

16TH INTERNATIONAL CONFERENCE
e-SOCIETY 2018

Lisbon, Portugal

14 - 16 April

PROCEEDINGS

Edited by:

Piet Kommers

Pedro Isaías



16th INTERNATIONAL CONFERENCE

e-Society 2018

**PROCEEDINGS OF THE
16th INTERNATIONAL CONFERENCE
e-Society 2018**

LISBON, PORTUGAL

14 - 16 APRIL, 2018

Organised by



international association for development of the information society

Copyright 2018

IADIS Press

All rights reserved

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Permission for use must always be obtained from IADIS Press. Please contact secretariat@iadis.org.

Edited by Piet Kommers and Pedro Isaias

Associate Editor: Luís Rodrigues

ISBN: 978-989-8533-75-3

TABLE OF CONTENTS

FOREWORD	ix
PROGRAM COMMITTEE	xi
KEYNOTE LECTURES	xiii

FULL PAPERS

AN ICT-BASED HIGHER EDUCATION ECOSYSTEM FOR ENHANCING STUDENTS' LEARNING AND OPERATIONAL PERFORMANCE <i>Haydar Moussalem, Bassam Hussein, Mohamad Raad and Amin Haj-Ali</i>	3
MEASURING ON-CAMPUS STUDENT ENGAGEMENT IN THE CONTEXT OF BLENDED-MOOC <i>Fadiyah Almutairi and Su White</i>	11
TANGLE: A BLENDED EDUCATIVE SUITE FOR THE QUANTUM INFORMATION EDUCATION AND DEVELOPMENT OF RESEARCH SKILLS <i>Francisco Javier Delgado-Cepeda</i>	23
ENHANCING READING SKILLS THROUGH ADAPTIVE E-LEARNING <i>Julie Villessèche, Christophe Quaireau, Olivier Le Bohec, Jérémie Nogues, Sandrine Oriez, Anne-Laure Besnard, Fanny De La Haye, Yvonnick Noël and Karine Lavandier</i>	30
IMPLEMENTATION AND OUTCOMES OF A UNIVERSITY COURSE IN STATISTICS FOR BUSINESS UNDER THE FIT BLENDED MODEL <i>Faustino Yesca Martínez and Francisco Javier Delgado-Cepeda</i>	40
SOCIAL MEDIA PLATFORMS: CO-CREATING CODES OF CONDUCT <i>Jenny Eppard, Awil Hashi, Sandra Baroudi and Marlieke Rodjen Helder</i>	47
DESIGNING AN IOT-ENABLED GAMIFIED APP FOR EMPLOYEE ENERGY CONSERVATION AT THE WORKPLACE <i>Dimosthenis Kotsopoulos, Cleopatra Bardaki, Thanasis G. Papaioannou, Stavros Lounis and Katerina Pramatari</i>	55
ORGANIZATIONAL INTEROPERABILITY: AN ASSESSMENT OF THE IMPACT OF THE SEMANTIC LAYER <i>Asmae Rhazale and Bouchaib Bounabat</i>	67
THE TRANSFORMATIONAL POTENTIAL OF CIVIL SOCIETY <i>József Veress</i>	75
MODELLING THE ENROLMENT ESERVICE OF A UNIVERSITY USING MACHINE LEARNING TECHNIQUES <i>Ainhoa Yera, Iñigo Perona, Olatz Arbelaitz and Javier Muguerza</i>	83
MEDIA INNOVATION AND BUSINESS MODELS: THE CASE OF END-TO-END IMMERSIVE AUDIOVISUAL SERVICES <i>Heritiana Ranaivoson and Simon Delaere</i>	92

