresistible to policy makers in view of its promise of high-wage jobs in sectors of economic activity that are mostly environmentally friendly and promise to upgrade the urban fabric. In a number of cases, practical attempts to pursue the idea have been complemented by efforts to mount displays of architectural master strokes designed to attract the attention of potential visitors and inward investors and to establish dramatized points of reference in the global race for economic and cultural influence. Florida (2002, 2004) has been the most forthright instigator of a normative agenda like this, but his ideas find both implicit and explicit support in other work, including the "consumer city" concept as formulated by Glaeser et al. (2001), and the view of the city as an "entertainment machine" that Lloyd and Clark (2001) have proposed. Florida's suggested strategy for building the creative city can be schematized – with only a touch of willful skepticism – in terms of three main brush strokes. First, municipal authorities are advised to encourage the development of amenities that are claimed to be valued by the creative class. Bikeways and fashionable restaurants figure prominently in the suggestions offered here (and regression analysis suggests that warm winters also help things along[1]); movie theaters and art galleries are apparently of much less consequence. Second, Florida proposes that once appropriate packages of amenities are in place in any given city, members of the creative class will then be induced to take up residence, especially if an atmosphere of tolerance and openness also prevails. As this occurs, diverse creative energies will then presumably be released. Third, and consequently, the dynamism of the local economy can be expected to accelerate along with further upscaling of the built environment and general enhancement of the prestige-*cum*-attractiveness of the city as a whole.

I have criticized this approach elsewhere (Scott 2004, 2006), and reaffirm that cities are subject to path-dependent growth trajectories in which both the supply and the demand for labor move in patterns of mutually cumulative causation. The primary engine of this process is not the inward and unilateral migration of particular types of workers, but the complex apparatus of the urban production system (i.e., the network of interrelated industrial and service activities generating locationally polarized economic development). This type of developmental engine was obviously at work in earlier periods of capitalism, and it is still detectable as the major motive force of urbanization in cognitive-cultural capitalism today. Consider the case of factory towns in 19th century England. It was not the prior massing together of dense working-class populations that explains the formation of these towns, even though the presence of a working-class population is essential for a factory town to function. Equally, the growth of Silicon Valley in the second half of the 20th century is not to be accounted for by invoking the prior existence of some undifferentiated creative class in the local area, just as it would surely be absurd to claim that the driving force of the Valley's long-term expansion can be ascribed to continual incursions by members of that class in search of amenity value. On the contrary, the historic accumulation in Silicon Valley of a labor force comprising specialized semiconductor technicians, computer scientists, software and engineers is comprehensible only when we set this trend in the context of an evolving web of specialized production activities and employment opportunities tied in to ever widening final markets for semiconductors, computers and software. Yes, the supply of labor is a crucial moment in the chain of temporal intermediations through which cognitive-cultural centers of production and work evolve, but it remains a subordinate moment in the sense that the generative power of local economic development resides preeminently in the path-dependent logic of production, agglomeration, and regional specialization. By the same token, dissipation of that power is a virtually inevitable road to ruin even where large numbers of workers with high levels of human capital continue to reside in the local area. Policy makers neglect these aspects of the problem at their peril.

Beyond the analytical flaws that underlie much recent work on the creative city, an odd reticence can be detected in many of the claims advanced about the possibilities for revival of the social life and physical environment of cities by tapping into the expansionary powers of the cognitive-cultural economy. While cognitive-cultural forms of production and work offer new and dynamic possibilities for urban regeneration, it bears repeating that there is a dark side to the developmental dialectic of contemporary cities, and the deepening neoliberalism trend is exacerbating the problem. This premise raises issues about the reconstruction of urban society that go well beyond simple pleas for openness, tolerance and diversity; while these are excellent goals, they do not guarantee transcendence of social isolation, fragmentation and inequality. To the contrary, even if these qualities were universally present, the ingrained structural logic of contemporary economic and social order would still give rise to conspicuous inequities and injustices in large cities. In contrast with the neoliberal political agenda that currently holds sway in the United States, and that is endemically associated with high levels of urban poverty and deprivation, only some sort of conscientious program of social democracy with a strong focus on redistribution, decent jobs for all and the re-engagement of the citizenry in the political realm seems appropriate to address social reform. Beyond the implementation of elementary principles of social equity, justice and participatory democracy, an additional challenge looms. As cities shift into cognitive-cultural modes of economic activity, the search for meaningful forms of solidarity, sociability and mutual aid in everyday work and life becomes increasingly urgent – not just because these attributes are important in their own right – but also because they enlarge the sphere of creativity, learning, innovation, social experimentation and cultural expression and are essential for the further economic and cultural flowering of contemporary cities. It is just possible that some of the goals of this search may be realizable if, as Judis and Teixeira (2002) expect, a new and socially progressive majority begins to take shape in what they call "postindustrial" cities.

Finally, an even broader social imperative is brought to the fore as the cognitive-cultural economy continues its ascent and as the symbolic-affective content of final outputs becomes ever more pervasive. Consumption of these outputs has potent direct and indirect impacts on human consciousness and ideological orientation, and this process, by the same token, generates massive externalities for all. These externalities give rise to complex dilemmas for they reappear in various social and political guises with deep implications for modes of social being. And precisely because they are externalities, they can never be adequately processed via market rationality alone. A persistent public debate and mutual education about the personal meanings and political consequences of the consumption side of the cognitive-cultural economy – and about the possibilities of more critically informed participation – is a further prerequisite of a progressive and democratic social order in contemporary capitalism.

Note

1. There is much in this kind of analytical maneuver that recalls the obdurate tautologies of neoclassical economics. For example, if we observe a significant tendency for individuals of type x to live in proximity to attributes of type y it is said that the same individuals must have a "revealed preference" for y. Revealed preference then accounts for their presence in proximity to y. How do we know this? Because they live close to y!

References

Arai, Y., H. Nakamura and H. Sato. 2004. "Multimedia and Internet Business Clusters in Central Tokyo." *Urban Geography* 25:483-500.

Autor, D.H., L.F. Katz and M.S. Kearney. 2006. "The Polarization of the US Labor Market." *American Economic Review* 96:189-94.

Autor, D.H., F. Levy and R.J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118:1279-1333.

Batt, R., S. Christopherson, N. Rightor and D. Van Jaarsveld. 2001. *Net Working: Work Patterns and Workforce Policies for the New Media Industry*. Washington, DC: Economic Policy Institute.

Beck, U. 2000. *The Brave New World of Work*. Polity Press.

Bell, D. 1973. *The Coming of Post-Industrial Society; a Venture in Social Forecasting*. Basic Books.

Blakely, E.J., and M.G. Snyder. 1999. *Fortress America: Gated Communities in the United States*. Brookings Institution Press.

Boltanski, L., and E. Chiapello. 1999. *Le Nouvel Esprit du Capitalisme*. Gallimard.

Bourdieu, P. 1979. *La Distinction: Critique Sociale du Jugement*. Le Sens Commun.

Boyer, R. 1986. *La Théorie de la Régulation: Une Analyse Critique*. Algalma.

Cooke, P., and K. Morgan. 1998. *The Associational Economy: Firms, Regions, and Innovation*. Oxford University Press.

Currid, E. 2006. "New York as a Global Creative Hub: A Competitive Analysis of Four Theories on World Cities." *Economic Development Quarterly* 20:330-50.

Florida, R. 2002. *The Rise of the Creative Class*. Basic Books.

______. 2004. *Cities and the Creative Class*. Routledge.

______. 2005. *The Flight of the Creative Class: The New Global Competition for Talent*. Harper Business.

Glaeser, E.L., J. Kolko and A. Saiz. 2001. "Consumer City." *Journal of Economic Geography* 1:27-50.

Gouldner, A. 1979. *The Future of Intellectuals and the Rise of the New Class*. Seabury.

Hall, P. 1998. *Cities in Civilization*. Pantheon.

Hesmondhalgh, D. 2002. *The Cultural Industries*. Sage.

Jameson, F. 1992. *Postmodernism, or, the Cultural Logic of Late Capitalism*. Duke University Press.

Jessop, B. 2004. "Critical Semiotic Analysis and Cultural Political Economy." *Critical Discourse Studies* 1:1-16.

Judis, J.B., and R. Teixeira. 2002. *The Emerging Democratic Majority*. Scribner.

Landry, C. 2000. *The Creative City: A Toolkit for Urban Innovators*. Earthscan.

Lasch, C. 1978. *The Culture of Narcissism: American Life in an Age of Diminishing Expectations*. Norton.

Lash, S., and J. Urry. 1994. *Economies of Signs and Space*. Sage.

Levy, F., and R.J. Murnane. 2004. *The New Division of Labor: How Computers are Creating the Next Job Market*. Russell Sage Foundation.

Lloyd, R. 2002. "Neo-Bohemia: Art and Neighborhood Development in Chicago." *Journal of Urban Affairs* 24:517-32.

Lloyd, R., and T.N. Clark. 2001. "The City as an Entertainment Machine." Pp. 357-78. *Critical Perspectives on Urban Redevelopment*. K.F. Gotham, editor. Elsevier.

Markusen, A. 2006. "Urban Development and the Politics of a Creative Class: Evidence from a Study of Artists." *Environment and Planning A* 10:1921-40.

McRobbie, A. 2004. "Making a Living in London's Small-Scale Creative Sector." Pp. 130-43 . *Cultural Industries and the Production of Culture*. D. Power and A.J. Scott, editors. Routledge.

Morris, M., and B. Western. 1999. "Inequality in Earnings at the Close of the Twentieth Century." *Annual Review of Sociology* 25:623-57.

Neff, G., E. Wissinger and S. Zukin. 2005. "Entrepreneurial Labor among Cultural Producers: Cool Jobs in Hot Industries." *Social Semiotics* 15:307-34.

Power, D., and A.J. Scott. Editors. 2004. *Cultural Industries and the Production of Culture*. Routledge.

Pratt, A.C. 1997. "The Cultural Industries Production System: a Case Study of Employment Change in Britain, 1984-91." *Environment and Planning A* 29:1953-74.

Putnam, R. 2000. *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster.

Rantisi, N. 2004. "The Designer in the City and the City in the Designer." Pp. 91-109. *Cultural Industries and the Production of Culture*. D. Power and A.J. Scott, editors. Routledge.

Reich, R. 1992. *The Work of Nations*. Vintage.

Sanders, J., V. Nee and S. Sernau. 2002. "Asian Immigrants' Reliance on Social Ties in a Multiethnic Labor Market." *Social Forces* 81:281-314.

Sassen, S. 1994. *Cities in a World Economy*. Pine Forge Press.

Schoales, J. 2006. "Alpha Clusters: Creative Innovation in Local Economies." *Economic Development Quarterly* 20:162-77.

Scott, A.J. 2000. *The Cultural Economy of Cities: Essays on the Geography of Image-Producing Industries*. Sage.

______. 2002. "Competitive Dynamics of Southern California's Clothing Industry: the Widening Global Connection and its Local Ramifications." *Urban Studies* 39:1287-1306.

______. 2004. "Cultural-Products Industries and Urban Economic Development: Prospects for Growth and Market Contestation in Global Context." *Urban Affairs Review* 39:461-90.

______. 2006. "Creative Cities: Conceptual Issues and Policy Problems." *Journal of Urban Affairs* 28:1-17.

Sennett, R. 1998. *The Corrosion of Character: the Personal Consequences of Work in the New Capitalism*. W.W. Norton.

Sklair, L. 2000. *The Transnational Capitalist Class*. Blackwell.

Smith, N. 2002. "New Globalism, New Urbanism: Gentrification as Global Urban Strategy." *Antipode* 34:427-50.

Storper, M. 1997. *The Regional World: Territorial Development in a Global Economy*. Guilford Press.

Thrift, N. 2005. *Knowing Capitalism*. Sage.

Ursell, G. 2000. "Television Production: Issues of Exploitation, Commodification and Subjectivity in UK Television Markets." *Media, Culture and Society* 22:805-25.

Waldinger, R. 2001. "The Immigrant Niche in Global City-Regions: Concept, Patterns, Controversy." Pp. 299-322. *Global City-Regions: Trends, Theory, Policy*. A.J. Scott, editor. Oxford University Press.

Yun, M.S. 2006. "Earnings Inequality in USA, 1969-99: Comparing Inequality Using Earnings Equations." *Review of Income and Wealth* 1:127-44.

Zukin, S. 1982. *Loft Living: Culture and Capital in Urban Change*. John Hopkins University Press.

______. 1995. *The Cultures of Cities*. Blackwell.

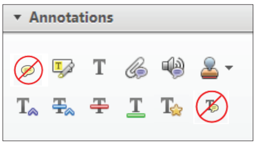
Page Proof Instructions and Queries

Journal Title: USJ
Article Number: 634002

Greetings, and thank you for publishing with SAGE. We have prepared this page proof for your review. Please respond to each of the below queries by digitally marking this PDF using Adobe Reader.

Click "Comment" in the upper right corner of Adobe Reader to access the mark-up tools as follows:

For textual edits, please use the "Annotations" tools. Please refrain from using the two tools crossed out below, as data loss can occur when using these tools.



For formatting requests, questions, or other complicated changes, please insert a comment using "Drawing Markups."

Detailed annotation guidelines can be viewed at: http://www.sagepub.com/repository/binaries/pdfs/AnnotationGuidelines.pdf Adobe Reader can be downloaded (free) at: http://www.adobe.com/products/reader.html.

Sl. No.	Query
	GQ: Please confirm that all author information, including names, affiliations, sequence, and contact details, is correct.
	GQ: Please review the entire document for typographical errors, mathematical errors, and any other necessary corrections; check headings, tables, and figures.
	GQ: Please ensure that you have obtained and enclosed all necessary permissions for the reproduction of art works (e.g. illustrations, photographs, charts, maps, other visual material, etc.) not owned by yourself. please refer to your publishing agreement for further information.
	GQ: Please note that this proof represents your final opportunity to review your article prior to publication, so please do send all of your changes now.
	GQ: Please confirm that the funding statements are accurate.
1	AQ: Glaeser is listed as 2011 in the reference section but as 2012 in the text – which is correct?
2	AQ: McKinsey, 2011 is not listed in the reference section. Please give details
3	AQ: Geertz (1973) is not listed in the reference section. Please give details
4	AQ: Please check whether these three references are to McFarlane 2011a, 2011b or 2011c
5	AQ: Is this McFarlane 2011a, 2011b or 2011c?

6	AQ: Caldeira (2001) is not listed in the reference section. Please give details
7	AQ: Is this McFarlane 2011a, 2011b or 2011c?
8	AQ: Roy A (2014) is not cited in the text. Please specify where it should be cited or if it can be deleted
9	AQ: Please give page range of chapter for Sheppard E (2014)
10	AQ: Please give volume, issue and page numbers, or DOI, for Walker R (2016)

Article

Urban Studies

Urban Studies
1–23

DOI: 10.1177/0042098016634002
usj.sagepub.com
SAGE

Current debates in urban theory: A critical assessment

Michael Storper
Department of Geography and Environment, London School of Economics, UK

Allen Scott
Department of Geography, University of California at Los Angeles, USA

Abstract

Urban studies today is marked by many active debates. In an earlier paper, we addressed some of these debates by proposing a foundational concept of urbanisation and urban form as a way of identifying a common language for urban research. In the present paper we provide a brief recapitulation of that framework. We then use this preliminary material as background to a critique of three currently influential versions of urban analysis, namely, postcolonial urban theory, assemblage theoretic approaches and planetary urbanism. We evaluate each of these versions in turn and find them seriously wanting as statements about urban realities. We criticize (a) postcolonial urban theory for its particularism and its insistence on the provincialisation of knowledge, (b) assemblage theoretic approaches for their indeterminacy and eclecticism and (c) planetary urbanism for its radical devaluation of the forces of agglomeration and nodality in urban-economic geography.

Keywords

agglomeration theory, assemblage theory, planetary urbanisation theory, post-colonial urbanism, urban theory

Received November 2015; accepted January 2016

Urban challenges and urban theory in the 21st century

The current period of human history can plausibly be identified not only as a global but also as an urban era. This is a period in which population, productive activity and wealth are highly and increasingly concentrated in cities.[1] Most cities offer a better standard of living for more people than ever before in human history; even the urban poor are better off, on average, than the rural poor around the world. Cities are primary centres of scientific, cultural and social innovation (Glaeser, ~~2012;~~ Hall, 1998). AQ1 Cities have also proliferated all over the

Corresponding author:
Michael Storper, Department of Geography and Environment, London School of Economics, Houghton Street, London, WC2A 2AE, UK.
Email: storper@ucla.edu

globe and have become increasingly interdependent so that where once we could speak quite meaningfully of "national urban systems" (most extensively developed in the Global North) the current situation is one marked by an increasingly integrated worldwide network of cities together with an extraordinary surge of urban growth in the Global South (McKinsey, 2011). But this era is also in some ways a dark age as marked by gutted-out old industrial cities, concentrated poverty, slums, ethnic conflict, ecological challenges, unequal access to housing, gentrification, homelessness, social isolation, violence and crime and many other problems. There has been a corresponding proliferation of academic and policy-related research on cities and a vigorous revival of debates about the content and theoretical orientation of urban studies.

In this paper we discuss three currently influential perspectives on these debates, namely, postcolonial urban analysis, assemblage theoretic accounts of the city, and the theory of "planetary urbanism". In their different ways, each of these three bodies of work attempts to provide bold understandings of the empirical trends referred to above. At the same time, each of them seeks to present an account of the city that poses strong challenges to much if not most hitherto existing urban theory. As such, these perspectives are prominent expressions of a renewed vibrancy and innovativeness in urban studies – reflecting the dramatically shifting geographies of urbanisation noted above – but in ways, as we shall argue, that often appear to be highly problematical. It should be stressed, at the outset, that these three bodies of work have different points of intellectual origin and different points of emphasis, though postcolonial and assemblage-theoretic approaches do share significant conceptual common ground, notably their focus on particularity, localism and difference, and an insistence on the empirical "complexity" of socio-spatial arrangements. Planetary urbanism for its part concentrates on an attempt to reformulate the relationship between "concentrated" and "extended" forms of human settlement, land use and spatial development by assimilating both of them into a theoretical urban landscape that is nothing less than global.

We will question these three approaches in a variety of ways. We will argue that each of them contains major blind spots and analytical distortions and that each has failed to offer a meaningful concept of urbanisation with generalisable insights about the logic and dynamics of cities. These weaknesses are not only regrettable in their own right, but are notably disabling in a field where the need to frame viable policy advocacies in search of social justice has become more and more insistent. In addition, we will argue that much of the current literature associated with these three approaches shares a predilection for certain kinds of convoluted philosophical and epistemological abstractions that actually present barriers to any understanding of the urban as a concrete social phenomenon. We begin our discussion by briefly re-stating ideas developed in an earlier paper (Scott and Storper, 2015) where we seek to establish a foundational concept of the urban. On that basis, we claim that there are fundamental common genetic factors underlying urban patterns, and a robust set of conceptual categories within which urbanisation processes and urban experiences can be analysed, wherever they may occur in the world. We then proceed to discuss in some detail what we take to be the most egregious weaknesses of the three main targets of our critique. As we work through this agenda we also offer a few replies – though less than a complete response – to a number of critical assessments of our earlier paper.

The nature of cities revisited

Towards an analytical understanding of the city

In Scott and Storper (2015) we dwelt in part on the high levels of diversity and disagreement in urban theory over the last century or so, and we asked if a coherent, stable theory of the city could be constructed. Such a theory, if it were possible, would need (a) to account for the genesis of cities in general, (b) to capture the essence of cities as concrete social phenomena and (c) to make it possible to shed light on the observable empirical diversity of cities over time and space.

Our approach to this theory-construction challenge was to build on the observation that cities are everywhere characterised by agglomeration involving the gravitational pull of people, economic activities and other *relata* into interlocking, high-density, nodal blocks of land use. The primary, but by no means the only mechanism driving this fundamental tendency, we argued, is the emergence of organic divisions of labour in which social and economic life (i.e. the production of goods and services, but also including cultural, religious and governmental pursuits) is organised and reorganised within networks of specialised but complementary units of human activity. This form of organisation means, in turn, that mutual geographical proximity or agglomeration of these units is crucial, for otherwise the time and distance costs of interaction would impede their operational effectiveness. In our earlier paper, we argued at length that all cities throughout history are based on this fundamental process of agglomeration The costs of covering distance were no doubt much higher at earlier periods of history, but as the copious literature on agglomeration dynamics reveals, proximity through co-location is imperative for certain types of activities even today (Cooke and Morgan, 1998; Fujita and Thisse, 2002; Krugman, 1991; Scott, 2012; Storper, 2013). A further major point must now be made to the effect that since the interdependent specialised activities that constitute the division of labour (and the residential housing associated with them) cannot all occupy a single point, they must necessarily sort themselves into a spatially extensive lattice or patchwork organised around their common centre of gravity and characterised by intricate internal patterns of geographic differentiation. We call any system of this sort an *urban land nexus* (cf. Scott, 1980).

These trans-historical and trans-geographical urban processes take on specific concrete attributes that reflect the wider – and ever changing – social, economic and political conditions within which urbanisation is always embedded. We can identify five basic variables or forces that shape the principal variations of the urban land nexus at different times and places. These can be enumerated as (a) the overall level and mode of economic development, (b) prevailing resource allocation rules, (c) forms of social stratification, (d) cultural norms and traditions and (e) relations of political authority and power. We do not have the space here to work out even a schematic description of the empirical diversity that these (and other) contextual variables are capable of generating, but they lead to a great deal of detailed variation in the urban land nexus from one instance to the next. For example, Imperial Rome, Xi'an in China, ancient Babylon, Timbuktu in the Empire of Mali, Tenochtitlan in 15th century Aztlán (contemporary Mexico), Manchester in the industrial era in Britain and Los Angeles, Mexico City and Hong Kong in the 21st century are all quite different from one another at one level of analysis, even as they all share in a common set of fundamental genetic forces. In view of the play of these differentiating variables, and notwithstanding our theoretical generalisations regarding the urban land nexus, we disagree with Dick and Rimmer (1998) who state that cities in various far-flung parts of the world are now converging towards a standard template. For the same reason, we also reject the claims of Roy (2015) when she describes our earlier paper as an attempt to construct a universal history whose objective is to obliterate "historical difference".

Not only does our analysis provide us with the tools for distinguishing between the general and the particular in urban outcomes, but also for separating out that which is distinctively and inherently *urban* from the rest of social reality. In particular, we must distinguish between phenomena that occur *in* cities but are not generated by urbanisation processes as such, and phenomena that are legitimately elements *of* cities in the sense that they play an active role in defining the shape and logic of urban outcomes. Thus, a hospital located in an urban area will usually play an important role as an element of the urban land nexus, both as a specific kind of service provider and as a catchment point for those who use its services, but its internal administrative arrangements are not likely to be of much relevance to any understanding of the city. Similarly, the interest rate, ideologies of imperialism or the price of sugar are not intrinsically urban; or rather, they can be said to have urban significance only insofar as they can be shown to play some role in the dynamics of the urban land nexus. A further illustration of these remarks is offered by the phenomenon of poverty, which has important urban dimensions but also has many substantive and relational manifestations that are not generated by the urban as such. To state this in another way, measurements of inequality or poverty in cities are not equivalent to the claim that inequality or poverty are basically engendered by cities (Sampson, 2012). In capitalist or market economies especially, poverty is not fundamentally caused by urban processes, but by the complex forces that shape income distribution in an economy marked by private property, competitive markets and wage labour. Equally, although researchers often use urban entities as units of observation in various kinds of statistical exercises (just as we use counties, states or countries for the same purpose), this alone does not endow these exercises with *intrinsically* urban meaning. The claim that any phenomenon occurring in a city is urban by nature is – without further specification – liable to the error of ecological fallacy. Political outcomes in the city, too, need to be carefully scrutinised in order to distinguish the specifically urban from what is merely contingently so. In particular, the urban land nexus is by its very nature subject to peculiar and endemic forms of politicisation. The tensions created by competition for land uses, the urge to secure access to positive externalities and to avoid the effects of negative externalities, the rent-seeking behaviour of property owners and the need to protect or enhance certain kinds of urban commons (such as agglomeration economies), among other frictions, all create constantly shifting circles of urban social collisions. Urban governance arrangements, too, or what Molotch (1976) called the urban "growth machine", are in significant ways caught up in these frictions through their functions as suppliers of public goods and services and their role as mediators of urban conflicts.

From these comments it follows (and even though we have affirmed that we live in an urban era in the sense that cities formally represent the principal geographic containers within which contemporary human society unfolds) that not all aspects of life, perhaps not even most aspects, can be understood as being necessarily (that is, "ontologically") urban phenomena in the very specific meaning as identified here. For these reasons, too, we are reluctant to accept Lefebvre's (1970) proposition that we are evolving in the direction of a full-blown "urban society" with its implied sub-text to the effect that society and the city are becoming one and the same thing. Similarly, the remark by Taylor (2013) that cities constitute the essential motors of all human society, politics and economy throughout history, and that hence all social science must become "city-centric", is clearly exaggerated. With a conscientiously delimited and focused concept of the city it is possible to identify how the urban generates specific kinds of social phenomena and sets them apart from non-urban phenomena. This is what provides a distinctive place for urban analysis in the academic division of labour and what, together with an appropriate analytical machinery, endows it with a central mission.

Some practical and theoretical implications

The urban land nexus emerges in the first instance out of dynamics of agglomeration and accompanying processes of land-use sorting, thus generating a complex lattice of locational activities over a shared gravitational field. In capitalist systems, significant parts of the urban land nexus are subject to the rule of private property and are hence commodified. In other types of social systems, land use decisions are apt to be directed by different kinds of mechanisms involving, say, limited or non-existent individual property rights or communal regimes of ownership (such as ethnic or clan rule).

Whatever the system, individual units of land ownership always have more than a purely private, atomised dimension. More specifically, agglomeration, proximity and density result in many different kinds of externalities (positive and negative) that circulate through the urban land nexus so that land use at one location invariably has impacts on other locations. Positive outcomes from agglomeration include processes of sharing (e.g. the joint usage of large-scale infrastructural artifacts), matching (e.g. the local availability of many alternative choices to purchasers and sellers of goods, services or labour) and learning (e.g. the rapid diffusion of cultural or technological information), which in part ~~accounts~~ for the dynamism we typically associate with cities throughout history and especially in capitalism (cf. Duranton and Puga, 2004). Negative outcomes may include the congestion, land use incompatibilities, incentives to crime, segregation and inequality, social conflicts and other undesirable consequences that arise out of the dense coexistence of highly differentiated social and economic activities in a relatively restricted spatial orbit. The importance and pervasiveness of these effects means, as already suggested, that some form of collective, non-individual control is necessary if the city is both to avoid internal blockage and if the individuals, households and firms that it contains are to seize jointly on strategic developmental opportunities. This explains in large degree why the individual decision-making and behavioural mechanisms of the urban land nexus are virtually everywhere regulated by collective governance arrangements designed to safeguard cities against implosion and stagnation (Roweis and Scott, 1977). Within the city, interrelated units of economic production typically form distinctive clusters interpenetrated by swaths of residential activity. Areas outside the city are sources of the food, resources and materials that are not produced internally; and they offer, in addition, markets for the city's tradable, specialised products. These areas are represented both by the immediate hinterland of the city and other cities and regions at more distant locations. Even in ancient times, long-distance trade was characteristic of many cities, as exemplified most dramatically by the case of Classical Rome. In the 21st century, cities interact with one another in a globally-integrated system of trade and information exchange as expressed in an emerging global mosaic of cities and city-regions.

In the light of these remarks, we can now state that the city represents a *very specific scale of economic and social interaction generated by agglomeration processes and focused on the imperative of proximity, and almost always endowed with governance arrangements that attempt to deal with the problematical effects of density and propinquity*. At the same time, the city is always embedded in a far-flung spatial economy that sustains it without compromising its integrity as a distinctive social phenomenon (Fujita et al., 1999). Accordingly, as we shall argue in more detail later, it cannot simply be dissolved away by fiat into a sort of overarching global plasma as theorists of "planetary urbanization" proclaim (e.g. Angelo and Wachsmuth, 2015; Brenner and Schmid, 2015). Our argument thus goes strongly against the grain of the main theses of planetary urbanists or those who, like Amin and Thrift (2002), claim that "the city is everywhere and nowhere". *A fortiori* we stand in opposition to those urbanists who state that the idea of the city is purely ideological; and in view of our characterisation of the urban land nexus as an overarching phenomenon that integrates urban space into a coherent social unit, we reject ~~the~~ Roy's (2015) characterisation of the city as a thing of "shards and fragments".

This is also an appropriate moment to allude to some of the criticism that has been made of our earlier analysis on the grounds that it is "economistic" (Mould, 2015; Roy, 2015). Given the primary role that we ascribe to economic forces in the genesis of the urban land nexus, this line of critique is entirely predictable but essentially misinformed. We assuredly do propose that the origins of the urban land nexus reside in the economic tensions engendered by the division of labour and agglomeration (and we offer strong justifications for this position), but our claim is very far indeed from any argument to the effect that cities are exclusively or monocausally structured by economic variables. Indeed, we have explicitly suggested that diverse other social, cultural and political forces are also at work in shaping the urban land nexus. Accordingly, our response to the charge of economism is twofold. On the one hand, we invite our critics to identify exactly what it is that they mean by "economism" (a term that is almost always vacuous in actual usage[2]). Our own suggestion here is that the most demanding sense in which the term can be used is to reserve it strictly for *situations where that which is not economic is erroneously proclaimed to be economic* (e.g. claims to the effect that the level of economic development determines the form of sociability in urban neighbourhoods or that the city is *nothing but* an economic phenomenon). On the other hand, we challenge our critics' attempts to characterise our work as economistic by asking them to go beyond purely gestural allegations and to demonstrate in disciplined critical detail how our formulations about the analytical origins of cities might actually be wrong and how they can be corrected. In fact, a close reading of our text should make it abundantly clear that our theory of the urban land nexus remains open to an enormous diversity of non-economic elaborations and hybridisations, and, indeed, to any number of complex reflexive relations between the economic and the social, political and cultural dimensions of urban life. Moreover, although this point surely should not need to be made explicit, the urban land nexus obviously is a fundamental key to understanding the city as a locational matrix of built forms and associated symbolic assets, which, according to the views of Walker (2016), are culpably absent from our own analysis.

This brief exposition of our theoretical views now serves as a point of reference against which we will review and criticise a number of currently fashionable theories of urbanisation that we take to offer seriously flawed accounts of both the scientific and political challenges posed by cities today.

Postcolonial urbanism: Cosmopolitan but provincial

Much contemporary postcolonial research originated in cultural and historical studies where it has functioned as a critique of numerous blind spots in Northern traditions of theoretical analysis. Above all, postcolonial thinking, as represented, for example, by Said (1978) and Spivak (2008), demonstrates how diverse intellectual legacies of colonialism (ethnocentric biases and prejudices in particular) enter unconsciously into scholarly writings about the Global South. Postcolonial scholars (such as Comaroff and Comaroff (2012)) are also, and correctly, intent on showing that the claims of universality that Euro-American theory has often arrogated to itself are sometimes demonstrably false. These same lines of thinking and critique have recently become strongly influential in urban studies. Robinson (2006, 2011) and Roy (2009, 2011), among many others (for example, Edensor and Jayne, 2012; Myers, 2014; Ong and Roy, 2011; Patel, 2014; Sheppard et al., 2013), have been notably vocal in this regard, and have been especially outspoken in decrying the application of urban theories constructed in Europe and North America to cities in the Global South.

These and other analysts have sought to correct what they see as imbalances and misrepresentations in Northern urban theories by means of two overlapping strategies. One is to call for more cosmopolitan forms of urban theory (what Ong and Roy (2011) refer to as "worlding") that take seriously the experiences of the cities of the Global South. The other is to insist on the irreducible core of idiosyncrasy that marks every city and to focus on the resulting play of empirical "difference" and "complexity". A further important point of departure for postcolonial urban scholars resides in the notion of the "ordinary city" developed by Amin and Graham (1997) to the effect that cities are all equally distinctive and unique and that none can be claimed to function as a privileged archetype or exemplar relative to the others. Robinson (2006), in particular, has appealed to this notion ~~as~~ by way of asserting the equivalent standing of all urban centres across the North-South divide, as well as by way of proclaiming that any meaningful problematic of the urban must focus intently on the essential character of cities as sites of difference. In a more radical vein, Roy (2009: 820) has advocated sweeping much of extant urban theory away with the peremptory injunction that "the center of theory-making must move to the Global South". However, as Peck (2015) points out, there is an apparently unresolved tension in postcolonial studies between constant calls for a worlding of urban analysis on the one side and the equally constant affirmation of a North/South binary on the other, and even, in some cases, as we shall see, a tendency to favour a wholesale "provincialization" of urban theory (Ren and Luger, 2015).

Postcolonial commentators are especially dissatisfied – not always incorrectly but frequently without appropriate nuance – with what they allege to be the pervasive modernist and developmentalist biases of urban theory as elaborated in the Global North. One of the most baleful cases of this kind of bias, in the view of these commentators, is represented by the Chicago School of Urban Sociology. A particular point of contention is the Chicago School's notion of the folk-urban continuum comprising primitive, non-urban social formations on the one side, and advanced, urbanised social formations on the other, and the extension of this notion in the work of Wirth (1938). Postcolonial urban theorists criticise modernism-developmentalism as a discourse that consigns the cities (and societies) of the Global South to the status of underdevelopment and backwardness, an outlook that is manifest, according to Roy (2011: 224) in "apocalyptic and dystopian narratives of the slum". She herself sees the poverty, informality, marginalisation and extensive slums of Southern cities as a *mode* of urbanisation (Roy, 2005; emphasis in the original). Quite what this phrase might mean is difficult to determine, but it presumably functions as a gesture intended to eliminate the allegedly pejorative implications of Northern theory. Modernism-developmentalism is further criticised by postcolonial scholars for its promotion of a teleological concept of cities in the Global South in which growth and change are alleged to be subject to evolutionary stages involving shifts from less to more modern and developed. The more specific claim here is that it is unreasonable to expect any linear movement from less formal to more formal arrangements in regard to settlement-building and property rights in the cities of the Global South (Roy, 2005)

The critique of postcolonial urban studies

Obviously, cities of the Global South have been severely overlooked in past research efforts;[3] obviously we must be careful to pay attention to the specificities of these cities; and obviously we need to acknowledge that urban theory must now range over the entire world for its sources of data and evidence while remaining fully open to new conceptual insights generated out of the experiences of the cities of the Global South. Equally obviously, we must beware of the dangers of Eurocentrism, by which we mean theoretical overreach based on limited evidence derived from Northern cities, but that is inappropriate or irrelevant with respect to Southern cities. Where postcolonial urban theory errs, we argue, is in its own peculiar forms of critical overreach and its overall commitment to what we have called a "new particularism" (Scott and Storper, 2015). In what now follows, we address what we take to be three major failures of postcolonial urban theory, namely, its exaggerated complaints regarding Euro-American epistemological bias in contemporary urban analysis, its highly selective critique of modernism-developmentalism and its strong methodological commitment to theoretically-unstructured comparativism. Note that all of these themes are essentially branches of a single meta-claim, that of a set of incommensurabilities: in point of view, in development and in representativeness.

Eurocentrism and the provincialisation of knowledge. To begin, then, postcolonial urban studies are broadly motivated by the claim that theory produced in the Global North is *inescapably* unable to account for empirical situations in the Global South (see Peck (2015) for an analogous characterisation of postcolonial theory). Roy (2009) adds the further damning claim that Euro-American urban theory "… keep(s) alive the neo-orientalist tendencies that interpret Third World cities as the heart of darkness, the Other" (though we can think of no scholarly paper on cities published in at least the last half-century that would bear this assertion out). In harmony with these judgments, many urban theorists with a postcolonial bent (notably Sheppard, 2014; Sheppard et al., 2013) state that theories must necessarily be local and confined in their empirical reach to specific segments of geographic reality. As Leitner and Sheppard write (2015), "… no single theory suffices to account for the variegated nature of urbanization and cities across the world, without asserting the necessity of different distinct theories for different contexts". And: "Our position, then, is that there can be no single urban theory of ubiquitous remit". Even though the authors fail to define what they mean by "different contexts" and how we might identify them, they then go on to call for self-conscious "provincialization" of urban theory as a virtue in itself and as a way of delegitimising what they see as the pervasive pretensions to universalism of European and American urban theory.

A first direct and simple answer to this call to provincialise theory is to ask for a clear and direct demonstration of the fundamental incommensurability of urban phenomena in different parts of the world, above and beyond assertions about empirical diversity. A second is to propose a counter-argument, as we have done (and which we offer for disconfirmation), to the effect that there are indeed theoretically generalisable features of urbanisation as a whole. Of course, we know from the work of Livingstone (2014) and others that theoretical work very often *does* unconsciously reproduce geographical and ideological biases reflecting the circumstances in which it arises, and urban theory is no exception to this observation. Moreover, various streams of philosophy and historiography, most especially since the middle of the 20th century, clearly recognise the social constructivist character of all intellectual activity (Haraway, 1988; Kuhn, 1962; Mannheim, 1952). This work points not only to the essential social and historical foundations of all forms of discourse, but also to the absence of any Archimedean point from which knowledge claims can be fully and finally adjudicated. These comments signify that knowledge is always provisional and motivated by human interests (Habermas, 1971), and in some cases (e.g. imperialist accounts of dominated peoples) can be grotesquely distorted representations. So far so good. It by no means follows, however, that ideas can never attain to universal value, or, more simply that an idea developed at place *a* must *invariably* fail when transferred to place *b*. This is a matter for step-by-step judgment, not for a blanket *diktat*. By contrast, commitment to the notion that theories must be provincialised as prescribed by Leitner and Sheppard (2015) calls for a clear identification of what constitutes a meaningful "province", and in the absence of any operational guidelines in this matter (as in the work under evaluation here) amounts to little more than an arbitrary and self-defeating preference for intellectual parochialism at the expense of more searching theoretical generalisation. At the same time, and as a corrective to the one-dimensional critique of Northern theory that is offered by postcolonial urban scholars, many of these same issues of bias and ethnocentricity are ones that theorists have struggled with since the Enlightenment, above all in regard to the question as to what constitutes the common or universal features of humanity and what in different contexts represents essential differences in human behaviours and aspirations (Pagden, 2013). The tensions in this duality were especially prominent in European debates over the 18th and 19th centuries (and even as far back as the 16th century if we consider Montaigne) about the nature of distant "others".

We can see in more detail why the critique of "Northern" theory by postcolonial urban scholars is unduly one-sided by examining how these scholars deal with modernism and developmentalism.

Modernism-developmentalism. There can be little doubt that some versions of modernist-developmentalist theory impose misguided concepts not only on the cities of the Global South but also on those of the Global North. The implausible mechanical model of stages of growth is one such theory. The same can be said for the organic-ecological model of neighbourhood succession as developed by the Chicago School, which is especially suspect given its Darwinian undertones and its association with the concept of the folk-urban continuum (Robinson, 2006).

Whatever the failures of these particular theories may be, scholars in both the Global South and the Global North are in practice faced with the evident empirical fact of the marked differences in levels of economic development and income exhibited by cities in different parts of the world and the effects of these differences on urban outcomes (cf. Smith, 2013). Acknowledgement of the powerful role of economic forces in shaping the urban milieu is not to advocate any sort of teleology of urban history, with all cities everywhere eventually converging to a state of achieved "modernity". To the contrary, we recognise that the empirical trajectories of development followed by individual cities vary markedly, both within the Global South and North as well as within single countries. Over time, some cities grow at an accelerated pace; some grow rapidly and then decline; some remain in a proto-capitalist state of development; some are prosperous while others are impoverished; some specialise in manufacturing while others are more given to service provision; some have dependent branch-plant economies while others become centres of innovative high-income entrepreneurialism; and so on through any number of possible variations. Throughout all of this diversity, however, there remains the burning issue of how specific forms and levels of economic development shape specific variants of agglomeration and high-density land use – in other words, the urban land nexus – and how this in turn feeds back upon those same forms and levels.

Postcolonial scholars' fixation on the supposed exceptionalism of the Global South is particularly evident in their treatment of such favoured themes as poverty, slums, informal labour markets, vulnerable property rights, inadequate infrastructure and lack of sanitation (Roy, 2005). These themes are frequently dealt with as though they had no family resemblance to similar issues in the Global North. Yet we only need think of Charles Dickens' London, Emile Zola's Paris and Sinclair Lewis' Chicago, or more recent cases of deprivation and spatial exclusion in Europe and America revealed in the studies of Chetty et al. (2014), Sampson (2012), Standing (2011) and Wilson (1987), to recognise that there is much in common between the cities of the Global North and the Global South in regard to poverty, and that examination of the former has much to offer to scholars of the latter, and vice versa. These remarks lead on to consideration of postcolonial scholars' approach to developmental issues generally, and in particular, as Chibber (2013) points out, their claims about developmental theories in the Global North being simple and linear as compared with the experience of the Global South where development is said to be complex and non-linear. Of course, as we have already pointed out, these claims about Northern development theories are a misrepresentation. Many different formulations of the diverse Northern routes to development have long constituted one of the principal axes of debate within Northern historical research (Allen, 2009; Aston and Philpin, 1987). Even so, post-colonial scholars continue to assert these claims as background to their view that urban development of the Global South is so unique as to defy any theoretical description that might establish commonalities with cities elsewhere. In other instances, post-colonial scholars (such as Robinson, 2011) effectively shift questions about the interrelations between economic development and urbanisation into the distant background as nothing but Northern theoretical fantasies irrevocably marred by Eurocentric parochialism, reductionism and teleological thinking. In fact, in both the

North and the South, despite many empirical differences of history and geography, the shifting forces of capitalism and markets and their expression in production, trade and employment pose a consistent set of conceptual problems. These include the ways in which capitalist and non-capitalist systems articulate with one another, as in the case of the co-agglomeration of producers in the informal and formal sectors in India as described by Mukim (2015; see also Rey et al., 1971) or the dynamics of informality in American cities offered by Mukhija and Loukaitou-Sideris (2014). Certainly, there is much in the way of difference and idiosyncrasy to investigate in cities around the world, but theory is required for this sort of investigation to take on any wider meaning. We have posed a generalised theoretical framework of the urban land nexus intertwined with five crucial processes shaping the specifics of urbanisation in different times and places (level of development; resource allocation rules; forms and levels of social stratification; cultural norms and traditions; authority and power). This framework precisely addresses the need to acknowledge diversity, but without falling into the sophism of particularism and thereby losing sight of the forces that affect all cities.

Comparativism and its limits. One of the ways in which postcolonial scholars seek to compensate for their deep scepticism about much current urban theory is by means of the "comparative gesture" stressing "thinking across differences" in ways that are "potentially open to the experiences of all cities" (Robinson, 2014: 57). A comparative approach is especially congenial to postcolonial studies because it is assumed to reduce the dangers of aprioristic thinking about cities and the inappropriate imposition of alien concepts on given empirical situations. Myers (2014), for example, offers what is intended to be an exemplary comparative account of multi-racial policy approaches to urban transportation policy in Capetown and Nairobi on the one hand and Hartford, CT on the other. This account turns on the proposition that the experiences of these African cities in the matter of community development could have usefully informed the design of the Hartford-New Britain transportation corridor, and it is no doubt interesting and pertinent within its own limited terms of reference. Yet like so much other work in this comparative genre it signally fails "to transform existing conceptualizations" as optimistically promised by proponents of the genre (Robinson and Roy, 2015: 3). Our own argument is that the well-travelled but narrow road represented by comparative and classificatory methods certainly adds a number of legitimate procedures to the social scientist's toolbox. However, if comparisons are to be effective, they can never proceed on the basis of theoretically uninformed choices about cases for comparison or the specific variables that are isolated for examination. *Prior conceptual labour about these matters is essential if comparative methodologies are to produce – other than by accident or good luck – significant results.* This means specifically that we need to have a degree of conceptual clarity or intuition about the issues under examination in order for comparison to proceed in a way that reveals consequential insights when different empirical situations are brought into conjunction with one another.

To be sure, the comparative gesture can be useful and interesting, but our point is that a more theoretically self-conscious pooling of data, experiences and investigative results is essential if urban investigations are to progress beyond localism, difference and the celebration of empirical complexity for its own sake. For this reason, there is much to be said in favour of identifying theoretically meaningful categories of cities that contain multiple cases of similar but not identical cities. This point is recognised by Robinson (2011), though her argument is strongly in favour of inductive analysis of the data. Other postcolonial urbanists appear to be rather more ambivalent about this issue. For example, Bunnell and Maringanti (2010) and Roy (2009) raise arguments against categories like world cities, international financial centres and city-regions, not so much because they may be misidentified but because they are said, somehow or other, to relegate excluded cities to secondary status while supposedly diverting our attention away from the full diversity of urban forms and experiences that the world has to offer. Here again we come face to face with the new particularism and the relegation of all cities to the status of "ordinariness". This insistence on difference and idiosyncrasy within ordinary cities is especially evident in the case of another favourite target of postcolonial critique, i.e. the notion of the representative or exemplary city, and, notably, the writings of the so-called LA School about the "paradigmatic" status of Los Angeles towards the end of the 20th century. We would be the first to acknowledge the incautiousness of many of the LA School's theses, and yet it is important to record that Los Angeles, in its pioneering status as a globalising post-fordist centre of flexible specialisation, disorganised labour, growing social inequality and polycentricity did indeed turn out to be an early and powerful expression of several incipient worldwide trends (Soja and Scott, 1986). Accordingly, the LA School called attention at an early stage to a developmental pathway that many other cities all over the world have subsequently followed, and from this perspective it was most certainly exemplary.

McFarlane's railway ticket

Over the last few decades, "assemblage theory" has emerged as a major genre of work in urban studies, as in the social sciences in general (see, for example, DeLanda, 2002; Latour, 2005). This theory, which has many affinities with postcolonial urbanism, has filtered into urban studies from the work of continental post-structuralist philosophers (in particular, Deleuze and Guattari, 1972). Assemblage theory is of considerable complexity in its philosophical representations, the understanding of which is not made any easier by the *langue de bois* favoured by its chief protagonists and the sharply conflicting interpretations of their work by secondary commentators. For our purposes, however, focused as they mostly are on abridged "applications" of this theory to urban analysis, a few essentials will suffice to motivate our critique.

Assemblage theory is first and foremost an ontological view of the world conceived as a mass of rhizomatic networks or finely-grained relationships constituting the fundamental character of reality. These networks bind together unique human and non-human objects within fluid, hybrid mosaics forming more or less temporarily stabilised systems of interconnections representing the current state of the observable world. Assemblages become stabilised by "territorialization" (as opposed to destabilising deterritorialisation) when they are anchored to particular tracts of geographical space. Importantly, any state of reality in this theory is taken to be "flat" in the sense that any perceived hierarchical or scalar ordering (from a top to a bottom) decomposes back again into the kaleidoscopic, rhizomatic and horizontal relations that are said to constitute it (DeLanda, 2002; see also Marston et al., 2005). This point is largely shared between assemblage theory and postcolonial theory, via the latter's emphasis on difference and its focus on the incommensurable uniqueness of cities.

There are several variants of assemblage theory, but one of the most influential is actor-network theory, a body of ideas associated above all with the work of Latour (2005). This is, again, a way of exploring the multiple relationships that tie human and non-human objects together, but with the additional claim that all of these objects are constituted as actants, i.e. capable of agency in the sense that they exert effects on other actants. In a number of methodological and theoretical publications, Farías (e.g. 2010, 2011) has outlined the main implications of assemblage and actor-network theory for urban studies. He writes that the city is "… an object which is relentlessly being assembled at concrete sites of urban practice, or, to put it differently, as a multiplicity of processes of becoming, affixing sociotechnical networks, hybrid collectives and alternative topologies" (Farías, 2010: 2).

Some of the multiple ways in which the urban might be assembled are then enumerated:

> … as a transport system, as a playground for skateboarders and free-runners (parkour), as a landscape of power, as a public stage for political action and demonstration, as a no-go area, as a festival, as a surveillance area, as a socialization space, as a private memory, as a creative milieu, as a jurisdiction, etc. (Farías, 2010: 14)

This conception then leads to a descriptive, anecdotal and notably indiscriminate approach to urban investigation. Farías (2011: 367) with apparent faith in the powers of inductive empiricism goes so far as to say that "we don't know what we are looking for until we find it". Little wonder, then, that Brenner et al. (2011) characterise this line of research as "naïve objectivism" and point to its failure to distinguish between the significant and the insignificant in urban analysis. Certainly, the assemblage approach is potentially of positive value in certain kinds of ethnographic and narrative accounts of the city such as those offered by de Boek and Plissart (2004), Mbembe and Nuttall (2004) or Simone (2014); and as Geertz (1973) has shown, thick description of social practices and material forms in cities or elsewhere can often provide sensitive depictions of the ways in which social lives are woven together. One example might be the complex manner in which we build high-rise downtown environments in major cities and the connections of this process to the construction industry, architectural practices and building norms, competing demands for space, visual conceptions of the built environment and office employment in the city. Our critique of assemblage theory therefore does not deny the possibility of certain important feedbacks between non-human objects and human society and it is emphatically not intended to repudiate the reflexive relations between technology, urban space and social life (cf. Graham and Marvin, 2001). However, we certainly do have strong reservations about the capability of inanimate objects to "act" as if whatever causal or generative powers they may possess were ontologically equivalent to sentient, purposive human behaviour. AQ3

Assemblage theory radically privileges the activity of assemblage itself, seeing no wider forces that might determine what assemblages are possible or not possible; rather, it advocates a methodology of building the elements of social organisation a posteriori from the ground up (Bender, 2010) and focusing on specific sites of daily life (Simone, 2011). The result is a largely indeterminate concept of the city as a complex, variegated, multifarious, open-ended, fluid, unique, hybrid, unruly, nonlinear, etc., etc. aggregate of disparate phenomena tied together in a haphazard mix of causal and contingent relationships. This concept, like Robinson's (2011: 13) (postcolonial) view of the city as "a site of assemblage, multiplicity, and connectivity" is at one level of observation certainly correct, but at another level interposes mere empirical convolution as a substitute for a deeper and more systematic level of (theoretical) comprehension. This naïve objectivism frequently also results in markedly indiscriminate bodies of information being packaged into empirical assemblages, perhaps especially where, in the words Acuto (2011: 553), those "missing masses of non-human actors often degraded to the role of mundane artefacts" are brought into concatenation with human life. In other words, there are no theoretical guideposts in assemblage theory for telling us how tease out significant relationships or to distinguish between the trivial and the important.

We may further pin down these remarks by reference to the work of McFarlane (2011a, 2011b), another prominent spokesperson for assemblage theory. In his lively defense of this theory, McFarlane (2011c: 216) draws on his research on poverty and informal housing in Mumbai. He insists that any attempt to understand "the everyday lives and hardships faced by the poor" requires us to pay attention to an eclectic collection of "urban materialities", which include in this instance such disparate objects as sackcloth, corrugated iron, brick, breezeblock, hydroform and infrastructures of drainage, sanitation, water and electricity. These elements are then organised into a description of poverty in Mumbai, but critically the account – which has the analytically "flat" quality prescribed by assemblage theory – is essentially devoid of useful explanatory ideas. The same can be said for the analogous work of Dovey (2012) who, like Simone (2011), puts forth an extended display of deleuzoguattarian[4] jargon in an attempt to illuminate descriptions of urban informality that nevertheless remain uninformative about the basic logic of social and economic marginalisation.

The fetishisation of inanimate objects as instruments of agency is dramatically highlighted in McFarlane's (2011: 217) discussion of the work of political activists in combatting poverty in Mumbai. He points to the way in which these activists discovered that they could make free telephone calls by inserting a railway ticket into a receiver, and by this means greatly extend their outreach. McFarlane then states that this is all part of the "experience and possibilities of urban life". Our point, by the way, is not to dismiss this kind of narrative as meaningless in principle. A good story is a good story, after all. Our point is rather that in the case of McFarlane's railway ticket a trivial contingency is in all seriousness offered as a link within a chain of agency that is supposed to function as a way of understanding urban poverty and as an informed account of the struggles that people engage in to escape from it. This picture sharply contrasts to the more analytically controlled realism about obstacles in the way of the poor that is painted by such diverse authors as Aw (2013), Boo (2012), Caldeira (~~2001~~) and Cole (2014). As Brenner et al. (2011: 233) write of actor-network theory: "This mode of analysis presupposes that the 'facts'—in this case, those of interconnection among human and nonhuman actants—speak for themselves rather than requiring mediation or at least animation through theoretical assumptions and interpretive schemata". Thus, in the flattened world of assemblage theory there is a perilous tendency to fail to distinguish between the inanimate character of material objects and the intentionality of humans, and to compound this oversight by undertheorised presentations of social interconnectivity (cf. Tonkiss, 2011). This flattening of the world also evacuates any meaningful political content from assemblage theory since everything is equally important (or equally trivial and unimportant).