SPORT TOURISM: IMPROVING VISITORS' EXPERIENCE THROUGH WEB 2.0 TOOLS IN SOCCER CLUBS' MUSEUMS AND STADIUMS <i>Deborah Moraes Zouain, Paola Bastos Lohmann, Marcio Ferreira Bezerra, Gabriela De Laurentis Cardoso, Kaarina Barbosa Virkki and Pedro Henrique Domingues Curvello</i>	103
HEART RATE VARIABILITY ANALYSIS AS A PREVENTIVE INSTRUMENT AGAINST STRESS FOR EMPLOYEES <i>Thomas Keller, Giovanni Annunzio and Elke Brucker-Kley</i>	112
TRUSTING IN THE SECURITY OF PRIVATE MESSAGES – A SURVEY-BASED ANALYSIS OF FINNISH INTERNET USERS <i>Jukka Vuorinen, Aki Koivula and Ilkka Koiranen</i>	120
CREATING COMPETENT TEAMS FOR EMERGENCY MANAGEMENT AND RESPONSE <i>Danilo Hasse, Fernando A. Ostuni Gauthier, Carlos Roberto De Rolt and Gisiela Hasse Klein</i>	129
INTERNET OF THINGS, NEW WAVES TOWARD ORGANIZATIONS <i>Fateme Tavallaei and Mohammad Ali Sarlak</i>	139
THE POSTAL SERVICE IN THE SMART CITY <i>Maria Pavlovskaya and Olga Kononova</i>	147
IMPROVING BIM WITH VR IN CONSTRUCTION <i>Alcínia Zita Sampaio</i>	155
SCALING E-HEALTH: AN ANALYSIS OF THE FUNCTIONS OF THE TECHNOLOGICAL INNOVATION SYSTEM AND THE ROLE OF GEOGRAPHIC PROXIMITY <i>Annemijn van Gorp and Bert Mulder</i>	163
AN IMPACT ASSESSMENT FRAMEWORK FOR ELECTRONIC PATIENT RECORDS MANAGEMENT: A CASE OF NAMIBIA <i>Ronald Karon and Nomusa Dlodlo</i>	171
TEACHING MOBILE DEVELOPMENT WITH APP INVENTOR AND PAIR PROGRAMMING <i>Paolo Musmarra</i>	178
INCLUSIVE JOBS: A PROJECT TO HELP PEOPLE WITH IMPAIRMENT FIND OPPORTUNITIES – HEURISTIC EVALUATION TO DIAGNOSE USABILITY PROBLEMS <i>Felipe Araujo and Adriano Bernardo Renzi</i>	185
PLAYER-CENTRIC ADAPTATION OF A CAR DRIVING VIDEO GAME <i>Boyan Bontchev, Dessislava Vassileva and Dobrin Ivanov</i>	193
CLASSIFICATION OF DRIVING TRAITS USING FUZZY-LOGIC CONTROL <i>Dylan Seychell and Steven Farrugia</i>	201
TOWARDS A MULTI-MODAL IT <i>Thomas Keller, Thierry Baumgartner and Elke Brucker-Kley</i>	209
DNA-BASED USER AUTHENTICATION SCHEMES FOR WIRELESS BODY AREA NETWORK <i>Esma Ergüner Özkoç</i>	217
PERSONAL DATA PROTECTION MODELS: ASPECTS OF OWNERSHIP <i>George Bouchagiar and Maria Bottis</i>	224
VISUALISATION – IT IS NOT THE DATA, IT IS WHAT YOU DO WITH IT <i>Chris Roast, David Patterson and Victoria Hardman</i>	231

CORPORATE ATTITUDES TOWARDS DATA PHILANTHROPY IN SOUTH AFRICAN BANKING, RETAIL AND TELECOMMUNICATIONS <i>Kungela Mzuku and Jean-Paul Van Belle</i>	239
THE EFFECTIVENESS OF USING SOCIAL MEDIA IN GOVERNMENT COMMUNICATION IN UAE <i>Elsayed B. Darwish</i>	247
DIGITAL ANALYTICS IN BUSINESS PRACTICE – USAGE, CHALLENGES AND RELEVANT TOPICS <i>Darius Zumstein and Seraina Mohr</i>	257
CONCENTRIC DIVERSIFICATION: GROWTH WITHOUT LOSING COMPETITIVENESS FOR A SMALL IT FIRM <i>Erick A. Arroyo-Blanco, Alexis E. Boza-Monge and Roberto Cortés-Morales</i>	265
THE DEVELOPMENT OF B2C E-COMMERCE IN MACEDONIA: CURRENT SITUATION AND FUTURE POTENTIAL <i>Shkelqim Nazifi and Adrian Besimi</i>	273
ENABLING SEMANTIC INTEROPERABILITY OF DISEASE SURVEILLANCE DATA IN HETEROGENEOUS HEALTH INFORMATION SYSTEMS IN NAMIBIA AS A STANDARD APPROACH <i>Nikodemus Angula and Nomusa Dlodlo</i>	281