The principal problems of assemblage theory as discussed in this section of the paper – the notion of reality as mere rhizomatic entanglements without underlying processes of structuration, the indiscriminate attribution of agency to things and the absence of concepts of human action – make this theory unable to detect urban dynamics, movement, change and causality in meaningful ways. Critical realism has long provided a way out of this kind of dead-end by insisting on the importance of *necessary* relationships, causal powers and theoretical abstraction as fundamental to the identification of the central properties and conditions of existence of social phenomena (Sayer, 2004). One searches in vain in assemblage theory and urban research based on it to know what larger difference assemblages make, which assemblages are important and which are insignificant and fleeting, which are empowering and which are disempowering and what kinds of policy interventions are most likely to bring about desired forms of social change.

Planetary perplexities

Some urban analysts today, most notably Brenner and Schmid (2015), suggest that in the 21st century a radical blurring of the category of the urban versus everything else has come about, and that what were formerly identified as urban areas can no longer be distinguished from the rest of geographic space, conceptually or empirically. These are the central doctrines of "planetary urbanism". As Brenner and Schmid (2014: 750) write:

> It is clear that settlement-based understandings of the urban condition have now become obsolete. The urban cannot be plausibly understood as a bounded, enclosed site of social relations that is to be contrasted with non-urban zones or conditions. It is time, therefore, to explode our inherited assumptions regarding the morphologies, territorializations and sociospatial dynamics of the urban condition.

Given the geographically intensive and extensive development of global capitalism, the authors are doubtless correct to refer to an integrated planet-wide socio-economic system. They are also right to claim that the notion of a purely "rural" realm occupying the interstitial spaces between cities is archaic and misleading. The notion has never, in any case, been entirely satisfactory given the diversity of these territories (deserts, forests, mountain ranges, sites of peasant farming, expanses of industrial agriculture, spaces of resource extraction, tourist regions, etc.). But we are at a loss to understand how these facts can lead to a claim that the idea of the city only "persists as an ideological framing" (Brenner and Schmid, 2015: 152), a phrase that is echoed by Merrifield's (2013) characterisation of the same idea as a "pseudo-concept". Above all, as we show below, Brenner and Schmid do not conclusively demonstrate that the city fades away as an identifiable geographic entity and scale of socio-economic interaction within planetary space, or that any distinction between the urban and the rest of geographic space (what they misleadingly insist on calling "the rural") must now be abandoned; and they are merely baffling when they write about the full extent of planetary space as being "urbanized", especially when this includes "rainforests, deserts, alpine regions, polar zones, and oceans and even the atmosphere" (Brenner and Schmid, 2015: 152–153).[5]

Brenner and Schmid hedge their bets rather clumsily by saying that there is something called "concentrated" urbanisation, or what we usually call cities, and something called "extended" urbanisation, which more or less corresponds to everything else. The puzzle is why they want to introduce the semantic confusion that ensues from applying the term "urban" with all its familiar city-centric connotations to everything else when numerous other descriptive terms are quite conceivable.[6] We shall argue that not only is there no conceptual (or what they relentlessly call "epistemological") gain by this manoeuvre, but considerable theoretical loss. Here, Angelo and Wachsmuth (2015) enter the fray with their commentary on something that they allude to as "methodological cityism". They identify this forbidding sin with the error of ascribing to the circumscribed geographic structure of the city processes that they say are more properly to be analysed within the wider framework of Brenner and Schmid's "planetary urbanization". The plot thickens when Brenner and Schmid point out (correctly as it happens) that there are usually no simple or intuitively-identifiable boundaries between the city (concentrated space) and the rest of the world (extended space) so that the continuity between the two appears to be unbroken. This is a familiar problem that has always perplexed urban analysts, but Brenner and Schmid are wrong to think that the issue goes away by assimilating the whole of geographic space into an urban problematic. There is in fact a more satisfactory way of approaching this problem.

Consider, to begin with, certain kinds of phenomena that exist at the intra-urban level, such as neighbourhoods, slums, industrial quarters, central business districts and suburbs. Each of these phenomena represents a distinctive and multifaceted type of socio-spatial outcome within a wider urban space (the urban land nexus) and none is divided from the rest of the city by a clear line of demarcation. Yet each appears to us as an ontologically distinctive scale of urban space not only because of its empirical character but also because *each poses uniquely problematical scientific and political questions* deriving from its mode(s) of operation. Sampson (2012), for example, has shown that there are many and sundry "neighbourhood effects" on people who live in poor communities, and Chetty et al. (2014) have shown how these effects also have an impact on intergenerational poverty rates. Similarly, the vast literature on local economic development reveals that intra-urban clusters of production units are marked by powerful spatial dynamics that are uniquely problematical as objects of inquiry. All of these phenomena are embedded in and marked by all manner of continuities with the urban land nexus, but in no case is it useful or meaningful simply to dismiss them as ideological constructions. Two related points now need to be made.

First, the city is a composite social, political, cultural and economic phenomenon (anchored and integrated by the urban land nexus) that is very much greater than the sum of its parts, signifying, in turn, that it has a potent collective presence. In particular, the city is a site of joint dynamics with a joint identity (e.g. "the San Francisco Bay Area", "Rio de Janeiro") deriving from its character as an agglomerated land nexus. This state of affairs means that the wider political interests of the individual firms, households and other behavioural units that make up the urban sphere become entangled with a concrete set of political interests specific to the city (including those forms of conflict, coalition, exclusion and deprivation peculiar to the urban land nexus). These political interests are partly mobilised in collective action and are almost always associated with formal institutions (especially governmental institutions) that endow cities with powers of taxation, managerial regulation and the capacity to make substantial public investments). Among the more important concerns of these institutions is the performance of the city as a centre of employment, earnings and quality of life, again rooted in the urban land nexus (Molotch, 1976). For all of these reasons, the city at large – especially given its foundations in agglomeration and its dense institutional and political overlay – poses questions that are quite specific to the urban arena both as an object of scientific enquiry and as a scale of human political and economic life. This is why proponents of so-called methodological cityism are mistaken in their characterisation of the city as nothing more than an ideological mirage.

Second, just as neighbourhoods, slums, industrial quarters, etc. are distinctive and idiosyncratic socio-spatial articulations (albeit within the urban land nexus), so the urban land nexus itself is a distinctive socio-spatial articulation (within wider global or planetary space). The city, in a nutshell, is in important ways an irreducible collectivity and, as we argued earlier, its peculiar character derives from its properties as a locus of agglomeration, gravitation and density as well as from its specific daily and weekly rhythms of life. These rhythms are embodied most notably in its local labour markets and its regular patterns of commuting (Cheshire and Hay, 1989; Kerr and Kominers, 2015). To state this latter point in another way, cities concern us because distance is not dead, and substantial elements of our lives are anchored in these spatially-and temporally-constrained urban systems. The day when we can move with no cost in time or effort from one place to another (i.e. a world of "magic carpets") is the day when can say that the city is dead. But the overwhelming situation in the contemporary world is one in which – despite the growth of long-distance linkages – proximity and density remain critically important as arrangements that facilitate the still expanding volumes of detailed, small-scale, intimate and ever-changing interactions that lie at the heart of human relationships within the urban land nexus.

There remains an unanswered question. Even given the above discussion, where and how do we draw the dividing line between the city and the rest of geographic space? We have argued, with specific reference to the city, that in spite of the continuity/indivisibility of geographic space (or of reality as a whole for that matter) there are differing scales and articulations of empirical phenomena, underlying processes and political interests that make it imperative to distinguish specific units and levels of interaction within the totality of planetary space as a whole. Moreover, there is no rigid line that separates the urban land nexus definitively from the rest of geographic space, but rather a series of spatial gradations in which we move from the one to the other. This does not mean that the urban land nexus and its dynamics as identified above are illusions, just as neighbourhoods, slums, industrial quarters, etc. do not dissolve away into an urban totality, and just as the fact that the seasons fade gradually and unevenly into one another does not mean that they do not exist as identifiable phenomena in their own right. The evident deduction from these remarks is that we almost always have considerable leeway in practice as to how we demarcate the spatial extent of the urban land nexus, but that the best bet is to define it in any given instance in a way that optimises our ability to deal with whatever given question(s) we may have in hand (e.g. economic development, public transport, ethnic conflict, neighbourhood blight, urban political strategy and so on) while eliminating from consideration as much irrelevant territory as possible. In practice, we have little option but to follow the pragmatic rule of thumb that has always been adopted by geographers and to locate the line of division in some more or less workable way relative to available data.

One possible objection to these lines of reasoning is that cities have diverse functional connections to other places in many different parts of the world. Indeed Brenner and Schmid (2014, 2015), among others (e.g. Amin and Thrift, 2002), make the explicit claim that the identity of the city as a spatial unit is deeply compromised by the widening external relations that form its so-called "constitutive outside". Our response here is simple. These relations are capable of inducing certain kinds of changes in cities, such as bursts of growth in central business districts or changes in given population categories, but their effects are virtually always assimilated into the urban land nexus as such without destroying its integrity as a complex social unit. For example, the New York financial district has strong external connections to far-flung customers and sources of capital, resulting in the growth of local firms with diverse impacts on land use patterns in Manhattan and on workers' residential behaviours. Whatever the effects of the constitutive outside of the city may be, however, these in no way undermine the theoretical notion of the urban land nexus as the critical *constitutive inside* of the city. Indeed, the urban land nexus gains in terms of its internal complexity even as these effects intensify and multiply. Equally, and despite the fact that in the world system of the 21st century spatial interconnections have attained unprecedented levels of volume and geographic extension, the need for proximity and local interaction has in many ways been bolstered within the urban land nexus (Duranton and Storper, 2008; Hummels, 2007). Anderson and van Wincoop (2004), for example, have demonstrated that trade costs remain so important in today's world that they frequently reinforce a distinctively local scale of interaction. We should point out in any case that long-distance interconnections between cities have always been a feature of urban life, beginning in Jericho 6500 years ago, and, as we have already stated, they are typically a crucial condition of continued urban viability. As such, they do not represent the negation of the identity of the city but one of the conditions that have made the existence of cities possible throughout history.

Summing-up: A challenge to urban theory and research

We have tried in all of the above to blast open a number of theoretical geographies of cities and urbanisation processes, and we have criticised in particular certain recent trends that for one reason or another deform or mischaracterise or conceal the essential functions and identity of the urban. At the same time, we have offered as background to our critique a concept of the city as a tangible phenomenon, distinct from but contained within society as a whole, and with specific genetic roots and unique internal organisational dynamics. This concept allows us to distinguish what is authentically urban from the merely contingently urban and hence to bring a degree of disciplined focus to the investigation of urban matters. We should add that precision of ideas in this respect is especially important in policy-relevant research (see Scott and Storper, 2015).

Against the backdrop of our own propositions about the nature of cities, we have examined three influential alternative views on urban matters, namely postcolonial theory, assemblage-theoretic approaches and planetary urbanism, and found them wanting. Postcolonial commentators argue for an approach to urban studies that is simultaneously provincial, comparativist and focused on difference, which in practice means particularity. While they invoke ambiguous notions of "worlding" they reject *as a matter of principle* the transfer of analytical results from cities of the Global North to cities of the Global South and by the same token any generalised theoretical concept of the urban, and presumably (at least for purists) any trans-provincial fertilisation of ideas.

Assemblage-theoretic approaches have much in common with these features, but in addition are intent on portraying social outcomes as relational kaleidoscopes in ever-changing combinatorial arrangements that offer few or no insights as to the genetics of indurated spatial and institutional arrangements. Not only are assemblage and actor-network approaches to the city notably weak in grasping fundamental social and economic processes, but they compound this weakness by suggesting that purely passive things lacking in intentionality and social discretion, like the door hinge mentioned by Acuto (2011), or the scallops studied by Callon (1984), or the railway ticket that enters into the account of poverty by AQ7 McFarlane (2011) are endowed with powers of agency akin to those of human subjects.

It should be noted that while postcolonial and assemblage-theoretic commentators have strong views about the conduct of urban research, none of them offers any coherent concept of the urban as such. Planetary urbanists for their part make strong claims about the deliquescence of the city as commonly understood and the assimilation of the urban into a world-wide space-economy. They provide little in the way of conceptual value-added by this manoeuvre while gratuitously deforming the usually accepted meaning of the term "urbanisation" pointing, as it does, to agglomeration, density and nodality and, by extension, to distinctive political, social, economic and identity-forming processes at the urban scale. Our own propositions regarding the material and relational structures of the urban land nexus suggest that the claims of planetary urbanists about the supposed waywardness of what they call methodological cityism and about the purely ideological status of the concept of the city are in the end seriously mistaken.

At least some of these aberrant tendencies in contemporary urban theory can be traced back to a remarkably uncritical faith among many contemporary analysts in the ability of abstracted philosophical ideas to orchestrate the shape and form of concrete investigations of cities. We are not opposed to incursions of philosophical ideas into the work of urban theorists; far from it. We are only too aware of how necessary philosophically-based criticality and clarity are to viable social analysis. Our concern here is focused primarily on what we take to be the unfortunate influence of post-structuralist philosophy in urban studies. We are referring here, first, to the semantically-inflated jargon that mars so much of the literature today. More importantly, and second, we also point to the overblown interpretative schemas that post-structuralism licenses and their tendency to crowd out analytically-oriented forms of social (and especially economic) enquiry in favour of a conceptually barren search for difference, particularity and localism. The ontologies of flatness favoured by post-structural theory are equally damaging to the vibrancy of urban studies especially in their denial of scalar dimensions to space in a manner that effectively dissolves the city away as a structured socio-geographic entity, and this encourages in turn a rampant eclecticism so that the city as such tends to shift persistently out of focus. Planetary urbanists are also at pains to secure this same dissolution, but this time on the basis of an enigmatic "epistemology" that in practice stands in for some rather unexceptional and, in our opinion, imperfectly digested observational statements. To repeat the message of our opening line, the current period of history can most certainly be characterised as an urban era, in the sense that more and more of humanity lives in distinctively urban settlements. If we are to come to some sort of understanding of the new and daunting challenges posed by this state of affairs (including a clear understanding of what is and is not ascribable to urban processes in modern life), we need an urban theory that is fully up to the task. We have tried to clear away some of the obstacles that we argue stand in the way of the accomplishment of this goal, and to propose some essential groundwork for building more effective theories of the urban and the urbanisation process.

Finally, we strongly advocate abandonment of the classification of cities in terms of a Global North and a Global South with its curious echo of First and Third Worlds. Postcolonial theorists, of course, have their own reasons for hewing to this terminology, and we ourselves certainly have no intention of suggesting that colonialism, even today, has not left deep traces on many different parts of the world and in many domains of human enquiry. That said, and in view of the prevailing, many-sided patchwork of spatial outcomes exhibiting many different empirical varieties of economic and political development in today's world, this schematic binary is quite definitely inadequate as an organisational framework for huge swaths of contemporary social investigation, and nowhere more so than in the case of urban studies.

Acknowledgements

We dedicate this paper to our late friend and colleague Edward Soja in the full knowledge that he would have been unable to resist expressing alternative perspectives on almost everything we say.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. In this paper, the term "cities" will generally be used to cover both small and large urban forms, including metropolitan areas and city-regions.
2. Much the same can be said for the term "determinism" that Mould (2015) invokes in criticism of our earlier paper. Any self-respecting determinist is likely to insist at a minimum that a deterministic approach involves the suppression of free will in favour of purely material or structural-functional causalities. Mould mobilises no reasoning or evidence as to how or why our theory of the urban land nexus involves any conception of this sort.
3. And now that opportunities for research on the cities of the Global South are expanding apace throughout the North and the South this relative neglect on the part of urban scholars will presumably fade rapidly away.
4. We use the term satirically.
5. A detailed response to Brenner and Schmid has been offered by Walker (2015). His main lines of critique differ from ours, but are equally adamant about the integrity of the city as an object of theoretical enquiry.
6. For example, among the many possible alternative terminologies are: "global space-economy", "planetary capitalism" or the "geographical anatomy of global society". These terminologies capture the spirit of what Brenner and Schmid seem to be saying, without obliterating the commonly received meaning of the term "urban".

References

Acuto M (2011) Putting ANTs into the *mille-feuille*. *City* 15: 552–562.

Allen RC (2009) *The British Industrial Revolution in Historical Perspective*. Cambridge: Cambridge University Press.

Amin A and Graham S (1997) The ordinary city. *Transactions of the Institute of British Geographers* 22: 411–429.

Amin A and Thrift N (2002) *Cities: Reimagining the Urban*. Cambridge: Polity.

Anderson JE and van Wincoop E (2004) Trade costs. *Journal of Economic Literature* 42: 691–751.

Angelo H and Wachsmuth D (2015) Urbanizing urban political economy: A critique of methodological cityism. *International Journal of Urban and Regional Research* 39: 16–27.

Aston T and Philpin CHE (1987) *The Brenner Debate: Agrarian Class Structure and Economic Development in Pre-Industrial Europe*. Cambridge: Cambridge University Press.

Aw T (2013) *Five Star Billionaire*. New York: Random House.

Bender T (2010) Postscript: Reassembling the city: Networks and urban imaginaries. In: Farías I and Bender T (eds) *Urban Assemblages: How Actor-Network Theory Changes Urban Theory*. London: Routledge, pp. 303–323.

Boo K (2012) *Behind the Beautiful Forevers: Life, Death and Hope in a Mumbai Undercity*. New York: Random House.

Brenner N and Schmid C (2014) The urban age in question. *International Journal of Urban and Regional Research* 38: 731–755.

Brenner N and Schmid C (2015) Towards a new epistemology of the urban. *City* 19: 151–182.

Brenner N, Madden DJ and Wachsmuth D (2011) Assemblage urbanism and the challenges of critical urban theory. *City* 15: 225–240.

Bunnell T and Maringanti A (2010) Practising urban and regional research beyond metrocentricity. *International Journal of Urban and Regional Research* 34: 415–420.

Callon M (1984) Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St Brieuc Bay. *Sociological Review* Special Supplement S1: 196–233.

Cheshire P and Hay D (1989) *Urban Problems in Western Europe: An Economic Analysis*. London: Hyman.

Chetty R, Hendren N, Kline P, et al. (2014) Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics* 129: 1553–1623.

Chibber V (2013) *Postcolonial Theory and the Spectre of Capitalism*. London: Verso.

Cole T (2014) *Every Day is for the Thief*. New York: Random House.

Comaroff J and Comaroff JL (2012) *Theory from the South: How Euro-America is Evolving toward Africa*. Boulder, CO: Paradigm.

Cooke P and Morgan K (1998) *The Associational Economy: Firms, Regions, and Innovation*. Oxford: Oxford University Press.

de Boek F and Plissart M-F (2004) *Kinshasa: Tales of the Invisible City*. Antwerp: Ludion.

DeLanda M (2002) *Intensive Science and Virtual Philosophy*. London: Continuum.

Deleuze G and Guattari F (1972) *Capitalisme et Schizophrénie*. Paris: Editions de Minuit.

Dick HW and Rimmer PJ (1998) Beyond the Third World city: The new urban geography of South-East Asia. *Urban Studies* 35: 2303–2321.

Dovey K (2012) Informal urbanism and complex adaptive assemblage. *International Development Planning Review* 34: 349–367.

Duranton G and Puga D (2004) Micro foundations of urban agglomeration economies. In: Henderson JV and Thisse JF (eds) *Handbook of Regional and Urban Economics, Vol. 4*. Amsterdam: Elsevier, pp. 2065–2118.

Duranton G and Storper M (2008) Rising trade costs? Agglomeration and trade with endogenous transaction costs. *Canadian Journal of Economics* 41: 292–319.

Edensor T and Jayne M (2012) Introduction: Urban theory beyond the West. In: Edensor T and Jayne M (eds) *Urban Theory Beyond the West: A World of Cities*. London: Routledge, pp. 1–27.

Farías I (2010) Introduction: Decentering the object of urban studies. In: Farías I and Bender T (eds) *Urban Assemblages: How Actor-Network Theory Changes Urban Theory*. London: Routledge, pp. 1–24.

Farías I (2011) The politics of urban assemblages. *City* 15: 365–374.

Fujita M and Thisse J-F (2002) *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge: Cambridge University Press.

Fujita M, Krugman P and Venables AJ (1999) *The Spatial Economy: Cities, Regions and International Trade*. Cambridge, MA: MIT Press.

Glaeser EL (2011) *Triumph of the City*. London: Macmillan.

Graham S and Marvin S (2001) *Splintering Urbanism: Networked Infrastructures, Technological Mobilities, and the Urban Condition*. London: Routledge.

Habermas J (1971) *Knowledge and Human Interests*. Boston, MA: Beacon.

Hall P (1998) *Cities in Civilization*. New York: Pantheon.

Haraway D (1988) Situated knowledges. *Feminist Studies* 14: 575–599.

Hummels D (2007) Transportation costs and international trade in the second era of

globalization. *Journal of Economic Perspectives* 21: 131–154.
Kerr W and Kominers SD (2015) Agglomerative forces and cluster shapes. *Journal of Economics and Statistics* 97(4): 877–899.
Krugman P (1991) *Geography and Trade*. Leuven: Leuven University Press.
Kuhn TS (1962) *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford: Oxford University Press.
Lefebvre H (1970) *La Révolution Urbaine*. Paris: Gallimard.
Leitner H and Sheppard E (2015) Provincializing critical urban theory: Extending the ecosystem of possibilities. *International Journal of Urban and Regional Research*. Epub ahead of print. DOI: 10.1111/1468-2427.12277.
Livingstone DN (2014) *Dealing with Darwin: Place, Politics, and Rhetoric in Religious Engagements with Evolution*. Baltimore, MD: Johns Hopkins University Press.
Mannheim K (1952) *Essays in the Sociology of Knowledge*. Henley-on-Thames: Routledge and Kegan Paul.
Marston SA, Jones JP and Woodward K (2005) Human geography without scale. *Transactions of the Institute of British Geographers* 30: 416–432.
Mbembe A and Nuttall S (2004) Writing the world from an African metropolis. *Public Culture* 16: 347–372.
McFarlane C (2011a) Assemblage and critical urban theory. *City* 15: 204–224.
McFarlane C (2011b) *Learning the City: Knowledge and Translocal Assemblage*. Oxford: Wiley-Blackwell.
McFarlane C (2011c) On context: Assemblage, political economy and structure. *City* 15: 375–388.
Merrifield A (2013) The urban question under planetary capitalism. *International Journal of Urban and Regional Research* 37: 909–922.
Molotch H (1976) The city as a growth machine: The political economy of place. *American Journal of Sociology* 82: 309–332.
Mould O (2015) A limitless urban theory? A response to Scott and Storper's 'The nature of cities: the scope and limits of urban theory'. *International Journal of Urban and Regional Research*. Epub ahead of print. DOI: 10.1111/1468-2427.12288.
Mukhija V and Loukaitou-Sideris A (eds) (2014) *The Informal American City*. Cambridge, MA: MIT Press.
Mukim M (2015) Coagglomeration of formal and informal industry: Evidence from India. *Journal of Economic Geography* 15: 329–351.
Myers G (2014) From expected to unexpected comparisons. *Singapore Journal of Tropical Geography* 35: 104–118.
Ong A and Roy A (eds) (2011) *Worlding Cities: Asian Experiments in the Art of Being Global*. Oxford: Wiley-Blackwell.
Pagden A (2013) *The Enlightenment and Why it Still Matters*. New York: Random House.
Patel S (2014) Is there a 'south' perspective to urban studies? In: Parnell S and Oldfield S (eds) *The Routledge Handbook on the Cities of the Global South*. London: Routledge, pp. 37–53.
Peck J (2015) Cities beyond compare. *Regional Studies* 49: 160–182.
Ren J and Luger J (2015) Comparative urbanism and the Asian city: Implications for research and theory. *International Journal of Urban and Regional Research* 39: 145–156.
Rey P, Duroux Y and Bettelheim C (1971) *Sur l'Articulation des Modes de Production*. Paris: Ecole Pratique des Hautes Etudes, Centre d'Etudes de Planification Socialiste.
Robinson J (2006) *The Ordinary City: Between Modernity and Development*. London: Routledge.
Robinson J (2011) Cities in a world of cities: The comparative gesture. *International Journal of Urban and Regional Research* 35: 1–23.
Robinson J (2014) New geographies of theorizing the urban: Putting comparison to work for global urban studies. In: Parnell S and Oldfield S (eds) *The Routledge Handbook on Cities of the Global South*. London: Routledge, pp. 57–70.
Robinson J and Roy A (2015) Global urbanisms and the nature of urban theory. *International Journal of Urban and Regional Research*. Epub ahead of print. DOI: 10.1111/1468-2427.12272.

Roweis ST and Scott AJ (1977) Urban planning in theory and practice: A re-appraisal. *Environment and Planning* 9: 1097–1119.

Roy A (2005) Urban informality: Toward an epistemology of planning. *Journal of the American Planning Association* 71: 147–158.

Roy A (2009) The 21st-century metropolis: New geographies of theory. *Regional Studies* 43: 819–830.

Roy A (2011) Slumdog cities: Rethinking subaltern urbanism. *International Journal of Urban and Regional Research* 35: 223–238.

AQ8 ~~Roy A (2014) Slum-free cities of the Asian century: Postcolonial government and the project of inclusive growth. *Singapore Journal of Tropical Geography* 35: 136–150.~~

Roy A (2015) Who's afraid of postcolonial theory? *International Journal of Urban and Regional Research*. Epub ahead of print. DOI: 10.1111/1468-2427.12274.

Said EW (1978) *Orientalism*. New York: Pantheon.

Sampson RJ (2012) *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago, IL: University of Chicago Press.

Sayer A (2004) Foreword: Why critical realism? In: Fleetwood S and Ackroyd S (eds) *Critical Realist applications in Organisation and Management Studies*. London: Routledge, pp. 6–20.

Scott AJ (1980) *The Urban Land Nexus and the State*. London: Pion.

Scott AJ (2012) *A World in Emergence: Cities and Regions in the 21st Century*. Cheltenham: Edward Elgar.

Scott AJ and Storper M (2015) The nature of cities: The scope and limits of urban theory. *International Journal of Urban Theory* 39: 1–15.

Sheppard E (2014) Globalizing capitalism and southern urbanization. In: Parnell S and Oldfield S (eds) *The Routledge Handbook on Cities of the Global South*. London: Routledge. AQ9

Sheppard E, Leitner H and Maringanti A (2013) Provincializing global urbanism: A manifesto. *Urban Geography* 34: 893–900.

Simone A (2011) The surfacing of modern life. *City* 15: 355–364.

Simone A (2014) *Jakarta, Drawing the City Near*. Minneapolis, MN: University of Minnesota Press.

Smith RG (2013) The ordinary city trap. *Environment and Planning A* 45: 2290–2304.

Soja E and Scott AJ (1986) Los Angeles: Capital of the late twentieth century. *Environment and Planning D: Society and Space* 4: 249–254.

Spivak GC (2008) Can the subaltern speak? In: Nelson C and Grossberg L (eds) *Marxism and the Interpretation of Culture*. Basingstoke: Macmillan, pp. 271–313.

Standing G (2011) *The Precariat: The New Dangerous Class*. London: Bloomsbury.

Storper M (2013) *Keys to the City: How Economics, Institutions Social Interaction and Politics Shape Developnent*. Princeton, NJ: Princeton University Press.

Taylor PJ (2013) *Extraordinary Cities: Millennia of Moral Syndromes World-Systems and City/State Relations*. Cheltenham: Elgar.

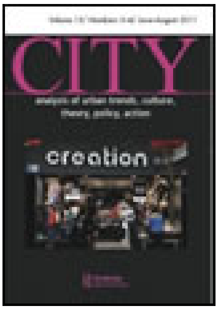
Tonkiss F (2011) Template urbanism: Four points about assemblage. *City* 15: 584–588.

Walker R (2015) Building a better theory of the urban: A response to 'Towards a New Epistemology of the Urban?' *City* 19: 183–191.

Walker R (2016) Why cities? A response. *International Journal of Urban and Regional Research*. AQ10

Wilson WJ (1987) *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago, IL: University of Chicago Press.

Wirth L (1938) Urbanism as a way of life. *American Journal of Sociology* 44: 1–24.



City: analysis of urban trends, culture, theory, policy, action
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ccit20

Emerging cities of the third wave

Allen J. Scott [a]

[a] Department of Geography and Department of Public Policy, University of California-Los Angeles

Available online: 30 Aug 2011

To cite this article: Allen J. Scott (2011): Emerging cities of the third wave, City: analysis of urban trends, culture, theory, policy, action, 15:3-4, 289-321

To link to this article: http://dx.doi.org/10.1080/13604813.2011.595569

CITY, VOL. 15, NOS. 3–4, JUNE–AUGUST 2011

Routledge
Taylor & Francis Group

Emerging cities of the third wave

Allen J. Scott, with images and captions by Elvin Wyly

I argue that three distinctive waves of urbanization can be recognized, each of them associated with a major historical phase of capitalist development. The leading edges of capitalism today can be typified in terms of a basic cognitive–cultural system of production that is transforming the economic foundations of many large metropolitan areas all over the world. This turn of events is evident in two further aspects of urbanization processes at the present time. First, a new division of labor is strongly under way with major implications for the restratification of urban labor markets and urban social life. Second, the economic and social transformations currently evident in large urban areas are provoking significant changes in the physical milieu and built form of the city, from gentrification to what I call aestheticized land use intensification. I attempt to synthesize important elements of the discussion by means of a disquisition on the city and the world, in which I point to some of the more outstanding institutional failures within the current system of neoliberal local–global development.

Key words: city-regions, cognitive–cultural economy, gentrification postfordism, urban economy, urbanization

1. Preamble: the city as a social phenomenon in capitalism

One of the primary features of urbanization processes in capitalism is that they are constituted, in part, as spatial accretions of economic activity. As such, cities in capitalism can always be represented in significant ways as dense agglomerations of capital and labor. An important corollary of this claim is that cities are both dependent outcomes of capitalist dynamics (for they reflect the search on the part of investors and producers for locational configurations that underpin their profitability) and way-stations in the social reproduction of capitalism (for they also provide critical foundations for production, work and life at high levels of complexity). Individual cities are further marked by a mass of emergent but reflexive effects that are incorporated in diverse ways into the recursive dynamics of capitalist society and urban development. The specifically intra-urban logic of this heterogeneous accumulation of phenomena is to be found in their mode of functional integration via the internal spatial structure of the city (Scott, 2008a).

In these senses, all cities in capitalism have certain generic qualities, though there are always marked species differences between selected groups of cities over space and time, depending partially on variations in the types of capitalism with which they are associated. My objective in this paper is, first of all, to argue that a major new phase of capitalist development has been in the course of formation since the 1980s, and second of all, to work out some

ISSN 1360-4813 print/ISSN 1470-3629 online/11/03–40289–33 © 2011 Taylor & Francis
DOI: 10.1080/13604813.2011.595569

Figure 1 Ford River Rouge plant, Dearborn, Michigan, July 2010. By 1910, Ford's Highland Park factory had begun the transformations that would eventually inspire Gramsci (1971, p. 280) to ask whether fordism might be 'the ultimate stage in the process of progressive attempts by industry to overcome the law of the tendency of the rate of profit to fall'. Almost immediately, however, Ford set his sights on a newer and larger innovation that combined horizontal one-story layout with intensified vertical integration in supply-chain management. At its peak, the two-mile long River Rouge plant employed more than 100,000 workers, but automation cut this figure by two-thirds by 1960 (Sugrue, 1995, p. 117).

elements of a systematic description of concomitant urbanization processes. Two of the distinguishing features of the cities that are most representative of this emerging phase are that they are cynosures of the so-called 'new' economy and that they function as key nodes in an intensifying network of relationships that is nothing less than global in extent.

2. Capitalism and urbanization: three waves?

In very schematic terms, we might recognize three broad historical phases of capitalist development, based—as regulation theory suggests (cf. Benko and Lipietz, 1995; Boyer, 1986; Talha, 1995)—on their characteristic technologies, leading sectors, labor market relations and competitive dynamics. Equally schematically, we can also distinguish three corresponding waves of urbanization.[1] It is important to bear in mind that I am not referring here to phases or waves that encompass all actually realized instances of capitalism and urbanization, but rather to episodic forms that appear to have some fairly wide paradigmatic value. One major episode can be identified as the 19th-century factory and workshop system, with its most advanced urban expression occurring in the burgeoning manufacturing towns of Britain at that time. Another coincides with the fordist mass production system that reached its highest expression in the large

Figure 2 Cleveland, Ohio, July 2010
Half a century ago, the lower reaches of the Cuyahoga River were choked with oily water, smoky skies, etc. and lots of workers. Today, the post-fordist city is clean, and emptying out. The city's population is half the peak attained in the early 1950s, and 'the quintessential blue-collar American city' has been 'thoroughly restructured around the prerequisites of post-fordism', with growth in a small high-technology sector, producer services and 'cultural consumption' (Warf and Holly, 1997, p. 208). Yet in the current crisis, a distressingly large portion of the city's economic base seems tied in one way or another to foreclosures and dispossessions (Saunig, 2010). The World Press Photograph of the Year for 2008 showed a Cuyahoga County sheriff's deputy moving through a ransacked home in Cleveland after foreclosure, his gun drawn in a sweep for squatters or (foreclosed) residents. US Treasury Secretary Hank Paulson's famous metaphorical 'bazooka' (2010) failed to reassure world financial markets in the summer of 2008, but the threat of evictions by sheriffs with handguns remains ruthlessly effective in the current era of accumulation by dispossession.

metropolitan areas of the American Manufacturing Belt over the middle decades of the 20th century (Figures 1 and 2). Yet another major episode of capitalist development and urbanization starts to emerge after the crisis of fordism in the 1970s. This episode is often characterized by reference to the term 'post-fordism' (cf. Amin, 1994), or more recently to terms such as the 'new economy' (Beyers, 2003), the 'knowledge economy' (e.g. Cooke and Piccaluga, 2006), the 'creative economy' (e.g. Franke and Verhagen, 2005), 'cognitive capitalism' (e.g. Moulier Boutang, 2007), 'cognitive–cultural capitalism' (e.g. Scott, 2008b), and so on. This still unfolding episode is associated above all with the great global city-regions, not only of the more developed North, but also and increasingly of the less developed South, though many smaller cities also participate in its logic in different ways.

Each of the three waves of urbanization alluded to above can be characterized by a sort of ideal-type city representing a notional condensation of the dominant functional systems that constitute intra-urban space as structured by prevailing capitalist social and property relations. Each, in other words, can be represented by an imaginary city with a particular kind of economic base, a

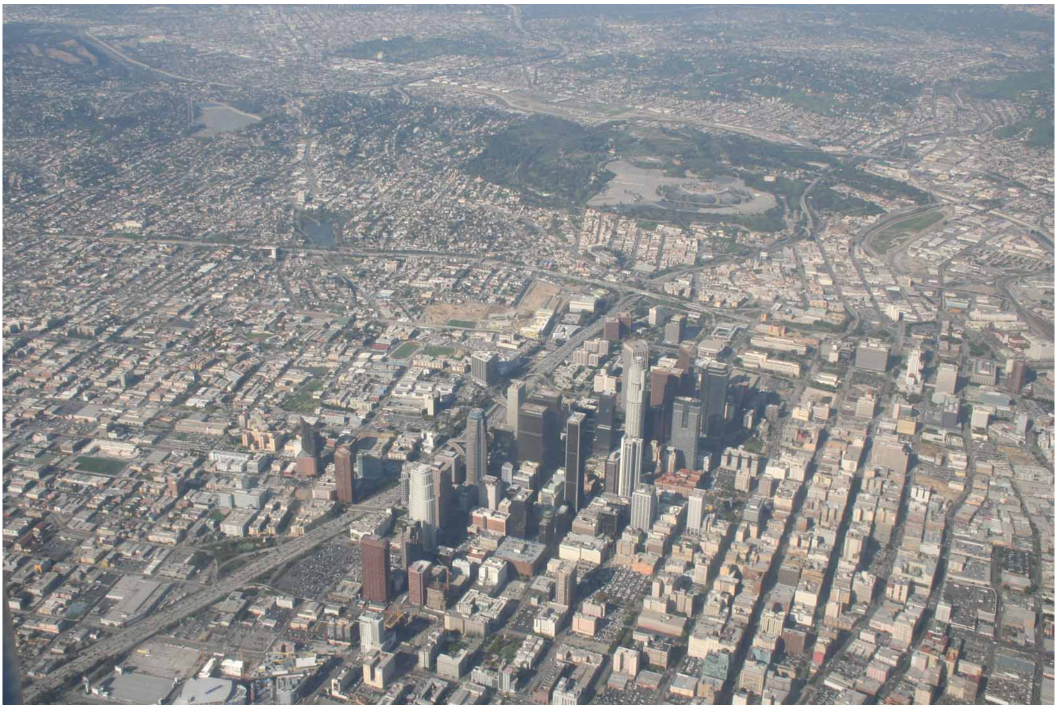


Figure 3 Los Angeles, once proclaimed as *the* paradigmatic city of post-fordism
'A simplified comparative urban epistemology for the past half-century', Dear (2005, pp. 250–251) writes, begins with 'a *scientific/quantitative approach*, which emerged in the 1960s, focused on the industrial city, with a positivist, Chicago-inspired modernism as its foundation.' A generalized 'Marxian modernism' less concerned with specific cities was followed by 'a postmodern urbanism … inspired primarily by Los Angeles'. The LA School 'mounted a sustained assault' on the defenses of positivist, modernist urban theory of the industrial age (Dear *et al.*, 2008, p. 102). The efforts of Dear and others to create a radical break in urban theory never quite achieved a consensus that Los Angeles was Chicago's successor as urban-theoretical crucible. But the force of the Angelistas' challenge shook the foundations of existing urban theory and thus made urban studies more receptive to cosmopolitan analysis of an assertively polycentric, polyvocal and urbanizing capitalism (Robinson, 2006).

specific division of labor, a definite spatial pattern of neighborhoods, and so on, together with a set of problems and predicaments that shape the city's policy and planning agendas. Of course, no precisely corresponding city will ever be found in empirical reality. The best we can do no doubt is to pick illustrative but invariably debatable cases, such as, say, Manchester in 19th-century England, Chicago and Detroit toward the middle of the 20th century, or Los Angeles at the turn of the 20th century (Figure 3). The so-called LA School of urban analysis that briefly flourished in the 1980s and 1990s (Cenzatti, 1993; Scott and Soja, 1996; Soja and Scott, 1986) proclaimed Los Angeles as *the* paradigmatic city of post-fordism, though the idea encountered considerable opposition at the time and still does today (cf. Dear *et al.*, 2008). Even so, and despite its flaws, the work of the LA School represents a sort of premonition of the kinds of theoretical research and ideological critique that, as I shall argue here, are called for as the third wave of urbanization begins to move beyond its incipient stages of formation.

With the wisdom of hindsight it is now clear that Los Angeles—and Southern California more generally—is only one of a number of cities that have shifted rather decisively beyond the wave of urbanization that coincided with fordism. Again, there can be

Table 1 The 75 worldwide centers of commerce as defined by MasterCard Worldwide (2008). Names of cities lying in peripheral and formerly peripheral areas of the world system are set in bold

Rank	City	Country	Index value
1	London	UK	79.17
2	New York	USA	72.77
3	Tokyo	Japan	66.60
4	**Singapore**	Singapore	66.16
5	Chicago	USA	65.24
6	**Hong Kong**	Hong Kong	63.94
7	Paris	France	63.87
8	Frankfurt	Germany	62.34
9	Seoul	South Korea	61.83
10	Amsterdam	Netherlands	60.06
11	Madrid	Spain	58.34
12	Sydney	Australia	58.33
13	Toronto	Canada	58.16
14	Copenhagen	Denmark	57.99
15	Zurich	Switzerland	56.86
16	Stockholm	Sweden	56.67
17	Los Angeles	USA	55.73
18	Philadelphia	USA	55.55
19	Osaka	Japan	54.94
20	Milan	Italy	54.73
21	Boston	USA	54.10
22	**Taipei**	Taiwan	53.32
23	Berlin	Germany	53.22
24	**Shanghai**	China	52.89
25	Atlanta	USA	52.86
26	Vienna	Austria	52.52
27	Munich	Germany	52.52
28	San Francisco	USA	52.39
29	Miami	USA	52.33
30	Brussels	Belgium	52.16
31	Dublin	Ireland	51.77
32	Montreal	Canada	51.60
33	Hamburg	Germany	51.53
34	Houston	USA	51.30
35	Dallas	USA	51.25
36	Washington, DC	USA	51.19
37	Vancouver	Canada	51.10
38	Barcelona	Spain	50.90
39	Dusseldorf	Germany	50.42
40	Geneva	Switzerland	50.13
41	Melbourne	Australia	49.93
42	**Bangkok**	Thailand	48.23
43	Edinburgh	UK	47.79
44	**Dubai**	United Arab Emirates	47.23
45	Tel Aviv	Israel	46.50
46	Lisbon	Portugal	46.46
47	Rome	Italy	45.99
48	**Mumbai**	India	45.71
49	Prague	Czech Republic	45.50
50	**Kuala Lumpur**	Malaysia	45.28
51	Moscow	Russia	44.99
52	Budapest	Hungary	44.52

(*Continued*)

Table 1. Continued

Rank	City	Country	Index value
53	**Santiago**	Chile	44.49
54	**Mexico City**	Mexico	43.33
55	Athens	Greece	43.25
56	**São Paulo**	Brazil	42.70
57	**Beijing**	China	42.52
58	**Johannesburg**	South Africa	42.04
59	Warsaw	Poland	41.26
60	**Shenzhen**	China	40.04
61	**New Delhi**	India	39.22
62	**Bogotá**	Colombia	38.27
63	**Buenos Aires**	Argentina	37.76
64	**Istanbul**	Turkey	36.14
65	**Rio de Janeiro**	Brazil	35.91
66	**Bangalore**	India	35.78
67	St Petersburg	Russia	35.55
68	**Jakarta**	Indonesia	35.40
69	**Riyadh**	Saudi Arabia	35.37
70	**Cairo**	Egypt	35.29
71	**Manila**	Philippines	35.15
72	**Chengdu**	China	33.84
73	**Chongqing**	China	33.13
74	**Beirut**	Lebanon	31.81
75	**Caracas**	Venezuela	26.11

no unambiguous statement about which cities constitute the essential core of this new wave or about their common characteristics, but there is one helpful inventory of cities that seems to capture at least some of what is at stake in the present investigation. Here, I am referring to the 75 cities identified in the MasterCard Worldwide Centers of Commerce Index for 2008 (MasterCard Worldwide, 2008). Like any of the competing city rankings currently available this one has its deficiencies, but it has the peculiar merit for present purposes of reflecting many of the preconceptions among global business elites about what constitutes a successful metropolis in the early 21st century. It includes a reasonably large number of cities ranked not only in relation to their economic environments but also in relation to such matters as their legal and political frameworks, their role in knowledge creation and their livability.[2] This particular ranking is probably about as good a representation of a first-cut geography of third wave urbanization as we are likely to get at the present time. The full set of 75 cities and their aggregate scores are set forth in Table 1. Not surprisingly, the cities of the traditional core capitalist countries are clearly dominant, with London, New York and Tokyo occupying first, second and third places, respectively. Perhaps even more significant is the strong presence of cities from East and Southeast Asia, and notably Singapore (Figure 4), Hong Kong (Figure 5), Seoul, Taipei and Shanghai. In addition as we scan through the bottom half of the rankings a large number of cities from other parts of Asia start to make their appearance, together with cities in the Middle East and Latin America. With the exception of Cairo and Johannesburg, cities in Africa are notably absent.

The cities mentioned in Table 1 are obviously all very different from one another in terms of history and geographic situation, and as the wide numerical variation of their scores suggests they are far from

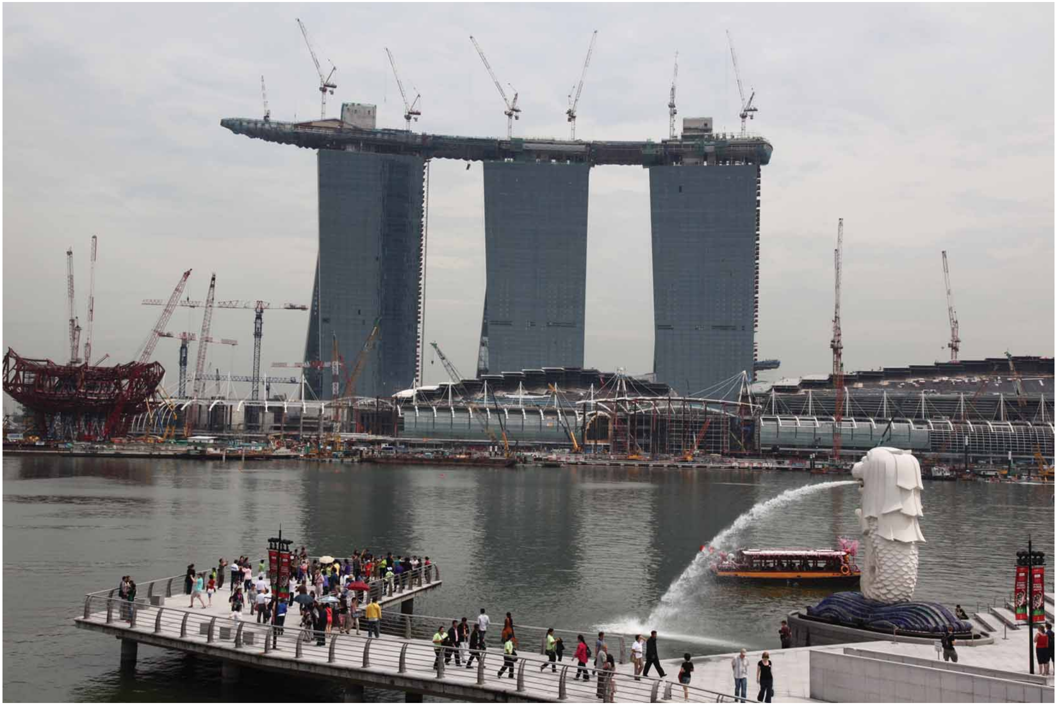


Figure 4 Marina Bay Sands, Singapore, January 2010
Marina Bay Sands opened in mid-2010, with 2500 rooms and suites, 'the hottest night clubs', a 'Las Vegas-style casino' and a convention center facility with capacity for 45,000 delegates. The Sands proclaims that it 'seamlessly combines business and leisure into a singular destination like no other' (Adelson, 2011).

participating equally in what I have provisionally identified as the third wave of urbanization. Even those with the top scores are more accurately thought of as still emerging than as fully emerged cities of the third wave. Moreover, some of the cities toward the bottom of the rankings can at best be only marginally included in the third wave, and many of these cities (especially in the global South) also contain large impoverished populations that remain far outside of any wider capitalist reality (Roy, 2011). My argument in what follows is that since about the 1980s there has been a widespread but by no means ubiquitous shift in capitalism in the direction of a global cognitive–cultural economy, with correspondingly profound impacts on urbanization processes. The cities listed in Table 1 are simply meant to suggest some of the places around the world where elements of this shift are definitely if irregularly discernible on sections of the urban landscape.

3. Economy and urbanization in the third wave

Shifting economic winds

As the crisis of fordism deepened over the 1970s, geographers and regional scientists began to produce deeply pessimistic accounts of the corresponding quandaries that were then proliferating in the principal cities and regions of that period (Figure 6) (Blackaby, 1978; Bluestone and Harrison, 1982; Carney *et al.*, 1980; Massey and Meegan, 1982). Even as fordism was approaching its climacteric, however, a number of unobtrusive albeit significant counter-tendencies were already in the air, and these can now be

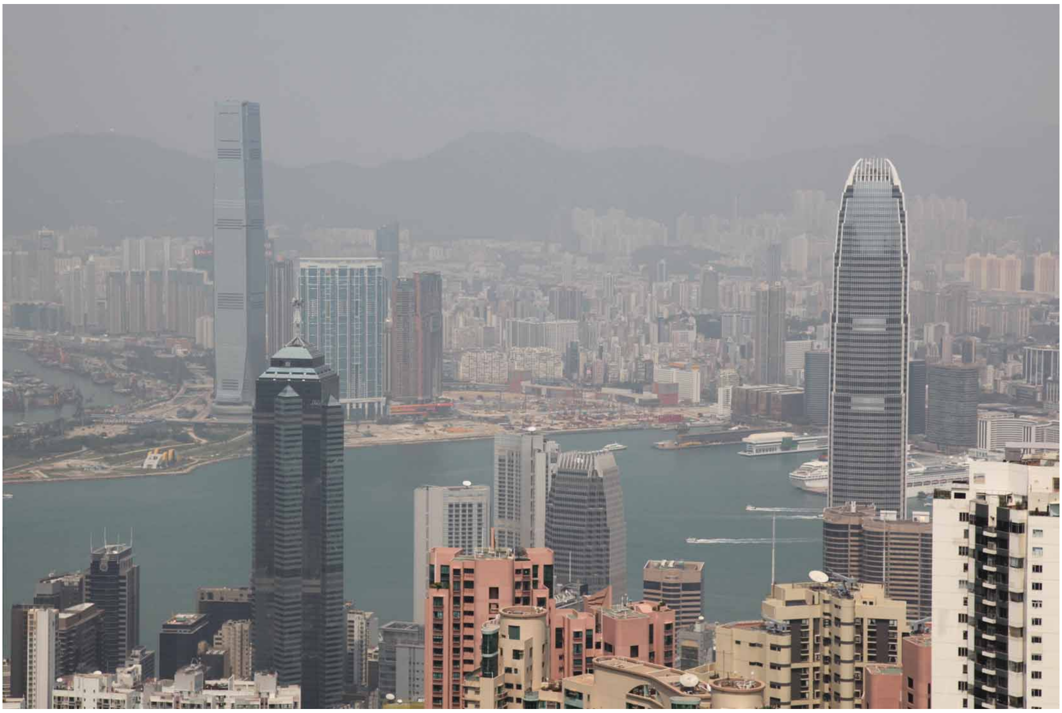


Figure 5 Hong Kong, March 2010
Above and beyond its increasingly prominent role as a global financial center, Hong Kong is part of an emerging network of East Asian 'creative industries … with aspirations to develop media capitals' (Flew, 2010, p. 42). In Hong Kong, Singapore, Shanghai and other leading centers, 'media space is developing its own dynamics of capital accumulation, creative migration, and socio-cultural variation … whose dynamism exceeds that of Hollywood' (Flew, 2010, p. 44).

seen as early anticipations of the later more resolute turn to a new economy with very intensive requirements in regard to cognitive and cultural labor. In the so-called Third Italy, a revival of small-scale craft industries was starting to become apparent in a number of places (Bagnasco, 1977). In the Sunbelt of the USA, several high-technology industrial agglomerations were moving rapidly ahead, with Silicon Valley in the lead (Rogers and Larson, 1984; Saxenian, 1983). The 'service economy', too, was now starting to be recognized as an important segment of the capitalist system, most notably in the major metropolitan areas of the more economically advanced countries (Fuchs, 1968; Stanback, 1979).