SHORT PAPERS

STUDENT RETENTION AS AN INDICATOR OF QUALITY ASSURANCE IN ECUADORIAN UNIVERSITIES <i>Marco Veloz, Mayra Albán, Idalia Pacheco, Gloria Vizcaíno and Mónica Salazar</i>	291
EVIDENCE-BASED TEACHING FOR PERSONALIZED LEARNING <i>Boong Yeol Ryoo, Shannon H Degenhart, and Gary Bolts</i>	296
EDUCATIONAL DATA ANALYSIS APPLYING A KDD METHODOLOGY <i>William Villegas-Ch, Diego Buenaño-Fernández and Sergio Luján-Mora</i>	301
SERIOUS GAMES AND BOARD GAMES IN THE COMPUTER SCIENCE CLASS <i>Dawid B. Jordaan</i>	306
PROTOTYPING A PERSONAL LEARNING ENVIRONMENT <i>Timothy Arndt</i>	311
"NOBODY KNOWS WHAT THE BODY CAN DO": A COMPLEX APPROACH OF AUTISTIC SPECTRUM DISORDERS <i>Nize Maria Campos Pellanda and Maria de Fátima de Lima das Chagas</i>	315
EXTENDING THE UTAUT TO MEASURE THE ADOPTION OF WEB-BASED KNOWLEDGE SHARING SYSTEMS IN SAUDI UNIVERSITIES <i>Nouf Almjally and Mike Joy</i>	319
ONLINE DATA ABOUT BRAZILIAN JUDICIARY: COMPARATIVE ANALYSIS OF JUSTICE IN NUMBERS 2004-2017 <i>Lahis Pasquali Kurtz, Amanda Nunes Lopes Espiñeira Lemos and Alexandra Krastins Lopes Souto Maior</i>	325

E-READINESS METHODOLOGIES: LITERATURE REVIEW <i>Jaouad Chennouf and Najib Belkhayat</i>	330
PROTECTION OF INTELLECTUAL PROPERTY RIGHTS IN POLAND FROM THE PERSPECTIVE OF POLISH YOUTH <i>Lukasz Tomczyk</i>	335
DATA IN THE CLOUD: WHOSE CONTROL AND WHOSE CONSENT? <i>George Bouchagiar and Maria Bottis</i>	341
SEMANTIC WEB APPROACHES FOR BIG DATA INTEGRATION <i>Danilo Hasse, Fernando A. Ostuni Gauthier, Carlos Roberto De Rolt and Gisiela Hasse Klein</i>	347
INVESTIGATION OF EHEALTH LITERACY AMONG PATIENTS WITH DIABETES <i>Sophie Huey-Ming Guo</i>	353

REFLECTION PAPERS

REFLECTIONS ON DIGITAL IDENTITY IN THE ONTOLOGICAL CYBERSPACE <i>Ziyed Guelmami and François Nicolle</i>	361
PRIVACY IN LOCATION-BASED APPLICATIONS FOR AIRPORT USERS <i>Nayra Rodríguez-Pérez and Pino Caballero-Gil</i>	365
MEDICAL E-LEARNING PLATFORM "IMAGE OF THE WEEK": ADVANCES AND CHALLENGES <i>Luísa Bernardino Valério, Giovanna Vieira Moreira, Laio Bastos de Paiva Raspante, Lucas Augusto Carvalho Raso, Bruno Campos Santos and Viviane Santuari Parisotto Marino</i>	369
TECHNOLOGICAL EXURBIA: LIMINALITY AND KNOWLEDGE CITIES' CONCEPTUAL BOUNDARIES <i>Lucia Elizondo</i>	373
PROJECTION OF KNOWLEDGE SOCIETIES TOWARDS THE PAST <i>Arnulfo Felix Jr.</i>	377

POSTERS

AUTOMATIC REFLECTIVE WRITING ANALYSIS BASED ON SEMANTIC CONCEPTS <i>Huda Alrashidi and Mike Joy</i>	383
LIFELONG LEARNING ON TELEMEDICINE AT MEDICAL INSTITUTE OF RUDN-UNIVERSITY <i>Elena Lukianova, Valery Stolyar, Tatiana Lyapunova, Ekaterina Shimkevich and Maya Amcheslavskaya</i>	386

AUTHOR INDEX

FOREWORD

These proceedings contain the papers and posters of the 16th International Conference on e-Society 2018, organised by the International Association for Development of the Information Society, in Lisbon, Portugal, 14 - 16 April 2018.