By the 1980s, these sorts of phenomena were increasingly being categorized under the rubric of 'post-fordism'. Along with this theoretical advocacy came a number of associated concepts about production and economic geography, such as flexible specialization (Piore and Sabel, 1984), flexible accumulation (Harvey, 1987) and new industrial spaces (Scott, 1988), together with a revival of the Marshallian concept of the industrial district (Becattini, 1987; Garofoli, 1987). These ideas are still widely invoked today, and there can be no doubt that the notion of post-fordism has played an important role in urban and regional analysis generally over the last three decades. With the continued advance of socio-spatial theory, however, the time has evidently come for an honorable retirement of the term. In particular, the expression 'post-fordism' explicitly alludes only to what it is not, and a more affirmative designation is obviously desirable. Accordingly, in this paper, I shall use

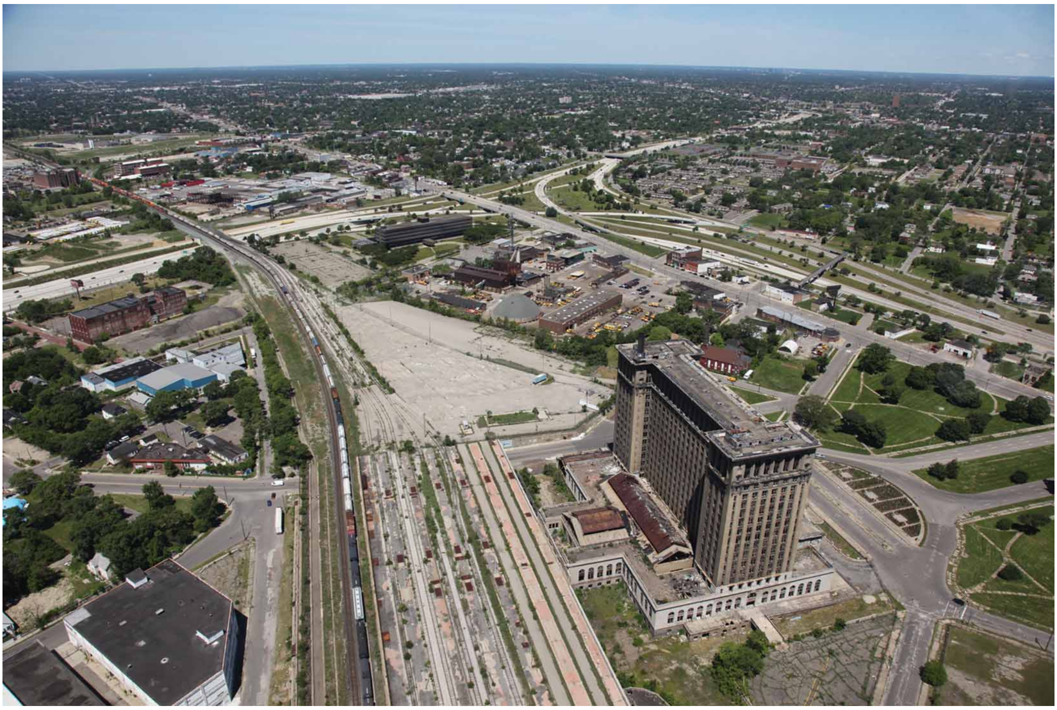


Figure 6 Abandoned Michigan Central Station, Detroit, Michigan, July 2010
The Michigan Central Station, opened in 1913, symbolized the optimism of urban growth at the birth of fordism: the station was situated two miles west of the CBD in anticipation of slum clearance 'as downtown inevitably expanded westward' (Kavanaugh, 2001, p. 1). Expansion never came, and slum clearance only exacerbated the decline of the fordist city. Planning failures that doomed the majestic station foreshadowed the broader crisis of fordist urbanism, making it 'important to see the Michigan Central as not only an exquisite structure, but also a metaphor for the entire city of Detroit' (Kavanaugh, 2001, p. 1). The train sheds on the left side were dismantled in 2000 to make way for temporary transfer yards for the Canadian Pacific Railway in the post-NAFTA expansion of cross-border trade. Now, more than two decades after the departure of the last train, the building serves two key roles in Detroit's putative position in the emergent cognitive–cultural economy: as a film-shoot for movies, television shows and documentaries, and as an attraction for 'ruingazers'. Detroit's mid-20th-century prosperity, segregation and automobility drove a dramatic decentralization, and now 'the footprint of this racialized fordist urbanism is the structural foundation of the city's forlorn appearance . . .' (Steinmetz, 2008, p. 231). Most of the 'ruingazers' are 'white suburbanites who left the city or whose parents fled a generation or two ago', chasing a nostalgia for the upward class mobility—and outward residential mobility—provided by the fordist metropolis (Steinmetz, 2008, p. 217).

the terms 'cognitive–cultural capitalism', or the 'cognitive–cultural economy', or sometimes just the shorthand 'new economy', as a way of capturing something of the wider import of capitalism in the current conjuncture and its effects on urbanization processes. Three major points must now be dealt with.

First, while classical fordism was founded on large-scale electro-mechanical technologies, capitalism today has forged ahead on the basis of digital methods of calculation, communication and information storage, with profound repercussions on the organization of production and work. These methods supplement and galvanize many non-routine forms of labor, and they have stimulated an enormous expansion of jobs that call upon the diverse cognitive and cultural capacities of workers. Employment of workers endowed with these kinds of human capital has in fact been growing rapidly over the last few decades in the American economy (Autor *et al.*, 2003). Concomitantly, routine work of all kinds is being steadily eliminated from the employment system, and this remark is as true for office work as it is for manual work in the manufacturing sector. Even where blocks of routine work continue to resist assimilation into digital technology systems they are being steadily shunted out of high-wage areas, and particularly out of large metropolitan areas, to peripheral labor depots with significantly lower wage levels.

Second, and in line with these same shifts, the overall division of labor in the more economically advanced centers of contemporary capitalism is undergoing far-reaching qualitative transformation as compared to the situation in fordism (Levy and Murnane, 2004). In contrast to the old white-collar/blue-collar split that emerged out of the logic of fordism, we are now seeing a rather different division of labor comprising at its extremes two very much more sharply differentiated strata, in terms of both occupational structure and income. The upper echelon of this division is made up of what some analysts have referred to as 'symbolic workers' (Reich, 1992), the 'cognitariat' (Moulier Boutang, 2007; Vercellone, 2007), 'knowledge workers' (Kunzmann, 2009), the 'creative class' (Florida, 2002), or what I designate here as cognitive–cultural workers. The lower echelon coincides to a significant degree with a 'new servile class', an identification that reflects the decline of blue-collar manufacturing work in major metropolitan areas and the expansion of low-wage service-oriented functions.[3] Even so, and in spite of their marginal social and economic status, the members of the lower echelon are far from being devoid of cognitive and cultural capacities, though they may be devoid of formal qualifications. Rather, much of their work—from child-minding through janitorial work to waiting on tables and taxi driving—involves considerable amounts of discretionary decision-making and many informal skills.

Third, as Engels' Law suggests, consumption over and above that which is required for simple subsistence tends to increase disproportionately with rising incomes. In fordism, this phenomenon gave rise to what we might label 'consumer society mark I', in which households accumulated masses of relatively standardized material products. Today, we are moving into 'consumer society mark II' in which households expend more and more of their budgets not only on material products, now relatively destandardized, but also on a large and constantly varying palette of goods and services that, for better or worse, have potent experiential significance (Pine and Gilmore, 1999).

These diverse trends are associated with a number of striking changes in the core sectors of the capitalist system. To be sure, mass production has never actually disappeared from the system, and much contemporary output is still based on assembly-line methods, though usually—thanks to computerization—with a much greater degree of flexibility than half a century ago. More to the point for present purposes has been the rise to centrality within the system of a number of new and revivified economic segments based above all on the digital technologies and finely honed cognitive and cultural forms of human capital (including handicraft skills) already alluded to. One of these segments comprises high-technology industry, which not only produces outputs like aerospace equipment, biotechnology, pharmaceuticals, and so on, but is also the source of the driving technological instruments of the new economy. A second segment revolves around advanced service activities with their main products ranging from financial and business services to medical and personal services. A third is made up of a variety of media, cultural and fashion industries with outputs like television programs, film, music, electronic games, architecture, clothing, and so on. The outputs of this third segment overlap with a diversity of hybrid cultural-*cum*-utilitarian products, such as shoes, handbags, furniture or even cars. All three of these segments are marked by rising levels of economic competition, particularly as globalization and neoliberalism continue to extend

Figure 7 Empty Packard plant, Detroit, Michigan, July 2010
In the 1940s, this was 'the heartbeat of the industrial metropolis', amidst 'one of the most remarkable concentrations of industry in the United States' (Sugrue, 2005, p. 125). Near this area was a Dodge factory that employed 35,000 workers, and at Packard's mile-long, 95-building complex 'at shift change time . . . cars, buses, and pedestrians clogged the streets' (Sugrue, 2005, p. 125). Packard was slow in readjusting to civilian production after the lucrative wartime military contracts (Ward, 1995). With the exception of temporary, transitional and squatter uses, most of the complex has been empty for decades.

Marge Piercy (2010) puts it best:
'I remember Detroit when it hummed
with factories like an army of bees
all day, all night . . .
. . .
. . . Who sucked
the money out? Who sent the jobs
overseas to gut the unions?
They live out in the more distant
Suburbs where we never went
except as maids, way past
the end of the bus lines.

Invisible unless we appear
where we don't belong among
decisionmakers who view us
as beasts of burden. They are
persons; we are only numbers.'

their reach. However, competition in this new economy is far from resembling classical laissez-faire. On the contrary, and precisely because of the underlying flexibility of digitized and handicraft production methods and the resulting high capacity of firms to differentiate their products in terms of both performance and appearance, markets at every geographic scale today are much more likely to be characterized by monopolistic competition *à la* Chamberlin (1933), that is, competition based on producer-specific and place-specific product qualities rather than uniquely on price.

Urban economic responses

Over much of the 20th century fordist machinofacture displayed a distinct proclivity to locational agglomeration and hence contributed greatly to urban growth, especially in the great manufacturing belts of North America and Western Europe. As it happens, fordism was also associated with a degree of locational dispersal, but for much of the 20th century this was subordinate to more powerful centripetal forces. Then, as the crisis of fordism moved into full spate in the 1970s, this process of dispersal accelerated greatly, and many cities succumbed to stagnation and decline as a result (Figure 7). Subsequently, over the 1980s and 1990s, trends to locational re-agglomeration have re-asserted themselves as a more cognitively and culturally inflected form of capitalism has come to prominence. These trends have been responsible for new rounds of urban growth in places within and beyond the traditional geographic range of fordist capitalism, including both the global South and former state socialist societies.

An extended corpus of literature is devoted to issues of agglomeration (e.g. Becattini, 1987; Cooke and Morgan, 1998; Scott, 1988; Storper, 1997), and there is accordingly little need here to provide a technical account as to why and how this phenomenon is once more on the upsurge. We do, however, need to recall that agglomerated production systems typically emerge out of three main lines of force, namely, (a) dense networks of specialized but complementary producers forming functional complexes of interrelated firms; (b) large and many-sided local labor markets; and (c) learning and innovation processes based on the grids of socio-economic interaction that characterize these kinds of systems, especially when tacit knowledge effects are present (Asheim and Coenen, 2005; Gertler, 2003). Virtually all of the major segments of the cognitive–cultural economy—above all, high-technology production, business and financial services, and the cultural industries—embody various combinations of these lines of force, and they are very much in the vanguard of re-agglomeration processes today. Processes like these engender what we might refer to as proto-urban forms, that is, economic spaces comprising specialized industrial districts and zones and related outlying production foci. These forms take on a fully fledged urban character as contingent socio-political phenomena unfold and act reflexively back upon the local production system to create intra-urban space in its full sweep, including the capital-intensive infrastructures that further augment agglomeration economies. The corresponding geography of the city is hence of great complexity, especially where multiple industrial districts occur within the same urban area. That said, the intra-urban geography of the three major segments of the cognitive–cultural economy as enumerated earlier is often expressed in concrete patterns that recur with some regularity from one city to another. Thus, high-level financial activities typically gravitate to core central business district locations; cultural industries are found in both central business districts and surrounding inner city areas as Hutton (2008) shows for places as diverse as London, San Francisco, Singapore and Vancouver (Figure 8); and high-technology industry is commonly concentrated in dense technopoles that coincide with the more suburban areas of city-regions (Castells and Hall, 1994; Scott, 1990).

These clusters of cognitive–cultural producers are almost always sites of persistent industrial innovation, if not hyper-innovation. In contrast to the top-down R&D model of innovation that was so characteristic

Figure 8 Central Vancouver, June 2008
Tom Hutton (2008, p. 222) notes that 'Vancouver has emerged as a largely post-fordist production centre . . . without first developing as a major site of fordist manufacturing.' At bottom-left is the Olympic Village under construction, not long before the global financial crisis frightened off the hedge fund providing bridge financing, forcing a City bailout. To the far-right of the photograph (just where the float-plane's trailing cable appears) is the redeveloped Woodward's complex, under construction. Situated in the heart of the poor and working-class Downtown Eastside, Woodward's condos were marketed under the slogan, 'Be bold or move to suburbia.' Promoted as 'one of the most ambitious urban redevelopment projects in Canadian history', the mixed-use complex has its own creative-class documentation in the form of a book, *Body Heat* (Enright, 2010).

of fordism (and that has by no means disappeared), this alternative model of innovation operates mainly on the basis of small-scale informal changes that may be virtually imperceptible on an individual basis, but that sometimes result cumulatively in quite significant system-wide advances. Small wonder, then, that concepts like the 'innovative milieu' and the 'learning region' should come to figure so largely on the agendas of urban scholars and policy-makers after the early to mid-1980s, particularly, in the first instance, in regard to high-technology industrial development (Aydalot, 1986; Camagni, 1991; Maillat and Vasserot, 1986). Equally unsurprising is the observation that research on these issues has been complemented in more recent years by a growing interest in creativity, and above all in the question of the creative city (Cooke and Lazzeretti, 2007; Currid, 2006; Landry and Bianchini, 1995; Storper and Scott, 2009). Then again, as Hall (1998) has shown in enormous detail, cities have always been places in which learning, innovativeness and creativity have flourished, though the shape and substantive content of these processes differ greatly from time to time and from place to place depending on prevailing social and economic circumstances. In 19th-century Lancashire innovation revolved insistently around the cotton textile industry, and was focused most especially on spinning and weaving technologies; in Detroit in the period of

Figure 9 Guangzhou, China, March 2010
Pannell (2002, p. 1582) observes that 'to take a train or bus ride for the approximately 110 miles between Hong Kong and Guangzhou, the two largest cities along the right flank of the Pearl River estuary, offers a quick and insightful transect of China's urban future'. This future involves reconfigured relations between production and consumption, and intensified market relations. Guangzhou is 'the key city within one of the most dynamic local economies in the world, and has been subjected to pervasive and far-reaching marketizing influences' (Forrest and Yip, 2007, p. 62).

fordism it was heavily concentrated on the car industry; in the great business hubs of London, New York and Tokyo today it is expressed in the circulation of bewildering arrays of new financial instruments.

This point about learning, innovation and creativity being rooted in some sort of concrete social base (whether economic or not) is critical. It is all the more crucial because much of the contemporary literature on creative cities unduly abstracts creativity away from prevailing social and economic circumstances, particularly in the accounts that are most closely associated with the work of Florida (2004). More specifically, innovation and creativity do not emanate unilaterally out of the minds of a 'creative class', even if innovation and creativity are always in part a function of individual talent and capability. Rather, innovation and creativity (together with related processes of learning) are always mobilized in concrete ways in relation to actual life forms and daily practices (Csikszentmihalyi, 1990; Grabher, 2001; Scott, 2010a; Törnqvist, 2004). The latter remark translates directly into the proposition that whatever forms of creativity may be operative in the urban economy today, they reside primarily, though certainly not exclusively, in the organizational peculiarities of the production system and in the specific challenges and opportunities that it throws out to individuals positioned within and around it. By the same token, cities in which the cognitive–cultural economy is highly developed are often unusually 'creative' in the sense that work, life and landscape all bear the marks of the technocratic and affective

Figure 10 Shenzhen, China, March 2010
By the beginning of the 21st century, Shenzhen had attained the highest per capita income of China's major cities. The 'factory of the world' is not without its consumption and retail spaces, and shopping in Shenzhen seems to echo the declaration of Jim Hightower, President of the South Street Seaport Museum, in 1979: 'The fact is that shopping is the chief cultural activity in the United States' (quoted in Defilippis, 1997, p. 405).

aptitudes and sensibilities that are mobilized within this peculiar economic and urban process (Figures 9 and 10).

Developments like these are by no means confined to the richer areas of the world, but are also increasingly to be observed in many poorer countries in Asia, Latin America and even in parts of Africa. Moreover, the cities in which these developments have advanced most vigorously are caught up in ever-escalating global competition with one another as well as in collaborative projects and joint ventures. In numerous cases, the competitive advantages of these cities are underscored not only by reason of the increasing returns effects that accrue to agglomerated economic activity, but also to their command of Chamberlinian benefits based on the local idiosyncrasies, such as skills, know-how, traditions, design ideologies, historical circumstances, and so on, that generally adhere to particular places. Many attempts by cities to brand themselves capitalize on the same idiosyncrasies. Thus, in addition to the Silicon Valleys, Alleys, Gulches, Glens, Casbahs and Wadis that have sprung up over the last few decades, Singapore once the electronics subcontracting capital of the world, now refers to itself as the 'global city of the arts', Bangkok is seeking to formulate policies that will turn it into a 'global creative city', and Seoul has branded itself as the 'city of design'. Cultural producers all over the globe play on similar registers by incorporating local aesthetic and semiotic values into their outputs, though these often need considerable tweaking in order to widen their appeal on global markets. The film industry of Bollywood exemplifies this point with some insistence (Lorenzen and Taube, 2008).

4. Social geography and local labor markets in third wave cities

The restratification of urban society

As we have seen, the industrial apparatus of fordism was the site of a dominant white-collar/blue-collar division of labor. This division of labor was in turn projected out into urban space, as it were, where it reappeared as an irregular but nonetheless discernible geographic division of neighborhoods. Urban analysts still refer widely to the patterning of urban social space in terms of white-collar and blue-collar neighborhoods, though just as the advent of the new economy calls for updated notions about urban production space, so also does it call for revision in the ways in which we think about the social geography of the city and the local labor markets that are interwoven with it.

The essential issue to be emphasized here is that the new division of labor that has come into being alongside the cognitive–cultural economy is manifest in a corresponding restratification of urban society, and hence in a recomposition of urban social space. This restratification is reflected both in the widening social and economic gap between the top and bottom halves of the labor force (especially in major metropolitan areas) and in the changing occupational engagements of individuals on either side of this divide (see, for example, Cai *et al.*, 2010; Fainstein, 2001; Hamnett and Cross, 1998; Walks, 2001; Yun, 2006). The factors sustaining the divide are unquestionably many-sided, but are at least in part related to the shifting technological foundations of contemporary capitalism, and the ways in which digitization substitutes for routine forms of work, but simultaneously complements and enhances high-level cognitive and cultural skills (Autor *et al.*, 2003).

The upper tier of the labor force, then, can no longer be simply assimilated into some undifferentiated 'white-collar' category, much less into the figure of the 'organization man' identified by Whyte (1956) in the mid-1950s. Today, the upper tier (which now includes substantial numbers of women) is considerably less regimented and bureaucratized than it was when the old hierarchical model of fordist corporate organization lay at the center of the capitalist system. This change was already intimated by Gouldner (1979) with his theory of the 'new class'. Elements of the old hierarchical model can still be found, but the upper fraction of the labor force is nowadays increasingly called upon to exercise its personal talents and creativity in much more open-ended tasks that entail key skills such as analytical shrewdness, deductive reasoning, technical insight, inter-personal judgment, imaginative thinking, cultural sensibility, story-telling abilities, and so on. At the same time, these tasks must frequently be undertaken in the context of shifting collaborative teamwork that in turn requires special skills of social interaction and empathy so as to ensure smooth progress toward desired (business) goals (Grabher, 2001).

The lower tier of the labor force in major metropolitan areas is also undergoing major transformation as the old blue-collar working class dwindles in size and is replaced by a sort of proletariat in low-wage service-oriented occupations, or what I alluded to earlier as a new servile class (cf. Drayse, 2004; Hondagneu-Sotelo, 2002; Peck and Theodore, 2001). In large cities all over the world, unusually high proportions of the servile class are made up of culturally or ethnically distinct immigrants often undocumented and with a heavy over-representation of women (Fernandez-Kelly and Garcia, 1989; McDowell, 2009). Out of the abundant exemplary cases that could be cited here, we might mention Caribbean and Hispanic workers in New York or Los Angeles, North Africans and Turks in Paris or Munich, and Filipinos and Indonesians in Hong Kong or Singapore. The same phenomenon is observable in the case of many migrants from impoverished rural parts of China to Beijing, Guangzhou and Shanghai. The members of this burgeoning servile class are employed in jobs that sustain the social life and physical infrastructures of the city, and in this sense they are very much—directly and indirectly—at the service of the upper tier of the labor force. For one thing, they are employed in diverse domestic and commercial settings as child-minders, gardeners, kitchen hands, janitors, hotel maids, and so on; for another, they undertake much of the physical work, from utilities maintenance to sign painting and vehicle operation that sustains the material fabric and built environment of the city (see, for example, Aguiar and Herod, 2006; Kang, 2003; Soja, 2010). Some of these service-oriented jobs, moreover, have been growing with unusual rapidity of late. For example, among the fastest growing jobs in large US cities in the first decade of the 21st century we find such low-wage service occupations as couriers and messengers, dining room attendants, dishwashers, food preparation workers, grounds maintenance workers, highway maintenance workers, hotel clerks, miscellaneous motor vehicle operators, painters, parking lot attendants and service station attendants (Scott, 2009, 2010b). Because these jobs cannot be geographically disarticulated from the wider stream of life and work in urban areas, they are not susceptible to amalgamation into standardized packets of work that can then be dispatched to offshore locations (Gatta *et al.*, 2009).

Workers in both of these tiers of the labor force are subject to significant and increasing employment risks as a result of the high incidence of temporary, limited-contract and part-time work in cities where the contemporary cognitive–cultural economy is strongly present. These instabilities in turn are manifestations of the flexible, intermittent and project-oriented nature of much of the new economy and its gyrating demands for different categories of labor. Members of the upper tier, however, are customarily integrated into finely honed information networks that enable them to short-circuit at least some of these problems (Scott, 1998), and many of them, in any case, have sufficient material resources to carry them relatively smoothly through periods of under-employment or unemployment. Members of the lower tier are generally faced with even more erratic employment prospects, and have more narrowly circumscribed job-related networks than upper tier workers, and these circumstances, combined with low wages, mean that substantial numbers of them are endemically engaged in a constant struggle to make ends meet.

The recomposition of urban social space

In the years preceding and immediately following World War II, the white-collar and blue-collar neighborhoods of the fordist city were commonly, if schematically, laid out in the form of suburban and inner city rings, respectively. From the 1950s onward, mounting rounds of industrial decentralization from the center to the periphery of the city started to occur. As a consequence, many blue-collar workers began to withdraw from old inner city areas and to take up residence near newly decentralized manufacturing jobs. Over the following few decades, these shifts resulted in a patchwork of white- and blue-collar neighborhoods in what had once been the relatively homogeneous suburbs. The social patchwork of American suburbs subsists down to the present moment, and is all the more intricate today by reason of rapidly growing complements of racially and ethnically distinctive enclaves in the outer fringes of the city (Figures 11–13) (Phelan and Schneider, 1996). Many of the residents of these enclaves are employed in production worker jobs in the high-technology firms and other industrial establishments that still occupy suburban locations in American cities.

Figures 11–13 Insta-ghost town suburb, Summerlin, Nevada, April 2009; foreclosed house in Lehigh Acres, Florida, dubbed 'the Ponzi state' by *The New Yorker* (Packer, 2009), May 2009; the 100-mile city in Phoenix, Arizona, June 2010 If the social ecology of the fordist metropolis was driven by the dispersal of once-centralized factory production and the acceleration of residential suburbanization, many of today's third wave city-regions represent a version of the '100-mile city' (Sudjic, 1993). If a key part of second wave suburbanization involved a *productive* spatial fix of the creation of peripheral built environments (Walker, 1981), the third wave involved the use of these places and people as collateral for new transnational circuits of *parasitic* mortgage debt, investment and speculation. Top-tier global financial corporations promoted sophisticated instruments of leverage—hybrid adjustable-rate and pay-option mortgages, collateralized debt obligations, credit default swaps—but to deploy them for accumulation required real people (targets) in real places. Mortgage exploitation was refined in America's racially marginalized inner cities, but eventually the industry's voracious 'appetite for yield' (Ashton, 2009) led to a push into the vast, white middle-class terrain of Sunbelt suburbs in what financial analysts dubbed the 'sand states' of California, Nevada, Arizona and Florida (Immergluck, 2009).

In parallel with these trends, the social complexion of inner city areas also started to change quite markedly after the 1950s (Figures 14 and 15). In the early 1960s, Glass (1964) had observed that a number of poorer neighborhoods in central London

Figure 14 Redevelopment of inner city neighborhoods. 'The Beasley' condominium, named after a former senior city planner, Vancouver, BC, September 2010

(e.g. Islington, Paddington, North Kensington) were undergoing social transformation as a result of upper middle-class individuals moving in and upgrading the local housing stock. She coined the term 'gentrification' to designate this phenomenon, a term still widely in use today, though its meaning has become increasingly blurred as it has been extended to cover not only residential redevelopment of inner city neighborhoods but also commercial redevelopment in central business districts. In what follows, I want to distinguish between these two rather different forms of urban change—despite a degree of overlap between them at the edges—and I shall use the term 'gentrification' strictly in its original meaning as given by Glass to signify upgraded residential redevelopment. Gentrification in this specific sense has continued to make major inroads into inner city neighborhoods, not only in North America and Western Europe, but also in many other parts of the globe, and as I shall argue, the wider unfolding of this trend is fundamentally linked to growing pressures emanating from the cognitive–cultural economy in urban areas.

Full-blown gentrification in any given neighborhood is often preceded by a

Figure 15 Redevelopment of inner city neighborhoods. Beacon Towers and Beacon Mews, luxury condominium in Harlem, February 2011

transitional phase when working-class housing (some of it already abandoned) and adjacent properties start to be colonized by artists and bohemians. This phase is usually accompanied by a corresponding growth of local services (such as bars and music venues) catering to their needs (Lloyd, 2002; Zukin, 1982). Accelerated outmigration of the original working-class residents also occurs as rents in the local area rise and as inner city manufacturing jobs continue to decline. The pace of change in this phase is frequently boosted by unscrupulous landlords eager to reap the benefits of higher property values and by over-zealous city councils anxious to enhance the image of the city (Lees *et al.*, 2008; Slater, 2006; Wacquant, 2008). As this happens, inward movement by members of the upper tier of the labor force to the central city takes place in the guise of inmigration by cognitive and cultural workers with demographic profiles like young professional families, cohabiting couples, people in same sex unions, apartment sharers, metrosexual singles, and so on (Harris, 2008; Rubino, 2005; Islam, 2005; Hamnett and Whitelegg, 2007; Haase *et al.*, 2010).