The e-Society 2018 conference aims to address the main issues of concern within the Information Society. This conference covers both the technical as well as the non-technical aspects of the Information Society. Broad areas of interest are e-Government / e-Governance, e-Business / e-Commerce, Technology and Society, e-Learning, New Media and E-Society, e-Health, Information Systems, and Information Management. However innovative contributions that don't fit into these areas have also been considered since they might be of benefit to conference attendees.

The e-Society 2018 had 173 submissions from more than 31 countries. Each submission has been anonymously reviewed by an average of 4 independent reviewers, to ensure the final high standard of the accepted submissions. Out of the papers submitted, 34 received blind referee ratings that signified acceptability for publication as full papers (acceptance rate of 20%), while some others were published as short and reflection papers and posters. The best papers will be selected for publishing as extended versions in the Interactive Technology and Smart Education (ITSE) journal (ISSN: 1741-5659) and also in the IADIS International Journal on WWW/Internet (ISSN: 1645-7641).

In addition to the presentation of full, short and reflection papers and posters, the conference also includes one keynote presentation. Special thanks go to Professor Chee Kit Looi (Learning Sciences & Technologies Academic Group Founding Head, Learning Sciences Lab, Nanyang Technological University (NIE/NTU), Singapore), for accepting our invitation as keynote speaker.

As we all know, a conference requires the effort of many individuals. We would like to thank all members of the Program Committee for their hard work in reviewing and selecting the papers that appear in this book. We would also like to thank all the authors who have submitted their papers to this conference. We wish to thank all members of our organizing committee.

Last but not least, we hope that everybody has a good time in Lisbon, and we invite all participants for next year's edition of the International Conference on e-Society.

Piet Kommers, University of Twente, The Netherlands
Program Chair

Pedro Isaias, The University of Queensland, Australia
Conference Chair

Lisbon, Portugal
April 2018

PROGRAM COMMITTEE

PROGRAM CHAIR

Piet Kommers, University of Twente, The Netherlands

CONFERENCE CHAIR

Pedro Isaias, The University of Queensland, Australia

COMMITTEE MEMBERS

Adam Wojciechowski, Poznan University of Technology, Poland

Adam Wojtowicz, Poznan University of Economics and Business, Poland

Agapito Ledezma, Universidad Carlos III de Madrid, Spain

Albena Antonova, University of Sofia, Bulgaria

Aleksandar Dimov, University of Sofia St. Kliment Ohridski, Bulgaria

Alfredo Milanim, University of Perugia, Italy

Alfredo Pina Calafi, Public University of Navarra, Spain

Andrej Brodnik, University of Primorska, Slovenia

Ann M. Shortridge, University of Oklahoma-HSC, USA

Arif Altun, Hacettepe University, Turkey

Arthur Tatnall, Victoria University, Australia

Atilla Elçi, Aksaray University, Turkey

Bardo Fraunholz, Deakin University, Australia

Basilio Pueo, University of Alicante, Spain

Bill Davey, RMIT University, Australia

Boong Ryoo, Texas A&M University, USA

Boyan Bontchev, Sofia University St. Kl. Ohridski, Bulgaria

C K Lok, HKCAAVQ, Hong Kong

Cesar Alberto Collazos, University Of Cauca, Colombia

Charles Owen, Michigan State University, USA

Christian Russ, University Klagenfurt, Austria

Demetrios Sampson, University of Piraeus & CERTH, Greece

Denise Whitelock, The Open University, United Kingdom

Denize Piccolotto Carvalho, Universidade Federal do Amazonas, Brazil

Despina Garyfallidou, University of Patras, Greece

Dessislava Petrova-antonova, University of Sofia, Bulgaria

Dessislava Vassileva, Sofia University "St. Kliment Ohridski", Bulgaria

Dimitris Alimisis, School of Pedagogical and Technological Education, Greece

Dorina Kabakchieva, University of National and World Economy, Bulgaria

Efthymios Alepis, University Of Piraeus, Greece

Egbert Sánchez Vanderkast , National Autonomuos University of Mexico, Mexico

Elpida Tzafestas, University of Athens, Greece

Eltahir Kabbar, ADWC - HCT, United Arab Emirates

Euclid Keramopoulos, Alexander Technological Educational Institute of Thessaloniki,
Greece

Eugenia Kovatcheva, University of Library Studies and Information Technologies,
Bulgaria

Fausto Fasano, Università Degli Studi del Molise, Italy

George S. Ioannidis, University of Patras, Greece

Giuseppe Lugano, University of Zilina, Slovakia

Gladys Benigni, Universidad de Oriente - Nueva Esparta, Venezuela

Gustavo Rossi, Lafia - Universidad Nacional de la Plata, Argentina

Haluk Unaldi, Paradigma Consultancy, Turkey

Hans-Dieter Zimmermann, FHS St. Gallen University of Applied Sciences, Switzerland

Hiroyuki Mitsuhashi, Tokushima University, Japan

Inga Hunter, Massey University, New Zealand

Isabel Solano Fernández, University of Murcia, Spain

Jarogniew Rykowski, Poznan University of Economics, Poland

Javier Sarsa, Universidad de Zaragoza, Spain

Jeanne Schreurs, Hasselt University, Belgium

Jeff Byford, University of Memphis, USA

Jiying Zhao, University of Ottawa, Canada

Joumana Boustanly, DICEN-IdF, Université Paris-Est Marne-la-Vallée, France

Kamelia Stefanova, University of National and World Economy, Bulgaria

Katerina Kabassi, TEI of the Ionian Islands, Greece

Kateryna Synytsya, International Research And Training Center, Ukraine

Laura M. Castro, University of A Coruña, Spain

Levent Yilmaz, Auburn University, USA

Lourdes Villalustre, Universidad de Oviedo, Spain

M. Esther Del Moral, Universidad de Oviedo, Spain

Manuel Gentile, The National Research Council (CNR), Italy

Manuel Gomez Olmedo, University of Granada, Spain

Mario Allegra, Cnr-itd, Italy

Mario Vacca, Italian Ministry of Education, Italy

Michele Risi, University of Salerno, Italy

Mieczyslaw Owoc, University of Economics Wroclaw, Poland

Oleksiy Voychenko, International Research and Training Center, Ukraine

Piotr Gawrysiak, Warsaw University of Technology, Poland

Rafał Wojciechowski , Poznan University of Economics, Poland

Rajavelu Loganathan, Abu Dhabi Women's College, United Arab Emirates

Rajendra Kumar, Massachusetts Institute of Technology, USA

Rob Van Kranenburg, Sociotal.eu, Belgium

Roger Hartley, University of Leeds, United Kingdom

Ryo Takaoka, Yamaguchi University, Japan

Takao Terano, Tokyo Institute of Technology, Japan

Tassos Mikropoulos, University of Ioannina, Greece

Tereza Trencheva, University of Library Studies and Inform. Technoli, Bulgaria

Tobias Meisen, RWTH Aachen University, Germany

Vasiliki Spiliotopoulou, ASPAITE Patras, Greece

Vincent Ng, The Hong Kong Polytechnic University, Hong Kong

Vincenzo Deufemia, Universita' Degli Studi Di Salerno, Italy

Wojciech Cellary, Poznan University of Economics and Business, Poland

Yongjian Fu, Cleveland State University, USA

KEYNOTE LECTURE

WHITHER SEAMLESS LEARNING: PERSPECTIVES, CHALLENGES AND OPPORTUNITIES

**By Professor Chee Kit Looi, Learning Sciences & Technologies Academic Group
Founding Head, Learning Sciences Lab,
Nanyang Technological University (NIE/NTU)
Singapore**

Abstract

In this talk, I will attempt a review of the latest state of research and development in seamless learning. Seamless learning is: "... when a person experiences a continuity of learning, and consciously bridges the multifaceted learning efforts, across a combination of locations, times, technologies or social settings." Over the past ten over years, there have been different perspectives, visions, and practices on seamless learning from various diversified (e.g., techno-centric, techno-pedagogical, socio-situated and socio-cultural; informal learning and formal learning; crossover learning, incidental learning and context-based learning) perspectives. While there are researchers in mobile learning, learning sciences, and educational technologies who have been explicitly researching on seamless learning, they are also those who did not think they were studying seamless learning but have indeed accomplished some research work that resembles the spirit and the salient characteristics of the learning approach. It is opportune to review the work in all these fields to inform us of the designs and learning mechanisms for seamless learning. I will discuss their theorization of learning, and the conditions, resources and frameworks for understanding such learning, and probe their relevance and connections to seamless learning.