This broad dynamic of change was first theorized by Smith (1982) in terms of a so-called 'rent gap'. The origins of the rent gap lie in the difference between the low property values (sustained by potent neighborhood effects) that prevail in inner city areas when working-class families dominate the local social environment, as compared with the augmented values that are obtainable once gentrification paves the way for 'better and higher' use of the land. As plausible as it may be as a starting point, the rent gap theory nevertheless fails to provide all the elements of a coherent explanation. More specifically, as Smith (1986) himself intimates, we need some further account of the historical timing of gentrification in relation to a wider view of changing urban structures as fordism gives way to post-fordism and then to cognitive–cultural capitalism. I have already indicated that one of the reasons for working-class abandonment of inner city residential areas in the post-war decades was local deindustrialization and a corresponding shift of families to suburban neighborhoods with access to expanding employment opportunities in manufacturing. In addition, and notably after the 1970s, the remarkable growth of the cognitive–cultural economy was accompanied by an enormous expansion of employment demands in central business districts and surrounding areas of the city for highly qualified workers. This expansion created, and still creates, a major incentive for well-paid cognitive and cultural workers to move into residential locations nearby, leading to corresponding upgrading of local properties (cf. Hutton, 2008; Ley, 1996) and hence to the phenomenon of the rent gap. A simple illustration of the forcefulness of this trend can be found in the fact that in the 1980s the center of Manchester had a population of only a few hundred, whereas today, after gentrification, the population is over 20,000 (Ward *et al.*, 2010). However, even as gentrification moves forward, dense residential enclaves of low-wage workers (especially ethnic immigrants engaged in the downgraded service and sweatshop segments of the urban economy) typically continue to maintain a foothold in selected areas of the inner city (Salway, 2008; Sassen-Koob, 1982; Sassen, 1991).

5. Urban milieu and built form

Architectural design is at once an important sector of the new cognitive–cultural economy and an instrument of its deployment in urban space. Another way of saying much the same thing is that the socio-economic shifts chronicled here have in part been ushered in on the basis of parallel transformations of the urban milieu and built form. Gentrification as described above is one element of this process. An allied phenomenon can be observed in the insistent redevelopment that has been occurring over the last few decades of central business district properties devoted to business and commercial ends. I shall refer to this phenomenon here in the awkward but evocative term 'aestheticized land use intensification'.

Land use intensification in general is nothing more than an increase in floor space and/or gross product per unit of land area as a way of augmenting Ricardian rent. Any such increase has a multiplier effect on existing rent, which means that the most insistent forms of intensification will tend to occur on land where rent is already high, subject to further conditions regarding the relationship between the cost of intensification and the expected change in total site-specific revenue (Nowlan, 1977; Scott, 1980). This general principle is evident in the unusually dense concentrations of floor space that are always to found in central city areas, as well as in the constant redevelopment to which properties in these areas are subject. Insistent land use intensification has been episodically present in the central business districts of capitalist cities virtually from the beginning, often with political backing as a way of breaking through indurated forms of spatial inertia. The visible signs of this dynamic are the ever escalating vertical expansion of large city centers and the revitalization of formerly derelict areas around these centers where large-scale land assemblies can be put together and where vigorous rehabilitation (as in the case of gentrification) is able to restore dramatically improved value to the land.

The long-standing practice of land use intensification has always been subject to various forms of aestheticization, but this process has taken on new and greatly enlarged meaning in the current conjuncture as a function of the creative urges and possibilities unleashed by urbanization in the global cognitive–cultural economy (Bontje *et al.*, 2011; Schmid *et al.*, 2011). Among other things, the built structures and landscapes of the central city undergird and reflect the upper tier creative, cerebral and cultural work activities that now dominate in these areas. Here, the more chaste geometries of modernism are steadily giving way to idiosyncrasy and bombast. Sklair (2005, 2010) has pointed out that forms of architectural expression in the cores of major world cities typically exhibit something of the spirit of the global corporations and financial organizations that favor these locations. As such, these forms also accord well with the aspirations and ideology of an ostentatiously wealthy transnational capitalist class that links these locations into worldwide business networks. At the same time, the corporate spaces of the central city are increasingly integrated with more consumer-oriented facilities such as gallerias, shopping malls, music centers, museums, art galleries, conference facilities, sports stadia, and so on. These amenities are principally at the service of the upper fraction of the labor force, and, by making the central business district yet more enticing to this fraction, they add their weight to the glamour of the city and further encourage gentrification of adjacent residential neighborhoods.

The aestheticization of intensified land uses in major metropolitan areas is stimulated yet more by the increasing exploitation of urban design and built form to serve as adjuncts to local economic development strategies. Large-scale iconic architectural gestures play an especially critical role in this respect by virtue of their capacity to serve as urban branding and marketing devices. The point is well illustrated by recent megaprojects (often signed by celebrity architects) (Figure 16) such as Los Angeles' Disney Hall, London's Docklands, Bilbao's Guggenheim Museum, Kuala Lumpur's Petronas Towers (Figure 17), Shanghai's World Financial Center, and so on, as well as by a host of comparable cases. What is more, it is precisely the new digital technologies of cognitive–cultural capitalism that make it possible to achieve the characteristically complex topological forms that typify so many of these projects. In these and other ways, the physical landscape of the city functions as a trump card in the global inter-city contests for status, inward investment, mega-events (such as the Olympic Games) and tourist dollars that recur at ever shorter intervals of time in the contemporary world (Figures 18 and 19)

(Kaika, 2010; Knox, 2011). In the absence of complementary initiatives, however, and as the Bilbao experience seems to indicate, large-scale architectural projects do not always live up to expectations about their local economic development potentials (Miles, 2005; Plaza, 2008).

As a number of the architectural projects cited above suggest, some of the most dramatic changes wrought on the urban landscape by aestheticized land use intensification are to be observed in large metropolitan areas of North America and Western Europe that were once at the forefront of fordist economic development. In these centers, formerly gritty and dispiriting landscapes in inner city areas are giving way before advancing waves of a regenerated milieu rich in visual experiences and spectacular urban symbols (Hannigan, 1998). Equally, the real or perceived incompatibilities between residential space and production space that once seemed to be a virtually inescapable aspect of life in earlier versions of capitalist urbanization are now being partially erased. Instead, a new sort of balance appears to be emerging as these spaces become more spatially interwoven and less mutually repellent in terms of their social functions and needs. Even parts of suburbia display a similar syndrome, as represented by verdant industrial parks with sleek buildings accommodating high-technology firms and other commercial ventures alongside the sweeping residential estates and gated communities of upper-income denizens of the urban periphery. Still, the paradox of the large metropolis in cognitive–cultural capitalism is the contrast between the dazzle and glamour of much of its immediate outward form and the squalor of its darker underside (Currid-Halkett and Scott, 2011).

6. The city and the world

Urbanization in the new global economy

Cognitive–cultural capitalism has developed in parallel with the great amplification of globalization that has occurred over the last few decades. The relations between these two phenomena are enormously complex and interlaced, though both owe much to the underlying technological changes and the drive to market extension that are characteristic of this moment of history. The same moment has witnessed the massive expansion of a far-flung system of cities, the most important of which constitute the global city-regions of the new world order (Etherington and Jones, 2009; Hall, 2001; Scott *et al.*, 2001). Many of these are represented in Table 1. Each of the world's great city-regions comprises a nucleus constituted by a large metropolitan area, or, increasingly, multiple and spatially overlapping metropolitan areas, surrounded by an extended hinterland area that may itself be the site of other urban settlements. City-regions function to a high degree as motors of the global economy as a whole. They represent the main concentrations of value-adding activity in the contemporary world economy, and are the chief command and control centers of the large multinational corporations at the pinnacle of the entire capitalist system (Sassen, 1991). They are also the privileged centers of the new cognitive–cultural economy.

City-regions are caught up with one another in intense and multifaceted relationships spanning the globe. Some of these relationships entail collaborative interactions, as exemplified by burgeoning international networks of intra-corporate activities and inter-corporate partnerships in many different sectors (especially in the cognitive–cultural economy) including business services such as banking, law, accountancy, insurance, management consulting, advertising, graphic design, and so on (Taylor *et al.*, 2008). Others are more competitive, though in today's economy, competitive relationships between city-regions are deeply marked by Chamberlinian processes, for as shown earlier, the products of cognitive–cultural labor tend to be highly differentiated in terms of performance and/or symbolic meaning as a function of unique production assets at their points of origin. This proposition

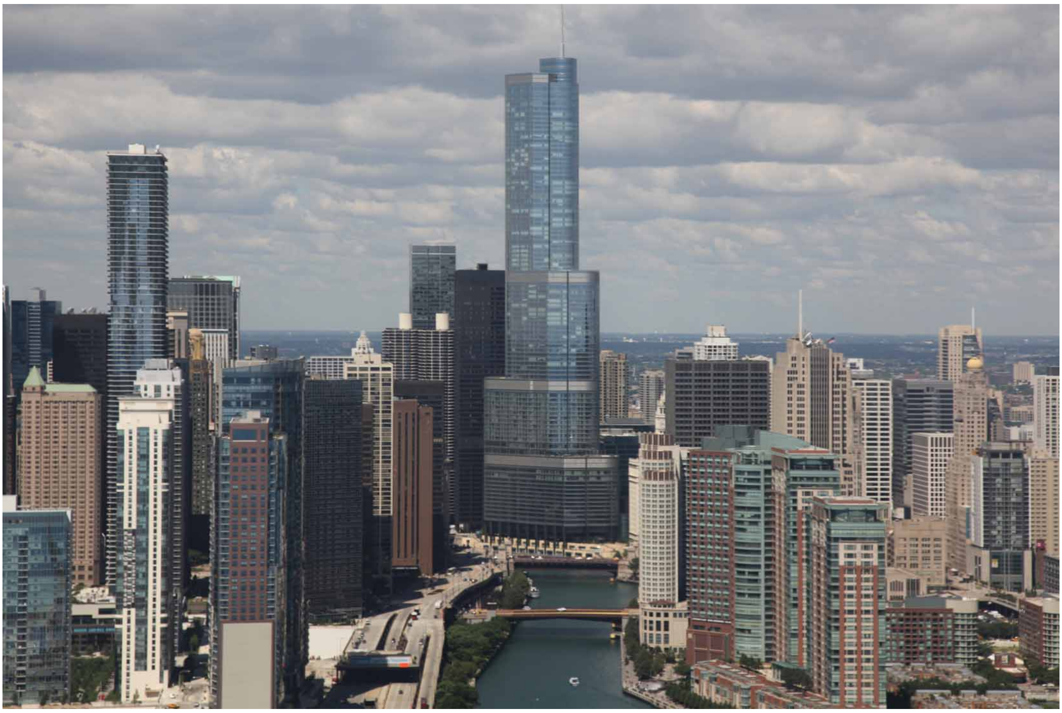


Figure 16 Land use intensification and the aestheticized landscape of the central business district. Chicago, Trump Towers, July 2010

has notable significance in regard to specifically cultural products such as film, television programs, music, fashion clothing, and so on. It implies (even in a world of rapidly diminishing transport costs) that different places producing similar types of goods or services can maintain durable competitive advantages over extended periods of time. In the absence of location- or place-based product differentiation, less developed centers would in the long run tend to succumb to the competitive advantages of more developed centers, and, ultimately, of the most developed center, based on the superior increasing returns effects generated by the latter. I recognize, of course, that decreasing returns will often set in when cities exceed some critical threshold, but the point is that decreasing returns are an essentially transitory phenomenon in the face of persistent urban planning and engineering initiatives. In other words, globalization does not necessarily threaten the long-run sustainability of multiple centers of cultural production in different countries. As a corollary, the dire warnings of earlier culture theorists such as Mattelart (1976) and Michalet (1987) about the incipient standardization of world culture under the sign of the imperial US multinational corporation now have less force than they once did, though, certainly, important questions remain about the increasingly commercial origins of this culture and the threat it poses to many time-honored traditions and ways of life. Even Hollywood, as dominant as it may continue to be, is nowadays beset on all sides by competitive products emanating from film, television-production and music-recording centers in other parts of the world. The city-regions of the modern world system are far from converging toward some sort of monochromatic sameness; rather, they constitute the vital cores of what is currently crystallizing out as an assertively polycentric and polyvocal capitalism.

Figure 17 Land use intensification and the aestheticized landscape of the central business district. Petronas Towers, Kuala Lumpur, Malaysia, January 2010
The Petronas Towers are the landmark that led Asia's contemporary dominance of the race for superlatives in global 'skyscraper geographies' (McNeil, 2005). The towers function as symbolic instruments in 'realizing a state vision of national development', and helped 'to image Malaysia as a "world class" national player (and Kuala Lumpur as a "world city") as well as to promote new "ways of seeing" among national citizens' (Bunnell, 1999, p. 1). Such megastructures also enable 'megaprojection', for instance, through cinema. Bunnell's analysis of the towers' portrayal in the film *Entrapment* (2004, p. 300), however, reveals an underlying irony: the film cut scenes of the towers with scenes from slums in distant Malacca: 'Kuala Lumpur here remained imaginatively trapped within a "third world" … "slums", "pollution", "poverty"—all those signs of underdevelopment that a decade of investment had sought to erase or, at least render out of sight, had been collapsed into a single frame alongside Kuala Lumpur's world class architectural centerpiece.'

City-regions: institutional responses

City-regions today are caught between two major sets of opposing but interrelated forces coming, on the one side, from internal pressures residing in their internal structure, and, on the other side, from external dangers and opportunities at many different spatial scales right up to the global. To be sure, these forces extend well beyond the framework of the cognitive–cultural economy as such, but they interact in important ways with elements of this framework and the resulting outcomes have many implications for the performance and impacts of the new economy in urban areas.

Urban areas in capitalism have always been sites of breakdowns involving the spatial and temporal misallocation of resources, underinvestment in needed public services, market failures, and countless other inefficiencies and irrationalities. In the era of Keynesian-welfare statism, breakdowns of these sorts were generally subject to energetic policy attention when they threatened overall urban performance and welfare (Brenner, 2004). By comparison, in third wave cities, enveloped as they have been hitherto in a dominantly neoliberal policy environment (with its dampening effect on public spending), these breakdowns are often left in

Figure 18 Site preparations for the Shanghai World Exposition, March 2010. 'Better City, Better Life'
Urban entrepreneurialism has been thoroughly transnationalized, routinized and accelerated, creating a strange and dynamic urban hybrid of Berry's 'cities as systems within systems of cities' (1964) and Sheller and Urry's 'new mobilities paradigm' (2006). Every year, every month, cities are bidding to host conventions, expositions, international sporting and hallmark events like the Olympic Games. The narrative and imagery of bids invariably emphasize the unique, creative-class consumption appeal of hopeful host cities (McCallum *et al.*, 2005). At the same time, China's staging of the 2008 Olympics and the 2010 World Exposition was a model of hierarchical, state-directed fordism.

abeyance by policy-makers. Conversely, certain urban policy questions that were relatively subdued at an earlier period have now come much more firmly onto the agenda. These revolve primarily around issues involving the competitive pressures on cities due to globalization and the increasing heterogeneity of urban society due to large-scale immigration (see below). Any search for answers to these questions is apt to be at least partially vitiated by a further problem residing in the municipal fragmentation that almost always characterizes the internal space of city-regions today, and that, in and of itself, is a persistent source of inefficiency and social bias in the provision of urban public goods and services.

Many city-regions all over the world are now struggling to come to terms with the dilemmas and tensions engendered by this internal administrative splintering and to create some more rational and inclusive managerial order. Toronto, London, Tokyo and Hong Kong, among others, have all instituted one version or another of extended metropolitan governance. The specific solutions that have been implemented take on assorted forms such as municipal federations, multi-tiered jurisdictions and, as in the case of Los Angeles, region-wide special-purpose districts. Chinese city-regions, as Vogel *et al.* (2010) have indicated, offer some particularly revealing instances of this quest for overarching administrative command. The advent of capitalism in that country has been attended by precipitously rising levels of inter-municipal competition, a phenomenon that was unknown in state socialism. This turn of

Figure 19 Hibao, mascot for the Shanghai World Exposition

events has set in motion a related drive to establish city-region governmental structures, both as a way of securing higher levels of operating efficiency and as a mechanism for building more effective local economic development policies (Figure 20).

To put the matter in more general terms, numerous experiments in building new institutions of collective management and coordination at the level of the city-region are proceeding apace in response to the weakening of the protective umbrella of the national state and the deepening of global competition. The search for some measure of city-region governance is even more imperative in metropolitan areas with a high quotient of cognitive–cultural industries, for the persistent market failures and the sometimes wayward evolutionary pathways embedded in agglomerations of these industries mean that they are susceptible to various forms of collective underperformance (Cooke and Morgan, 1998; Scott, 1988; Storper, 1997). City-wide governance enhances policy opportunities for improving localized competitive advantages by dealing with critical common-pool assets and liabilities in their full deployment over urban space. This emphasis on competitiveness also helps to account for the relatively recent emergence of aggressively entrepreneurial mayors and city councils in the USA (Harvey, 1989). Similarly, in many countries, national urban policy has turned away from old fordist-era strategies of locational decentralization in favor of measures that encourage cities to develop their own growth potentials on the basis of their own local assets. The rationale underlying this policy shift is partly founded on the renewed innovativeness of localized industrial clusters in an economic environment where the cognitive and cultural assets of the labor force play an ever augmenting role in the development process. Perhaps nowhere is this shift more evident than in France, where one of the latest substitutes for the old decentralization policy mania of the fordist era can be found in the ambitious *Pôles de Compétitivité* program initiated in 2004, with its emphasis on advanced technology-intensive and creative industries (Darmon and Jacquet, 2005; Guillaume, 2008; Pecqueur, 2008).

Pressing social policy questions have also come prominently into sight in cities where the cognitive–cultural economy is highly developed and especially in the context of the widening gap between the incomes and prospects of upper and lower tier workers in these cities. These questions have been magnified by the persistent absorption of politically marginalized immigrants into the low-wage underbelly of the cognitive–cultural economy, and they fester all the more in the anti-welfare policy regimes that are nowadays all too common in large urban areas all over the world. They are further expressed in the rapidly diminishing levels of enfranchisement and democratic incorporation of large segments of the population in the lower reaches of urban society (Jonas and Ward, 2007). There are no simple technocratic solutions to the indurated social problems highlighted by these remarks, particularly in view of the resistances that can be expected to materialize from existing vested interests. It is one thing to

Figure 20 Shenzhen, China, February 2010
Key elements of third wave cities are signified in the 'ambitious urbanization project' that 'aims to turn the Pearl River Delta, the factory of the world, into a super-size metropolis' linking Guangzhou, Shenzhen and seven other cities into a 40,000 square kilometer city-region with a population of 42 million (Lau, 2011). Some 150 major infrastructure projects are planned to create an integrated regional network of transport, water supply, energy, and information and telecommunications. The region is developing a transnational urbanist fusion of cognitive–cultural production—with lightning-fast adaptation to the latest shifts in consumer demand—and updated forms of assembly-line proletarianization. Even before the worst months of the global financial crisis of 2008, more than 67,000 factories across China had closed in anticipation of collapsing demand, fueling a wave of labor protests over unpaid back wages and large-scale return migration to rural areas (Wong, 2008). Chan (2010, p. 659) has pointed to the exploitation of 'super-cheapened' rural migrant labor incorporated 'at the bottom of the global supply chain' through China's export-zone factories.

propose social justice programs focused on reforms like subsidized housing or a more equitable local system of taxation and redistribution—not to mention higher wages, improved benefits and better working conditions for members of the servile class—but such programs are not much likely to flourish in the absence of a considerably more forceful level of political contestation than is the case in most of the world's great city-regions today. That said, there is surely a point at which the tensions and inequities of life can be expected to tip the balance in favor of greatly increased political activism on the part of those who toil at or close to the bottom of the urban employment system. In addition, the neoliberal political practices that have accompanied the rise of the cognitive–cultural economy over the last few decades appear now to have become more or less exhausted as instruments of socio-economic regulation. The fiscal crises currently raging at every level of government today are just one symptom of this failure. If correct, these remarks mean that prospects for a new progressive political opening may be on the horizon. If not, one likely alternative is steady and persistent deterioration of the city as a locus of community, conciliation and conviviality, together with continued widening of the gulf between the rich and the poor.

7. Postscript

Over the last three or four decades a series of remarkable shifts have occurred in the space-economy of capitalism and its expression in the urban domain. These shifts are in significant ways being driven by a steadily expanding cognitive–cultural economy with very specific effects on the form and functional characteristics of the modern city, and they are greatly intensified by globalization. I have argued that the cities where these developments are most in evidence represent the vanguard of a third major wave of urbanization in historical capitalism. As cognitive–cultural capitalism and globalization maintain their current trajectory of development, this evolving third wave will no doubt continue to leave ever deeper imprints on the urban character of the world at large. Of course, any statement of this sort must also be accompanied by the proviso that the structure and dynamics of capitalism are invariably filtered through an inherited urban milieu that acts reflexively on their expression on the ground (cf. Kloosterman, 2010). For this and other reasons, my argument is far from insinuating that the notion of the third wave implies that convergence toward some sort of universal norm is occurring among the great cities of the contemporary world.

The logic of urban change today is intertwined with the evolutionary development of a globalizing cognitive–cultural capitalism in the context of a dominantly neoliberal policy milieu. However, the accumulating failures of neoliberal approaches to social and economic regulation, at whatever geographic scale we may care to consider, suggest that the need for a new dispensation is now pressing. The severity of these failures implies, in turn, that new objective opportunities may be opening up for a more progressive approach to political order in the cognitive–cultural economy. In the circumstances, some kind of pragmatic social democracy securing a politically progressive capitalism, would seem to be one of the best feasible bets for establishing a viable and stable pathway to the future, that is, a pathway to economic growth, social justice and democratic order (cf. Fainstein, 2010). An approach of this sort would also provide some of the basic tools for achieving three currently urgent tasks of urban reform, namely, effective institution-building in support of the collective action needs of the new economy, the restoration of redistributive equity and the rehabilitation of communal life.

Acknowledgements

I wish to thank Elvin Wyly for his insightful comments on an earlier draft of this paper as well as for the excellent images and captions that accompany the published version. Any errors of fact and theoretical emphasis are my responsibility.

Notes

1 Not to be confused with the three urban revolutions as posited by Soja (2000).
2 The seven major sub-components of the index are: (1) legal and political framework; (2) economic stability; (3) ease of doing business; (4) financial flow; (5) business center; (6) knowledge creation and information flow; (7) livability.
3 The new servile class has many points of overlap with the so-called 'precariat' (Perrin, 2004; Standing, 2011; Wacquant, 2006). However, the notion of the servile class identifies a social position derived specifically from the division of labor whereas the precariat refers to a generalized syndrome of social marginality that also incorporates the unemployable, the destitute and the 'dangerous classes'. Any attempt from a sociological or anthropological perspective to deepen the concept of third wave cities should certainly deal more explicitly with the latter social categories.

References

Adelson, S. (2011) 'About Marina Bay Sands', http://www.marinabaysands.com (accessed 2 June 2011).
Aguiar, L.M. and Herod, A., eds (2006) *The Dirty Work of Neoliberalism: Cleaners in the Global Economy.* New York: Wiley.
Amin, A., ed. (1994) *Post-Fordism: A Reader* (Oxford: Blackwell.
Asheim, B.T. and Coenen, L. (2005) 'Knowledge bases and regional innovation systems: comparing Nordic clusters', *Research Policy* 34, pp. 1173–1190.
Ashton, P. (2009) 'An appetite for yield: the anatomy of the subprime mortgage crisis', *Environment and Planning A* 416, pp. 1420–1441.
Autor, D.H., Levy, F. and Murnane, R.J. (2003) 'The skill content of recent technological change: an empirical exploration', *Quarterly Journal of Economics* 118, pp. 1279–1333.
Aydalot, P., ed. (1986) *Milieux Innovateurs en Europe* (Paris: GREMI.
Bagnasco, A. (1977) *Tre Italie: la Problematica Territoriale dello Sviluppo Italiano.* Bolgna: Il Mulino.
Becattini, G., ed. (1987) *Mercato e forze locali: il distretto industriale* (Bologna: Il Mulino.
Benko, G. and Lipietz, A. (1995) 'De la régulation des espaces aux espaces de la régulation', in R. Boyer and Y. Saillard (eds) *Théorie de la Régulation: L'Etat des Savoirs*, pp. 293–303. Paris: La Découverte.
Berry, B.J.L. (1964) 'Cities as systems within systems of cities', *Papers of the Regional Science Association* 13, pp. 147–163.
Beyers, W.B. (2003) 'On the geography of the new economy: perspectives from the United States', *Service Industries Journal* 23, pp. 4–26.
Blackaby, F., ed. (1978) *De-industrialization* (London: Heinemann.
Bluestone, B. and Harrison, B. (1982) *The Deindustrialization of America.* New York: Basic Books.
Bontje, M., Musterd, S. and Pelzer, P. (2011) *Inventive City-Regions: Path Dependence and Creative Knowledge Strategies.* Farnham: Ashgate.
Boyer, R. (1986) *La Théorie de la Régulation: Une Analyse Critique.* Paris: Algalma.
Brenner, N. (2004) 'Urban governance and the production of new state spaces in Western Europe, 1960–2000', *Review of International Political Economy* 11, pp. 447–488.
Bunnell, T. (1999) 'Views from above and below: the Petronas twin towers and/in contesting visions of development in contemporary Malaysia', *Singapore Journal of Tropical Geography* 201, pp. 1–23.
Bunnell, T. (2004) 'Re-viewing the entrapment controversy: megaprojection, misrepresentation and colonial performance', *GeoJournal* 59, pp. 297–305.
Cai, H.B., Chen, Y.Y. and Zhou, L.A. (2010) 'Income and consumption inequality in urban China: 1992–2003', *Economic Development and Cultural Change* 58(3), pp. 385–413.
Camagni, R., ed. (1991) *Innovation Networks: Spatial Perspectives* (London: Pinter.
Carney, J., Hudson, R. and Lewis, J., eds (1980) *Regions in Crisis: New Perspective in European Regional Theory.* New York: St Martin's Press.
Castells, M. and Hall, P. (1994) *Technopoles of the World: The Making of Twenty-First Century Industrial Complexes.* London: Routledge.
Cenzatti, M. (1993) *Los Angeles and the L.A. School: Postmodernism and Urban Studies.* Los Angeles: Los Angeles Forum for Architecture and Urban Design.
Chamberlin, E. (1933) *The Theory of Monopolistic Competition.* Cambridge, MA: Harvard University Press.
Chan, K.W. (2010) 'The global financial crisis and migrant workers in China: there is no future as a laborer; returning to the village has no meaning', *International Journal of Urban and Regional Research* 343, pp. 659–677.
Cooke, P. and Lazzeretti, L. (2007) 'Creative cities: an introduction', in *Creative Cities, Cultural Clusters and Local Economic Development'*, pp. 1–24. Cheltenham: Edward Elgar.
Cooke, P. and Morgan, K. (1998) *The Associational Economy: Firms, Regions, and Innovation.* Oxford: Oxford University Press.
Cooke, P. and Piccaluga, A., eds (2006) *Regional Development in the Knowledge Economy.* New York: Routledge.