Full Papers

AN ICT-BASED HIGHER EDUCATION ECOSYSTEM FOR ENHANCING STUDENTS' LEARNING AND OPERATIONAL PERFORMANCE

Haydar Moussalem¹, Bassam Hussein², Mohamad Raad³ and Amin Haj-Ali³

¹*Department of Information Technology*

²*Department of Industrial Engineering*

³*Department of Computer and Communication Engineering
Lebanese International University, Beirut, Lebanon*

ABSTRACT

As student-centered institutions, universities set their missions to enhance the learning experience and thus profile of their students targeting lifelong self-actuated graduates. Nowadays, the world of education has witnessed a great leap in the use of information and computer technology (ICT). ICT based tools are considered as important enablers for students' learning. Generally, they contribute to the enhancement of students' learning experience by increasing their efficiency and effectiveness. When students are well prepared and qualified, they enhance their professional performance efficiently and creatively. However, there are so many available ICT tools that provide a variety of features and functions both on the institutional and student levels. Selecting or implementing an appropriate and suitable set of tools is not always obvious and easy. Universities are under great pressure to become more efficient and provide increased value. They must look at new tools and approaches to reach students. This paper presents one part of an online university management system (UMS) that aims to increase the efficiency and effectiveness of teaching at the largest and fastest growing private university in Lebanon; the Lebanese International University, LIU. An effective system should necessarily have two educational interfaces, one for faculty and the other for students. The interface for faculty is one part of the UMS that has been implemented and will be the subject of this paper.

KEYWORDS

Information and Computer Technology, Higher Education, Student Learning Experience, On-Line Education Tools

1. INTRODUCTION

Many higher education institutions are facing challenges regarding the integration of student learning techniques into technology systems. There are technologies that help establishing a complete ICT based environment to help improve the students' learning while ensuring the quality of education. Recent trials to categorize discrete software components and their usage in learning and operational performance have been merely map mobile applications to parts of the "Knowledge Domains Hierarchy", yet it does not present or propose any coherent and homogeneous eco-system supporting this mapping, neither does it describe a working mechanism that prop the objective of learning and/or operational performance. The focus of the system described in this paper is to enable instructors to create targeted content. As such, the system must allow the association of content with course objectives and to associate courses with programs and groups of programs. The content must also adhere to the knowledge domains hierarchy pyramid (see Figure 1).



Figure 1. Knowledge domains hierarchy pyramid

The main functionalities of the system described in this paper are: Knowledge domain management, major management and association of majors with knowledge domains, and course management and association of courses with majors. The system is built to allow the collection of feedback by defining outcomes for each activity, course, major and domain and then tracking how well those outcomes have been achieved through surveys and student achievement monitoring. Student achievement is weighted according to the "knowledge domains hierarchy pyramid" with more emphasis placed on creative achievement.

2. INTEGRATING ICTS IN HIGHER EDUCATION INSTITUTIONS

Educational researchers as well as scholars and teachers are still concerned in incorporating technology in higher education. Even a few believe that it is risky in the use education, for instance, computer-based education could restrict the decent variety of instructive ways to deal with the main model of technological-based performance (Sasseville, 2004). Moreover, teachers not only should be trained to use technology but also they need the support of human and technical resources in order to attain progressive technology integration, otherwise, they would not be able to integrate technology into their daily instructions (Zhao & LeAnna Bryant, 2006) (Cagiltay, Yildirim, & Aksu, 2006).

Universities have sought to find alternative methods to teaching and learning to cope with such sophistication like tackling a wide range of online courses. Applications such as PowerPoint have some little influence on learning teaching strategies while other educational technologies like virtual learning environments (VLEs) have a considerable influence on teaching/learning strategies and on the educational institution itself. It is evident that the educational technology that plays a series of mixed and additional functions in education (Ezziane, 2007) (Kundi & Nawaz, 2010).

E-learning has undergone thorough and dramatic changes and stages over the past thirty years. Although researchers have not provided any specific learning advantages of e-learning, the use of computers in universities and educational institutions has been progressively expanding (Nawaz, Kundi, & Shah, 2007) (Cagiltay, Yildirim, & Aksu, 2006). The real problem is not from the systems itself, but weak development of information infrastructure of such universities. In fact, it seems that these universities lack adequate distribution of information organizational structure (Sife, Lwoga, & Sanga, 2007).