Csikszentmihalyi, M. (1990) 'The domain of creativity', in M.A. Runco and R.S. Albert (eds)', *Theories of Creativity*, pp. 190–212. Newbury Park, CA: Sage.

Currid, E. (2006) 'New York as a global creative hub: a competitive analysis of four theories on world cities', *Economic Development Quarterly* 20, pp. 330–350.

Currid-Halkett, E. and Scott, A.J. (2011) 'The geography of celebrity and glamour'. Unpublished paper, School of Policy, Planning and Development, University of Southern California, Los Angeles.

Darmon, D. and Jacquet, N. (2005) *Les Pôles de Compétitivité: Le Modèle Français*. Paris: La Documentation Française.

Dear, M. (2005) 'Comparative urbanism', *Urban Geography* 263, pp. 245–251.

Dear, M., Burridge, A., Marolt, P., Peters, J. and Seymour, M. (2008) 'Critical responses to the Los Angeles school of urbanism', *Urban Geography* 29, pp. 101–112.

Defilippis, J. (1997) 'From a public re-creation to private recreation: the transformation of public space in South Street Seaport', *Journal of Urban Affairs* 194, pp. 405–417.

Drayse, M.H. (2004) 'Local labor market restructuring and the employment of welfare recipients in Los Angeles County', *Urban Geography* 25, pp. 139–172.

Enright, R., ed. (2010) *Body Heat: The Story of the Woodward's Redevelopment* (Vancouver: Blueimage.

Etherington, D. and Jones, M. (2009) 'City-regions: new geographies of uneven development and inequality', *Regional Studies* 43(2), pp. 247–265.

Fainstein, S.S. (2001) 'Inequality in global city-regions', in A.J. Scott (ed.) *Global City-Regions: Trends, Theory, Policy*, pp. 285–298. Oxford: Oxford University Press.

Fainstein, S.S. (2010) *The Just City*. Ithaca, NY: Cornell University Press.

Fernandez-Kelly, M.P. and Garcia, A.M. (1989) 'Informalization at the core: Hispanic women, homework, and the advanced capitalist state', in A. Portes, M. Castells and L.A. Benton (eds) *The Informal Economy: Studies in Advanced and Less Developed Countries*, pp. 247–264. Baltimore: The Johns Hopkins University Press.

Flew, T. (2010) 'Cultural economic geography and global media studies: the rise of Asian media capitals', *Journal of the Oriental Society of Australia* 42, pp. 35–49.

Florida, R. (2002) *The Rise of the Creative Class*. New York: Basic Books.

Florida, R. (2004) *Cities and the Creative Class*. London: Routledge.

Forrest, R. and Yip, N.M. (2007) 'Neighbourhood and neighbouring in contemporary Guangzhou', *Journal of Contemporary China* 1650, pp. 47–64.

Franke, S. and Verhagen, E. (2005) *Creativity and the City: How the Creative Economy Changes the City*. Rotterdam: NAi.

Fuchs, V.R. (1968) *The Service Economy,* General Series No. 87. New York: National Bureau of Economic Research.

Garofoli, G. (1987) 'Il modello territoriale di sviluppo degli anni 70–80', *Note Economiche* 1, pp. 156–176.

Gatta, M., Boushey, H. and Appelbaum, E. (2009) 'High-touch and here-to-stay: future skills demands in US low wage service occupations', *Sociology—The Journal of the British Sociological Society* 43, pp. 968–989.

Gertler, M.S. (2003) 'Tacit knowledge and the economic geography of context, or, the undefinable tacitness of being there', *Journal of Economic Geography* 3, pp. 75–99.

Glass, R. (1964) 'Aspects of change', in Centre for Urban Studies (ed.) *London: Aspects of Change*, pp. xiii–xlii. London: MacGibbon and Kee.

Gouldner, A. (1979) *The Future of Intellectuals and the Rise of the New Class*. New York: Seabury.

Grabher, G. (2001) 'Ecologies of creativity: the village, the group, and the heterarchic organization of the British advertising industry', *Environment and Planning A* 33, pp. 351–374.

Gramsci, A. (1971) *Selections from the Prison Notebooks*. New York: International Publishers.

Guillaume, R. (2008) 'Des systèmes productifs locaux aux pôles de compétitivité: approches conceptuelles et figures territoriales du développement', *Géographie, Economie, Société* 10, pp. 295–309.

Haase, A., Kabisch, S., Steinführer, A., Bouzarovski, S., Hall, R. and Ogden, P. (2010) 'Emergent spaces of reurbanisation: exploring the demographic dimension of inner-city residential change in a European setting', *Population Space and Place* 16, pp. 443–463.

Hall, P. (1998) *Cities in Civilization*. New York: Pantheon.

Hall, P. (2001) 'Global city-regions in the twenty-first century', in A.J. Scott (ed.) *Global City-Regions: Trends, Theory, Policy*, pp. 59–77. Oxford: Oxford University Press.

Hamnett, C. and Cross, D. (1998) 'Social polarisation and inequality in London: the earnings evidence', *Environment and Planning C: Government and Policy* 16, pp. 659–680.

Hamnett, C. and Whitelegg, D. (2007) 'Loft conversion and gentrification in London: from industrial to post-industrial land use', *Environment and Planning A* 39, pp. 106–124.

Hannigan, J. (1998) *Fantasy City: Pleasure and Profit in the Postmodern Metropolis*. London: Routledge.

Harris, A. (2008) 'From London to Mumbai and back again: gentrification and public policy in comparative perspective', *Urban Studies* 45, pp. 2407–2428.

Harvey, D. (1987) 'Flexible accumulation through urbanization: reflections on post-modernism in the American city', *Antipode* 19, pp. 260–286.

Harvey, D. (1989) 'From managerialism to entrepreneurialism—the transformation in urban governance

in late capitalism', *Geografiska Annaler, Series B—Human Geography* 71, pp. 3–17.
Hondagneu-Sotelo, P. (2002) *Domestica: Immigrant Workers Cleaning and Caring in the Shadows of Affluence*. Berkeley: University of California Press.
Hutton, T.A. (2008) *The New Economy of the Inner City: Restructuring, Regeneration, and Dislocation in the Twenty-First Century Metropolis*. London: Routledge.
Immergluck, D. (2009) 'From the subprime to the exotic: excessive mortgage market risk and foreclosures', *Journal of the American Planning Association* 741, pp. 59–76.
Islam, T. (2005) 'Outside the core: gentrification in Istanbul', in R. Atkinson and G. Bridge (eds) *Gentrification in a Global Context: The New Urban Colonialism*, pp. 121–136. London: Routledge.
Jonas, A.E.G. and Ward, K. (2007) 'Introduction to a debate on city-regions: new geographies of governance, democracy and social reproduction', *International Journal of Urban and Regional Research* 31, pp. 169–178.
Kaika, M. (2010) 'Architecture and crisis: re-inventing the icon, re-imagining London and rebranding the City', *Transactions of the Institute of British Geographers* 35, pp. 453–474.
Kang, M. (2003) 'The managed hand: the commercialization of bodies and emotions in Korean immigrant-owned nail salons', *Gender and Society* 17, pp. 820–839.
Kavanaugh, K.B. (2001) *Detroit's Michigan Central Station*. Chicago: Arcadia.
Kloosterman, R.C. (2010) 'This is not America: embedding the cognitive–cultural urban economy', *Geografiska Annaler, Series B—Human Geography* 92B, pp. 131–143.
Knox, P.L. (2011) *Cities and Design*. London: Routledge.
Kunzmann, K.R. (2009) 'The strategic dimensions of knowledge industries in urban development', *Disp* 45(2), pp. 40–47.
Landry, C. and Bianchini, F. (1995) *The Creative City*. London: Demos.
Lau, M. (2011) 'Nine cities to be linked in delta megalopolis', *South China Morning Post*, 24 January, p. 5.
Lees, L., Slater, T. and Wyly, E. (2008) *Gentrification*. London: Routledge.
Levy, F. and Murnane, R.J. (2004) *The New Division of Labor: How Computers are Creating the Next Job Market*. New York: Russell Sage Foundation.
Ley, D. (1996) *The New Middle Class and the Remaking of the Central City*. Oxford: Oxford University Press.
Lloyd, R. (2002) 'Neo-Bohemia: art and neighborhood development in Chicago', *Journal of Urban Affairs* 24, pp. 517–532.
Lorenzen, M. and Taube, F.A. (2008) 'Breakout from Bollywood? The roles of social networks and regulation in the evolution of Indian film industry', *Journal of International Management* 14, pp. 286–299.
Maillat, D. and Vasserot, J.Y. (1986) 'Les milieux innovateurs: le cas de l'Arc Jurassien suisse', in P. Aydalot (ed.) *Milieux Innovateurs en Europe*, pp. 217–246. Paris: GREMI.
Massey, D. and Meegan, R. (1982) *Anatomy of Job Loss: The How, Why, Where, and When of Employment Decline*. London: Methuen.
MasterCard Worldwide (2008) *Worldwide Centers of Commerce Index*, http://www.mastercard.com/us/company/en/insights/pdfs/2008/MCWW_WCoC-Report_2008.pdf
Mattelart, A. (1976) *Multinationales et Systèmes de Communication: Les Appareils Idéologiques de l'Impérialisme*. Paris: Editions Anthropos.
McCallum, K., Spencer, A. and Wyly, E. (2005) 'The city as an image-creation machine: a critical analysis of Vancouver's Olympic bid', *Yearbook of the Pacific Coast Geographers* 67, pp. 24–46.
McDowell, L. (2009) *Working Bodies: Interactive Service Employment and Workplace Identities*. Chichester: Wiley-Blackwell.
McNeil, D. (2005) 'Skyscraper geography', *Progress in Human Geography* 291, pp. 41–55.
Michalet, C.-A. (1987) *Le Drôle de Drame du Cinéma Mondial*. Paris: Editions de la Découverte.
Miles, M. (2005) 'Interruptions: testing the rhetoric of culturally led urban development', *Urban Studies* 42, pp. 889–911.
Moulier Boutang, Y. (2007) *Le Capitalisme Cognitif, Comprendre la Nouvelle Grande Transformation et ses Enjeux*. Paris: Editions Amsterdam.
Nowlan, D.M. (1977) 'The land market: how it works', in L.B. Smith and M. Walker (eds) *Public Property?*, pp. 3–37. Vancouver: Fraser Institute.
Packer, G. (2009) 'The Ponzi state', *The New Yorker*, 9 February.
Pannell, C. (2002) 'China's continuing urban transition', *Environment and Planning A* 349, pp. 1571–1589.
Paulson, H.M. (2010) *On the Brink*. New York: Business Plus.
Peck, J. and Theodore, N. (2001) 'Contingent Chicago: restructuring the spaces of temporary labor', *International Journal of Urban and Regional Research* 25, pp. 471–496.
Pecqueur, B. (2008) 'Pôles de compétitivité et spécificité de la ressource technologique: une illustration grenobloise', *Géographie, Economie, Société* 10, pp. 311–326.
Perrin, E. (2004) *Chômeurs et Précaires au Cœur de la Question Sociale*. Paris: La Dispute.
Phelan, T.J. and Schneider, M. (1996) 'Race, ethnicity, and class in American suburbs', *Urban Affairs Review* 31, pp. 659–680.
Piercy, M. (2010) 'Working class nostalgia', *Monthly Review* 624, pp. 45.
Pine, B. and Gilmore, J. (1999) *The Experience Economy: Work is Theatre and Every Business a Stage*. Boston: Harvard Business School.
Piore, M. and Sabel, C. (1984) *The Second Industrial*

Divide: Possibilities for Prosperity. New York: Basic Books.
Plaza, B. (2008) 'On some challenges and conditions for the Guggenheim Museum Bilbao to be an effective economic re-activator', *International Journal of Urban and Regional Research* 32, pp. 506–517.
Reich, R. (1992) *The Work of Nations*. New York: Vintage.
Robinson, J. (2006) *Ordinary Cities*. London: Routledge.
Rogers, E. and Larson, J. (1984) *Silicon Valley Fever*. New York: Basic Books.
Roy, A. (2011) 'Slumdog cities: rethinking subaltern urbanism', *International Journal of Urban and Regional Research* 35, pp. 223–238.
Rubino, S. (2005) 'A curious blend? City revitalisation, gentrification and commodification in Brazil', in R. Atkinson and G. Bridge (eds) *Gentrification in a Global Context: The New Urban Colonialism*, pp. 225–239. London: Routledge.
Salway, S. (2008) 'Labour market experiences of young UK Bangladeshi men: identity, inclusion and exclusion in inner-city London', *Ethnic and Racial Studies* 31, pp. 1126–1152.
Sassen, S. (1991) *The Global City: New York, London, Tokyo*. Princeton, NJ: Princeton University Press.
Sassen-Koob, S. (1982) 'Recomposition and peripheralization at the core', *Contemporary Marxism* No. 5, pp. 88–100.
Saunig, M. (2010) 'Rebranding public nuisance: City of Cleveland v. Ameriquest Mortgage Securities, Inc. as a failed response to economic crisis', *Catholic University Law Review* 59, pp. 911–929.
Saxenian, A.L. (1983) 'The urban contradictions of Silicon Valley—regional growth and the restructuring of the semiconductor industry', *International Journal of Urban and Regional Research* 7, pp. 237–262.
Schmid, H., Dietrich, W.-D. and Urry, J., eds (2011) *Cities and Fascination: Beyond the Surplus of Meaning*. Farnham: Ashgate.
Scott, A.J. (1980) *The Urban Land Nexus and the State*. London: Pion.
Scott, A.J. (1988) *New Industrial Spaces: Flexible Production Organization and Regional Development in North America and Western Europe*. London: Pion.
Scott, A.J. (1990) 'The technopoles of Southern California', *Environment and Planning A* 22, pp. 1575–1605.
Scott, A.J. (1998) 'Multimedia and digital visual effects: an emerging local labor market', *Monthly Labor Review* 121, pp. 30–38.
Scott, A.J. (2008a) 'Inside the city: on urbanization, public policy, and planning', *Urban Studies* 45, pp. 755–772.
Scott, A.J. (2008b) *Social Economy of the Metropolis: Cognitive–Cultural Capitalism and the Global Resurgence of Cities*. Oxford: Oxford University Press.
Scott, A.J. (2009) 'Human capital resources and requirements across the metropolitan hierarchy of the United States', *Journal of Economic Geography* 9, pp. 207–226.
Scott, A.J. (2010a) 'Cultural economy and the creative field of the city', *Geografiska Annaler, Series B—Human Geography* 92, pp. 115–130.
Scott, A.J. (2010b) 'Space-time variations of human capital assets in the American economy: profiles of abilities and skills across metropolitan areas, 1980 to 2000', *Economic Geography* 86, pp. 233–249.
Scott, A.J. and Soja, E., eds (1996) *The City: Los Angeles and Urban Theory at the End of the Twentieth Century*. Berkeley and Los Angeles: University of California Press.
Scott, A.J., Agnew, J., Soja, E.W. and Storper, M. (2001) 'Global city-regions', in A.J. Scott (ed.) *Global City-Regions: Trends, Theory, Policy*, pp. 11–30. Oxford: Oxford University Press.
Sheller, M. and Urry, J. (2006) 'The new mobilities paradigm', *Environment and Planning A* 38, pp. 207–226.
Sklair, L. (2005) 'The transnational capitalist class and contemporary architecture in globalizing cities', *International Journal of Urban and Regional Research* 29(3), pp. 485–500.
Sklair, L. (2010) 'Iconic architecture and the culture–ideology of consumerism', *Theory, Culture and Society* 27, pp. 135–159.
Slater, T. (2006) 'The eviction of critical perspectives from gentrification research', *International Journal of Urban and Regional Research* 30, pp. 737–757.
Smith, N. (1982) 'Gentrification and uneven development', *Economic Geography* 58, pp. 139–155.
Smith, N. (1986) 'Gentrification, the frontier, and the restructuring of urban space', in N. Smith and P. Williams (eds) *Gentrification of the City*, pp. 15–34. London: Allen and Unwin.
Soja, E.W. (2000) *Postmetropolis: Critical Studies of Cities and Regions*. Oxford: Blackwell.
Soja, E.W. (2010) *Seeking Spatial Justice*. Minneapolis: University of Minnesota Press.
Soja, E.W. and Scott, A.J. (1986) 'Los Angeles: capital of the late twentieth century', *Environment and Planning D: Society and Space* 4, pp. 249–254.
Stanback, T.M. (1979) *Understanding the Service Economy: Employment, Productivity, Location*. Baltimore: The Johns Hopkins University Press.
Standing, G. (2011) *The Precariat: The New Dangerous Class*. London: Bloomsbury.
Steinmetz, G. (2008) 'Harrowed landscapes: white ruingazers in Namibia and Detroit and the cultivation of memory', *Visual Studies* 233, pp. 211–237.
Storper, M. (1997) *The Regional World: Territorial Development in a Global Economy, Perspectives on Economic Change*. New York: Guilford Press.
Storper, M. and Scott, A.J. (2009) 'Rethinking human capital, creativity and urban growth', *Journal of Economic Geography* 9, pp. 147–167.
Sudjic, D. (1993) *The 100-Mile City*. London: Flamingo.
Sugrue, T.J. (1995) 'Forget about your inalienable right to work: deindustrialization and its discontents at Ford, 1950–1953', *International Labor and Working-Class History* 48, pp. 112–130.

Sugrue, T.J. (2005) *Origins of the Urban Crisis*. Princeton, NJ: Princeton University Press.

Talha, L. (1995) 'Théorie de la régulation et développement', in R. Boyer and Y. Saillard (eds) *Théorie de la régulation: l'état des savoirs*, pp. 452–458. Paris: Editions la Découverte.

Taylor, P.J., Evans, D.M. and Pain, K. (2008) 'Application of the interlocking network model to mega-city-regions: measuring polycentricity within and beyond city-regions', *Regional Studies* 42, pp. 1079–1093.

Törnqvist, G. (2004) 'Creativity in time and space', *Geografiska Annaler* 86, pp. 227–243.

Vercellone, C. (2007) 'From formal subsumption to general intellect: elements for a Marxist reading of the thesis of cognitive capitalism', *Historical Materialism* 15, pp. 13–36.

Vogel, R.K., Savitch, H.V., Xu, J., Yeh, A.G.O., Wu, W.P., Sancton, A., Kantor, P., Newman, P., Tsukamoto, T., Cheung, P.T.Y., Shen, J.F., Wu, F.L. and Zhang, F.Z. (2010) 'Governing global city regions in China and the West', *Progress in Planning* 73, pp. 1–75.

Wacquant, L. (2006) *Parias Urbains: Ghetto, Banlieues, Etat*. Paris: La Découverte.

Wacquant, L. (2008) 'Relocating gentrification: the working class, science and the state in recent urban research', *International Journal of Urban and Regional Research* 32, pp. 198–205.

Walker, R. (1981) 'A theory of suburbanization: capitalism and the production of urban space in the United States', in M. Dear and A.J. Scott (eds) *Urbanization and Urban Planning in Capitalist Society*, pp. 383–419. New York: Methuen.

Walks, R.A. (2001) 'The social ecology of the post-Fordist/global city? Economic restructuring and socio-spatial polarisation in the Toronto urban region', *Urban Studies* 38, pp. 407–447.

Ward, J.A. (1995) *The Fall of the Packard Motor Company*. Stanford: Stanford University Press.

Ward, K., Fagan, C., McDowell, L., Perrons, D. and Ray, K. (2010) 'Class transformation and work–life balance in urban Britain: the case of Manchester', *Urban Studies* 47(11), pp. 2259–2278.

Warf, B. and Holly, B. (1997) 'The rise and fall and rise of Cleveland', *Annals of the American Academy of Political and Social Science* 551, pp. 208–221.

Whyte, W.H. (1956) *The Organization Man*. New York: Simon and Schuster.

Wong, E. (2008) 'Workers drifting away as plants close in China', *International Herald Tribune*, 15 November, p. 1.

Yun, M.S. (2006) 'Earnings inequality in USA, 1969–99: comparing inequality using earnings equations', *Review of Income and Wealth* 1, pp. 127–144.

Zukin, S. (1982) *Loft Living: Culture and Capital in Urban Change*. Baltimore: Johns Hopkins University Press.

Allen J. Scott, Department of Geography and Department of Public Policy, University of California–Los Angeles. Email: ajscott@ucla.edu

MEDIA CULTURE & SOCIETY

Commentary

The cultural economy: geography and the creative field

Allen J. Scott
UCLA, LOS ANGELES, USA

Introduction

The cultural economy comprises all those sectors in modern capitalism that cater to consumer demands for amusement, ornamentation, self-affirmation, social display and so on. These sectors comprise various craft, fashion, media, entertainment and service industries with outputs like jewelry, perfume, clothing, films, recorded music or tourist services. Such outputs have high symbolic value relative to utilitarian purpose (see Bourdieu, 1971; Lash and Urry, 1994).

Major portions of the modern cultural economy are concentrated in global cities like Los Angeles, New York, Paris, Milan or Tokyo (Hall, 1998; Scott, 1997). Significant geographic fragments of the cultural economy can also be found in other locales, ranging from Las Vegas or the French Côte d'Azur to the neo-artisanal industrial districts of north-east and central Italy. One common characteristic of such places is that their participation in cultural–economic activities is based upon dense networks of producers combined with a dependence on complex local labor markets. My argument in what follows is that these geographic underpinnings are decisive for understanding processes of creativity and innovation in the cultural economy.

Art, science, culture

There is an extended literature suggesting that culture – even in Williams's (1982) rather abstract sense of the 'informing spirit' of a whole way of life – can best be comprehended as a social phenomenon rather than as the expression of some transcendent personalized impulse. Culture, in short, is an immanent construct whose character can only be seized in terms of the wider systems of human relationships with which it is intertwined (Bordwell et al., 1985; Crane, 1992; Negus, 1998; Wolff, 1981).

This argument may be advanced not just for the case of culture as embodied in the everyday artifacts that constitute the materiality of social life, but also as it is expressed in such domains of human activity as art or science. To be sure, there are

Media, Culture & Society © 1999 SAGE Publications (London, Thousand Oaks and New Delhi), Vol. 21: 807–817
[0163-4437(199911)21:6;807–817;010767]

powerful versions of philosophy that arrogate to themselves special authority to issue warrants for aesthetic or scientific practices, but this view is increasingly in retreat in the light of scholarly work showing how these practices connect even at their most intimate moments of genesis with concrete social conditions. Writers on the sociology of art and culture, such as Becker (1974, 1976, 1982), Bourdieu (1983), Crane (1992) or White and White (1965), and on the sociology of knowledge, such as Barnes (1974), Barnes et al. (1996), Latour and Woolgar (1979), Mannheim (1952) or Mulkay (1972) have built up an imposing array of arguments and evidence in favor of this sociological approach.

Four main points need to be made.

1. What can be identified as viable (i.e. inter-subjectively meaningful) topics for art works or scientific projects are given out of conditions of practical and political life.
2. Artistic or scientific work is always molded by the context in which it occurs. One of the more significant variables at play here is the division of labor in cultural production, even in such apparently elusive cases as a painter's studio or a scientific laboratory.
3. Art and science depend on inter-personal norms, methods, languages and so on, in order to achieve communicability. Both practitioners and audiences will therefore have had to undergo some degree of common socialization if they are to connect together.
4. The social profile of consumers of art and science (alternatively, the market) invariably plays a role in how producers conceive and present their finished work. This is all the more so where intermediaries (such as agents, editors or gallery owners) exert an influence on cultural production (Hennion, 1989; White and White, 1965).

These remarks, if correct, begin the task of situating cultural production firmly within the domain of the social. They serve, too, as a prelude to the proposition that the cultural economy in capitalism is just another way of producing human culture, though it does not produce just any culture; on the contrary, the outputs of the modern cultural economy bear a determinate relationship to their social (capitalistic) conditions of production.

In none of this is there any necessary denial of the talents or dispositions of the individual cultural worker. The point is not that these qualities are always submerged in the anonymous apparatus of commodified production, but that they are mobilized and channeled by the manner in which the apparatus works, including the ways in which specialized but complementary workers come together in the tasks of cultural production (see DiMaggio, 1977; DiMaggio and Hirsch, 1976). This statement applies as much to sectors like film or music where (some) workers' identities are overtly inscribed on the final product, as to sectors like furniture or clothing where workers' identities tend to be relatively hidden from the consumer. It applies because modern cultural–economic systems almost always take the form of complex networks of workers *within* firms, linked together by tightly wrought networks of transactions *between* firms, in which many different hands are brought to bear on products as they go through the process of conception, fabrication and final embellishment. The attributes of these networks, and hence of final outputs, are subject to strong economic pressures, which in cultural-products sectors often assume one of two mutually exclusive types. One type derives from efforts to economize on costs by standardizing production processes; the other seeks to ward off competitive threats by means of constant product differentiation.

The cultural content of cultural products, then, needs to be treated as endogenous to the system of production, but also as an authentic vehicle of aesthetic and semiotic expression (no matter how well or badly achieved in any given instance).

Cultural communities

Many types of cultural production, whether in the commodity form or not, are rooted in communities of workers anchored to particular places. Examples range from traditional craft communities, such as the brahmin painters of Nathdwara in north-west India as described by Maduro (1975), through the artistic and intellectual circles of 19th- and 20th-century Paris (see Bourdieu, 1977; Hall, 1998; Menger, 1993), to the actors, directors, writers and so on, that make up the film colony of contemporary Hollywood (Scott, 1984; Storper and Christopherson, 1987). Place-based communities such as these are not just foci of cultural labor in the narrow sense, but are also vortexes of social reproduction in which critical cultural competencies are generated and circulated. They are, too, magnets for talented individuals from other places who migrate to these centers in search of professional fulfillment (Denisoff and Bridges, 1982; Menger, 1983, 1993).

These examples hint at one of the representative features of such communities, namely, that they are less miscellaneous jumbles of individuals following many different and disconnected pursuits than they are collectivities whose members become caught up in mutually complementary and socially coordinated careers (see Montgomery and Robinson, 1993). A major factor binding such collectivities together is the conventions that virtually always come into being in any human community that has subsisted over a period of time. As such, communities are the repositories of an accumulated cultural capital that is one (though only one) of the defining factors of the creative field.

Cultural capital in this sense is further sustained by the institutional infrastructures with which most such communities are endowed. Specialized schools, training establishments, apprenticeship programs and so on are a recurrent feature of well-established communities of specialized workers and provide a ready supply of appropriately socialized neophytes. Workers' organizations, such as unions or professional associations, contribute further to the maintenance of local standards of cultural and economic performance. Other, more idiosyncratic institutions, ranging from museums or associations devoted to keeping alive memories of past accomplishments, to annual festivals focused on celebrating the achievements of the immediate present also reinforce communal cultural frameworks.

This kind of overarching order is in large degree what Marshall (1919) meant by the term 'industrial atmosphere', a term that he applied to the specialized social assets of the manufacturing districts of 19th-century England. Atmosphere in this sense represents an externality, a common set of resources facilitating the adaptation of workers to employment routines and providing a platform for creative and innovative activity; it eases the tasks of intra-community communication, and represents the common ground on which localized groups of firms or workers come together in the solution of workaday problems. Clearly, though, it can also sometimes operate in a negative fashion by making certain options invisible and encouraging premature termination of the search for solutions.

Each community of cultural workers represents a unique and complex case in which inventiveness on the one hand and conventional or institutional constraint on the other give rise to many varied outcomes. In the traditional brahmin painter community described by Maduro (1975) a restrictive tradition keeps creativity and innovation within tight bounds. In the case of the country-music performers who congregate in large numbers in and around Nashville, Tennessee, adherence to a strongly demarcated musical genre is tempered by occasional bravura demonstrations of originality that periodically stretch its boundaries (see Peterson and DiMaggio, 1975; Ryan and Peterson, 1982). In contemporary Hollywood, a new wave of creativity and innovation has been unleashed by the digital effects revolution which has prompted a wholesale transformation of the art of film-making.

Place, community and the cultural economy are thus often closely interconnected. For this reason, many kinds of outputs are indelibly associated in the mind of the discriminating consumer with particular locales. Theater in London, Parisian *haute couture* or furniture from Italy all illustrate this reputation effect. As Molotch (1996) has argued, place endows products with a guarantee, somewhat akin to the accumulated symbolic value of the fashion designer's label (Bourdieu and Delsaut, 1975), and thus is the source of location-specific monopoly rents.

Production system and milieu

Communities of skilled and socialized cultural workers are one thing; mobilizing them into patterns of productive labor is another. It is only when we introduce the much more expansive notion of the production system and its milieu that we really start to grapple centrally with the logic of the creative field. In this regard, we need to go well beyond 'gatekeeping' models of cultural production (Hirsch, 1969, 1972). These models describe the cultural economy as a filtering process through which some products pass while others are rejected along the way; but they are silent on the central question of how the cultural economy itself begets the essential features of final outputs.

Let us begin with a brief allusion to the work of Hennion (1981, 1983, 1989) and Kealy (1979) on the internal operations of recording studios in the popular music industry, though we might equally as well have started off with such analogous sites of cultural labor as the fashion designer's workshop or a work-team creating a multimedia game.

In his 1989 article, Hennion likens the production of recorded music to the execution of scientific experiments in that the recording studio and the laboratory represent organized social milieux where groups of workers seek by trial and error to obtain results that can then be publicized. What is never made public, however (except by the inquisitive anthropologist), is how these results are influenced by the purely social and often quite messy internal order of the laboratory or studio.

In the recording studio, the interactions between the composer/arranger, the performers, the producer, the sound engineer and other critical individuals constitute an inherently collective sphere of artistic experimentation, and even the efforts of the performers themselves do not necessarily always comprise the most decisive ingredient of what is realized on the final recording (Kealy, 1979). The finished products that emerge from the cultural economy are the results of a collaborative labor process that involves many specialized operations by many different individuals (see Frith, 1992; Negus, 1996; Ryan and Peterson, 1982). Even the stars who occupy the pinnacle positions on the work ladder of the cultural economy are in important ways an endogenous expression of its logic.

There is no special reason, however, why we should call a halt to our investigations at the outer walls of the recording studio. The studio is only one element, a sort of microcosm, of a much more extensive domain of activities in the cultural economy, and hence of the creative field. This domain almost always takes the form of an agglomerated production system and its immediate geographic milieu. Two remarks are apposite.

First, the cultural economy is typically a site of dense inter-firm transactions, as illustrated by the modern film industry with its propensity to horizontal and vertical disintegration (Scott, 1999a; Storper and Christopherson, 1987). These transactions involve both traded and untraded interdependencies. In the cultural economy they are often extremely unstable, finely grained, frequent and mediated by face-to-face contact, which means also that they tend to absorb significant resources of time and energy. Interrelated firms thus commonly put a high premium on mutual proximity to one another.

Second, as production proceeds, workers need to be assembled daily at work sites, and labor markets are constantly recalibrated by means of job search and recruitment activities. Spatial clustering of the participants in the economy again has beneficial consequences since it means that firms and workers occupy mutually accessible locations and that local labor markets function as sets of quasi-pooled resources (see Scott, 1998a).

Inter-firm transacting and local labor market processes in the cultural economy thus encourage agglomeration, and this tendency is greatly boosted by the increasing returns effects (including ‘atmosphere’) that usually emerge. Agglomeration and related increasing returns effects not only enhance system *efficiency* but also *creativity*, and perhaps nowhere more so than in the case of cultural-products complexes. It is within such systems of spatial relations that the creative field emerges in its definitive form.

One of the claims of the literature on learning and change in industrial systems is that many improvements in product and process configurations flow steadily from the multiple, small, unrecorded, day-by-day transactions that occur in any production complex (Lundvall and Johnson, 1994; Von Hippel, 1988). Transactions like these are particularly common in agglomerations where production is deeply disintegrated. Transactions that involve negotiations over the design specifications of products and services as they pass along the input–output chain appear to be a remarkably rich source of creative and innovative energy. In a study of ceramic tile production in Sassuolo, Italy, Russo (1985) showed that a stream of small-scale but cumulatively important improvements in production practices could be traced back to detailed interactions between specialized firms as they engaged in business with one another.

In cultural-products agglomerations, these types of interactions are usually well developed (O’Connor, 1991). They are apt to be characterized by close collaboration between the different parties involved (e.g. in the planning of a television show, or in the design and fabrication of a piece of jewelry), so that the consequences in terms of learning will tend to be all the more evident. The gains in know-how and the beneficial on-the-job adjustments that occur in this manner, refer not only to concrete practices and techniques, but also to the emotive content of products. Workers caught up in this sort of activity are apt to emerge with an altered awareness of the imaginative possibilities that lie within their work (see Csikszentmihalyi, 1990).

The instability of markets for cultural products means that these interactions are often subject to rapid rotation, thus sharpening the associated learning effects. With sufficiently large pools of specialized firms and workers, the number of different combinatorial variations in the structure of (inter-firm) production teams is effectively unlimited, and in industries where there is an incessant search for fashion or novelty effects in final products, this system flexibility is critical.

Perhaps the most dramatic instance of this phenomenon can be found in the popular music industry where recording companies maintain an ever-changing flow of releases in the effort to win a place on the hit-parade. Accordingly, frequent access to a large variety of relevant skills is paramount. Los Angeles and New York, the main concentrations of the recording industry in America, have a disproportionate capacity for producing hit singles, and this is evident even after controlling for such complicating factors as the presence of music-industry majors in both places (Scott, 1999b).

These remarks evoke the notion of the learning region as identified by theorists such as Cooke and Morgan (1998) or Storper (1996). The notion is generally brought up in relation to technology-intensive production, though it can be usefully extended to include the cultural economy as well. Thus, it refers here not only to agglomerations of technologically dynamic firms but also to production places where qualities such as cultural insight and imagination are stimulated by the action and needs of the local economy. At the level of the region as a whole, these qualities can be described as reflections of a many-sided economic and geographic system of production, above and beyond the sum of individual creative efforts. This comment is illustrated by modern film-industry agglomerations where scores if not hundreds of different firms typically come together on particular projects. In these circumstances, any film – even a *film d'auteur* – is actually a huge collective venture, and the final product is always a cultural and an economic artifact at one and the same time. From this perspective we can finally make sense of the celebrated but otherwise cryptic afterthought that Malraux (1946) appended to his study of the cinema to the effect that film-making is not only an art but also, intrinsically, an industry.

Temporality

Creativity and innovation in the cultural economy are marked not only by strong spatial patterns, but also by a robust temporal logic. A pioneering attack on this question was mounted by Peterson and Berger (1975) in their study of diversity in the recorded music industry in the USA. Peterson and Berger defined diversity in terms of the number of different recording companies with titles listed on *Billboard*'s Hot 100 music charts. Their study of data from these charts over the 1950s and 1960s concluded that the diversity of hit records is subject to cyclical oscillations over time. These oscillations appear to depend on the competitive interplay between major recording companies and small independents in final markets. Peterson and Berger observed that when majors dominate the market, product diversity diminishes; when independents are in the ascendant, diversity increases. These phases succeed one another as follows:

1. When the diversity of hit records is at a low level due to majors' control of the market, new niches appear beyond the margins of the current mainstream.
2. Many of these niches are explored by risk-taking independents, and as some of them find popular favor the diversity of entries on the hit parade increases.
3. Some of these niches become commercially lucrative, and the majors will proceed to take them over, thus creating a new mainstream with eventually diminished diversity of products.
4. New margins and exploratory possibilities for risk-taking independents appear; and so on.

This approach has been corroborated by analysts such as Alexander (1996), Christianen (1995) and Gronow (1983). In addition, Burnett (1990, 1992, 1993), Lopes (1992) and Scott (1999b) have pointed out that since the end of the 1970s an independent trend toward increasing diversity in Hot 100 hits has become intertwined with – and is possibly beginning to override – this cyclical process. In all likelihood, this circumstance is due to the growing tendency for majors to absorb independent recording companies as affiliates in the quest for broader market penetration. This is in line with a shift to increasing product diversity within the cultural economy at large, and it represents – up to a point – a move away from the kind of massification that Frankfurt School theorists, such as Adorno (1991), thought was to be the fate of popular culture.

Agglomerations of producers usually follow developmental trajectories shaped by the interpenetrating logic of all their component elements. Accordingly, and in view of their intricate structural make-up, cultural-products agglomerations are evidently subject to the system-wide branching and lock-in processes described by Arthur (1990) and David (1985). Once any agglomeration shifts in a given direction, further evolutionary possibilities are pre-empted in the sense that only derivative branches can now be followed while others become closed off. Depending on the past history of the system, these branches may lead on to further evolutionary possibilities, or to stagnation.

Creativity and innovation are very much at stake as this process works its course. One of the most pervasive forms of lock-in in the types of systems under scrutiny here is the situation exemplified by Maduro's brahmin painters where the codes and styles of cultural performance become so conventionalized that possibilities for change are negligible. It might be argued that much of the French cinema today, with its overarching structure of protective institutions is, paradoxically, in danger of approaching lock-in (Scott, 1999a). Similarly, in the immediate post-war years, the film industry of Hollywood was relatively locked in to a studio system of production, and this greatly impeded it from responding to the competitive pressures now coming from the burgeoning television industry (Maltby, 1981). Under the stimulus of this external threat, the industry painfully restructured over the 1950s and 1960s, above all by pursuing strategies of vertical disintegration. In parallel with these changes, new rounds of creativity and innovation were unleashed in the industry, in both business practices and in the cultural content of films, culminating in the Hollywood of today with its myriad small adaptable firms moving rapidly from project to project and its wholesale embrace of new digital technologies.

These temporal logics are deeply etched onto the geography of the creative field. As they unfold, particular places become dynamically integrated into the system of cultural production as aspects of their identity are assimilated into final products, and as these products in turn define and redefine their places of origin. This relationship can be relatively open-ended or can converge toward lock-in, leading in the one case to continual shifts in the images projected by product and place, and in the other to a mutually reinforcing stability.

Consider the symbiosis between the film industry, qua production apparatus, and Hollywood, the place. The film industry of Hollywood draws on a web of local cultural assets in the form of street scenes, natural landscapes, ways of life and so on. These assets play a crucial role in imparting to the products of the industry their distinctive look and mental associations. But these products also create new images (real or imagined) of Hollywood/Southern California and imbue earlier images with fresh meanings. These are then assimilated back into the region's fund of cultural assets where they become available as inputs to new rounds of production. This dynamic flow of images is also frequently tapped by other cultural-products industries in Southern California, such as music recording, television show production, clothing and so on (Scott, 1996). Equally, industries like high fashion in Paris or country music in Nashville draw upon associations rooted in place while simultaneously helping to make those same places as cultural constructs. 'Nashville', indeed, is virtually entirely a creature of its cultural-products industry (see Peterson, 1975). Las Vegas, another prominent instance of the same phenomenon, is literally a place and a cultural-products industry at one and the same time.

Conclusion

I have sought in this brief review to demonstrate that creativity and innovation in the modern cultural economy can be understood as social phenomena rooted in the production system and its geographic milieu, i.e. the creative field. Three main issues have been investigated in this regard, namely, (a) the formation of cultural communities, (b) the organization of the cultural economy and its propensity to agglomeration and (c) the temporal logic of cultural production complexes. Once again, this claim about the immanence of creativity and innovation does not in any way depreciate the role of the individual as a repository of specific aptitudes and imaginative capacities. Indeed, the individual as the bearer of these endowments is indispensable to the whole process. Creativity and innovation, however, are also imbricated in spatial and temporal fields of social action in the sense that they come into being in organized production environments where the talents and abilities of different individuals assume an interdependent character directed to economic ends.

When the Frankfurt School theorists wrote their pessimistic prognostications about the looming eternal sameness and alienating effects of commercialized culture, they were making inductions from then current trends which they no doubt correctly apprehended. These trends were evident in the deepening corporate control of culture, and its expression in Fordist or proto-Fordist production methods. However, the world has not quite turned out as the members of the Frankfurt School anticipated. For one thing, in the light of increasing evidence that *all* culture grows out of concrete social situations, one might plausibly argue that while large segments of the capitalist cultural economy will always produce dross of one sort and another, there is no reason in principle why other segments cannot function at the leading edges of cultural progress and experimentation (see Featherstone, 1995; Frith, 1996; Garnham, 1987; Negus, 1998; Rowe, 1995). For another, there has been a turn to increasing diversity in many different sectors of the cultural economy, a turn that is likely to be intensified as new cultural-products agglomerations make their appearance in the global economy (Scott, 1998b).

References

Adorno, T.W. (1991) *The Culture Industry: Selected Essays on Mass Culture*. London: Routledge.

Alexander, P.J. (1996) ‘Entropy and Popular Culture: Product Diversity in the Popular Music Recording Industry’, *American Sociological Review* 61: 171–4.

Arthur, W.B. (1990) ‘Silicon Valley Locational Clusters: When do Increasing Returns Imply Monopoly?’, *Mathematical Social Sciences* 19: 235–51.

Barnes, B. (1974) *Scientific Knowledge and Sociological Theory*. London: Routledge & Kegan Paul.

Barnes, B., D. Bloor and J. Henry (1996) *Scientific Knowledge: A Sociological Analysis*. London: Athlone.

Becker, H.S. (1974) ‘Art as Collective Action’, *American Sociological Review* 39: 767–76.

Becker, H.S. (1976) ‘Art Worlds and Social Types’, *American Behavioral Scientist* 19: 703–18.

Becker, H.S. (1982) *Art Worlds*. Berkeley and Los Angeles: University of California Press.

Bordwell, D., J. Staiger and K. Thompson (1985) *The Classical Hollywood Cinema: Film Style and Modes of Production to 1960*. New York: Columbia University Press.

Bourdieu, P. (1971) 'Le Marché des biens symboliques', *L'Année sociologique* 22: 49–126.

Bourdieu, P. (1977) 'La Production de la croyance: Contribution à une économie des biens symboliques', *Actes de la recherche en sciences sociales* 13: 3–44.

Bourdieu, P. (1983) 'The Field of Cultural Production, or: the Economic World Reversed', *Poetics* 12: 311–56.

Bourdieu, P. and Y. Delsaut (1975) 'Le Couturier et sa griffe: Contribution à une théorie de la magie', *Actes de la recherche en science sociales* 11: 7–36.

Burnett, R. (1990) *Concentration and Diversity in the International Phonogram Industry*. Gothenburg Studies in Journalism and Mass Communication 1. Gothenburg, Sweden: University of Gothenburg.

Burnett, R. (1992) 'The Implications of Ownership Changes on Concentration and Diversity in the Phonogram Industry', *Communication Research* 19: 749–69.

Burnett, R. (1993) 'The Popular Music Industry in Transition', *Popular Music and Society* 17: 87–114.

Christianen, M. (1995) 'Cycles in Symbol Production? A New Model to Explain Concentration, Diversity and Innovation in the Music Industry', *Popular Music* 14: 55–93.

Cooke, P. and K. Morgan (1998) *The Associational Economy: Firms, Regions, and Innovation*. Oxford: Oxford University Press.

Crane, D. (1992) *The Production of Culture: Media and the Urban Arts*. Newbury Park, CA: Sage.

Csikszentmihalyi, M. (1990) 'The Domain of Creativity', pp. 190–212 in M.A. Runco and R.S. Albert (eds) *Theories of Creativity*. Newbury Park, CA: Sage.

David, P. (1985) 'Clio and the Economics of QWERTY', *American Economic Review* 75: 332–7.

Denisoff, R.S. and J. Bridges (1982) 'Popular Music: Who are the Recording Artists?', *Journal of Communication* 32: 132–42.

DiMaggio, P. (1977) 'Market Structure, the Creative Process, and Popular Culture: Toward an Organizational Reinterpretation of Mass-Culture Theory', *Journal of Popular Culture* 11: 436–52.

DiMaggio, P. and P.M. Hirsch (1976) 'Production Organizations in the Arts', *American Behavioral Scientist* 19: 735–52.

Featherstone, M. (1995) *Undoing Culture: Globalization, Postmodernism and Identity*. London: Sage.

Frith, S. (1992) 'The Industrialization of Popular Music', pp. 49–74 in J. Lull (ed.) *Popular Music and Communication*. Newbury Park, CA: Sage.

Frith, S. (1996) *Performing Rites: On the Value of Popular Music*. Cambridge, MA: Harvard University Press.

Garnham, N. (1987) 'Concepts of Culture: Public Policy and the Cultural Industries', *Cultural Studies* 1: 23–37.

Gronow, P. (1983) 'The Record Industry: The Growth of a Mass Medium', *Popular Music* 3: 53–75.

Hall, P. (1998) *Cities in Civilization*. New York: Pantheon.

Hennion, A. (1981) *Les Professionels du disque: Une sociologie des variétés*. Paris: Éditions A.M. Métailié.

Hennion, A. (1983) ‘The Production of Success: An Anti-Musicology of the Pop Song’, *Popular Music* 3: 159–93.
Hennion, A. (1989) ‘An Intermediary between Production and Consumption: The Producer of Popular Music’, *Science, Technology and Human Values* 14: 400–23.
Hirsch, P.M. (1969) *The Structure of the Popular Music Industry*. Ann Arbor: Institute for Social Research, University of Michigan.
Hirsch, P.M. (1972) ‘Processing Fads and Fashions: An Organization-Set Analysis of Cultural Industry Systems’, *American Journal of Sociology* 77: 639–59.
Kealy, E.R. (1979) ‘From Craft to Art: The Case of Sound Mixers and Popular Music’, *Sociology of Work and Occupations* 6: 3–29.
Lash, S. and J. Urry (1994) *Economies of Signs and Space*. London: Sage.
Latour, B. and S. Woolgar (1979) *Laboratory Life: The Social Construction of Scientific Facts*. Beverly Hills, CA: Sage.
Lopes, P.D. (1992) ‘Innovation and Diversity in the Popular Music Industry, 1969 to 1990’, *American Sociological Review* 57: 56–71.
Lundvall, B.A. and B. Johnson (1994) ‘The Learning Economy’, *Journal of Industrial Studies* 1: 23–42.
Maduro, R. (1975) *Artistic Creativity in a Brahmin Painter Community*. Research Monograph 14. Berkeley: Center for South and Southeast Asia Studies, University of California.
Malraux, A. (1946) *Esquisse d’une Psychologie du Cinéma*. Paris: Gallimard.
Maltby, R. (1981) ‘The Political Economy of Hollywood: The Studio System’, pp. 42–58 in P. Davies and B. Neve (eds) *Cinema, Politics, and Society in America*. Manchester: Manchester University Press.
Mannheim, K. (1952) *Essays in the Sociology of Knowledge*. Henley-on-Thames: Routledge & Kegan Paul.
Marshall, A. (1919) *Principles of Economics*. London: Macmillan.
Menger, P.M. (1983) *Le Paradoxe du Musicien: Le Compositeur, le Mélomane et l’État dans la Société Contemporaine*. Paris: Flammarion.
Menger, P.M. (1993) ‘L’Hégémonie Parisienne: Économie et Politique de la Gravitation Artistique’, *Annales: Économies, Sociétés, Civilisations* 6: 1565–1600.
Molotch, H. (1996) ‘LA as Design Product: How Art Works in a Regional Economy’, pp. 225–75 in A.J. Scott and E.W. Soja (eds) *The City: Los Angeles and Urban Theory at the End of the Twentieth Century*. Berkeley and Los Angeles: University of California Press.
Montgomery, S.S. and M.D. Robinson (1993) ‘Visual Artists in New York: What’s Special about Person and Place?’, *Journal of Cultural Economics* 17: 17–39.
Mulkay, M.J. (1972) *The Social Process of Innovation: A Study in the Sociology of Science*. London: Macmillan.
Negus, K. (1996) *Popular Music in Theory: An Introduction*. Hanover and London: Wesleyan University Press.
Negus, K. (1998) ‘Cultural Production and the Corporation: Musical Genres and the Strategic Management of Creativity in the US Recording Industry’, *Media, Culture and Society* 20: 359–79.
O’Connor, K. (1991) ‘Creativity and Metropolitan Development: A Study of Media and Advertising in Australia’, *The Australian Journal of Regional Studies* 6: 1–14.
Peterson, R.A. (1975) ‘Single-Industry Firm to Conglomerate Synergistics: Alternative Strategies for Selling Insurance and Country Music’, pp. 341–58 in J.F. Blumstein and B. Walter (eds) *Growing Metropolis: Aspects of Development in Nashville*. Nashville, TN: Vanderbilt University Press.

Peterson, R.A. and D.G. Berger (1975) 'Cycles in Symbol Production: The Case of Popular Music', *American Sociological Review* 40: 158–73.
Peterson, R.A. and P. DiMaggio (1975) 'From Region to Class, the Changing Locus of Country Music: A Test of the Massification Hypothesis', *Social Forces* 53: 497–506.
Rowe, D. (1995) *Popular Cultures: Rock Music, Sport and the Politics of Pleasure*. London: Sage.
Russo, M. (1985) 'Technical Change and the Industrial District: The Role of Inter-Firm Relations in the Growth and Transformation of Ceramic Tile Production in Italy', *Research Policy* 14: 329–43.
Ryan, J. and R.A. Peterson (1982) 'The Product Image: The Fate of Creativity in Country Music Songwriting', in J.S. Ettema and D.C. Whitney (eds) *Individuals in Mass Media Organizations: Creativity and Constraint*. Beverly Hills, CA: Sage.
Scott, A.J. (1984) 'Territorial Reproduction and Transformation in a Local Labor Market: The Animated Film Workers of Los Angeles', *Environment and Planning D: Society and Space* 2: 277–307.
Scott, A.J. (1996) 'The Craft, Fashion, and Cultural-Products Industries of Los Angeles: Competitive Dynamics and Policy Dilemmas in a Multisectoral Image-Producing Complex', *Annals of the Association of American Geographers* 86: 306–23.
Scott, A.J. (1997) 'The Cultural Economy of Cities', *International Journal of Urban and Regional Research* 21: 323–39.
Scott, A.J. (1998a) 'Multimedia and Digital Effects: An Emerging Local Labor Market', *Monthly Labor Review* (March): 30–8.
Scott, A.J. (1998b) *Regions and the World Economy: The Coming Shape of Global Production, Competition, and Political Order*. Oxford: Oxford University Press.
Scott, A.J. (1999a) 'French Cinema: Economy, Policy and Place in the Making of a Cultural-Products Industry', *Theory, Culture & Society*.
Scott, A.J. (1999b) 'The US Recorded Music Industry: On the Relations between Organization, Location, and Creativity in the Cultural Economy', *Environment and Planning A*.
Storper, M. (1996) 'Innovation as Collective Action: Conventions, Products, Technologies', *Industrial and Corporate Change* 5: 761–90.
Storper, M. and S. Christopherson (1987) 'Flexible Specialization and Regional Industrial Agglomerations: The Case of the US Motion Picture Industry', *Annals of the Association of American Geographers* 77: 260–82.
Von Hippel, E. (1988) *The Sources of Innovation*. New York: Oxford University Press.
White, H.C. and C.A White (1965) *Canvases and Careers: Institutional Change in the French Painting World*. New York: Wiley.
Williams, R. (1982) *The Sociology of Culture*. New York: Schocken Books.
Wolff, J. (1981) *The Social Production of Art*. New York: St Martin's Press.