3. LEVELS OF INTEGRATION

Different cultures of education can go through various periods of development with respect to change, progress, gradual decrease or without any change (Aaron, Dicks, Ives & Montgomery, 2004). According to Dinevski and Kokol (Dinevski & Kokol, 2005), e-learning may be faced with changes from:

1. Sequential learning arranged along a straight line to hypermedia learning
2. Guidance-based learning to construction and discovery
3. Education focused on teacher to education focused on student
4. Grasping material to know how to command and how to learn
5. Get educated through school to learn as long as you live
6. The same kind of learning for all to modified learning
7. Learning as kind of punishment to learning for pleasure
8. The teacher is only a transferring person of material to the teacher as an assistant

Multiple levels of technology should be integrated into any e-learning system whatever the number and kind of ICT related establishments. Advancing e-learning must be lined up with the developer and user needs (Carey & Gleason, 2006) and the user practice and experience must be based on analysing the user requirements (Sasseville, 2004) (Zhao & LeAnna Bryant, 2006) (Graff, Davies, & McNorton, 2001).

3.1 Integration at the Planning Levels

Aaron et al. stated that ICTs integration requires identifying clear goals on the pedagogical level in order to set a successful plan based on decisions made regarding technological needs (Aaron et al, 2004). However, it

is neither easy nor simple to create and execute an action plan that includes educational technology. In fact, the process of developing a strategic plan as well as implementing it is too complicated. Educational technology has a framework of basic organizational structures and a system of methods and techniques that should be all conducted within the teaching and learning environment. Through using technology and other resources collaboratively, instructors need to determine teaching requirements, plan for a strategic plan and implement it according to new demands, and assess classroom instruction (Willis, 2006).

Planners often know that every successful strategic plan should have a clear mission and vision. Therefore, integrating technology into a strategic plan of an educational institution should be related to the mission and vision of the institution itself and the strategy should be connected to its culture, values and history. Teaching technologies should be integrated systematically by using a set of various instruments, mechanisms and methods. Since e-learning improvement is an activity related to more than one branch of knowledge, it involves many inter-dependant participants and processes. Thus, institutions must set an educational strategy to enhance plans including IT integration to meet various other plans (Ezziane, 2007).

3.2 Integration at the Development Levels

Since learners may differ in nationality, gender, and the style of learning, student characteristics should be taken into consideration through setting and designing e-learning systems to avoid facing difficulties regardless that educators can benefit from these systems at the theoretical level (Graff, Davies, & McNorton, 2001). In spite of the fact that technology designers are in charge of streamlining the usage of their items, the educational training group such as instructors, directors, locale administrators, lawmakers, and so forth must be dependable buyers and clients of technological products of education (Buzhardt & Heitzman-Powell, 2005). However, technology developers face some challenges in application communication. In other words, sometimes it is hard to let application integrate and intercommunicate (Nawaz, Kundi, & Shah, 2007) (Carey & Gleason, 2006).

Programming designers need to share models and patterns of their own systems and exchange several branches of knowledge with instructors and students. There is no doubt that the educational institution would alter its teaching methods and shifts from lecturing to instructing through using technology after gaining experience and building up its pedagogical structure over several years. Therefore, the advanced development of mode of teaching is related to the number of years of the institution existence according to statistics (Nawaz & Kundi, 2010) (Mehra & Mital, 2007).

3.3 Integration at the Use Levels

Students' characteristics such as learning style, nationality, and gender differ between one learner and another. These characteristics should be taken into account through designing e-learning systems otherwise the educational team in the institution would face many challenges (Graff, Davies, & McNorton, 2001). Instructors are not completely able to use ICTs in their teaching; they are used to depend on traditional educational ways in classrooms. It seems that teachers need more time and training to properly integrate educational technologies in their classrooms (Sasseville, 2004). Although the educational sector largely invests educational media for learners, they lacked knowledge about how to use such media in education.

It is true that technical training for technology integration is essential to achieve ICTs goal, but it makes no sense without adding many other aspects of e-learning (Zhao & LeAnna Bryant, 2006). E-learning is applied in many different sectors. For example, the economic meaning defines it as a sector of e-business whereas the educational defines it as an approach for developing new instructional styles of teaching and learning. Thus, there is a relationship between e-learning and the academic teaching environment.

4. PROBLEMS OF INTEGRATION

Instructors and educational program designers have been working out for testing the different uses of technology in education. Despite their hard attempts, research shows that scholars and instructors aren't completely able to incorporate technology into their practices and they feel uncomfortable with technology integration. Some teachers believe that it seems hard to integrate ICT into their classrooms because it restricts

them from having the opportunity to use diverse educational approaches in their academic systems (Sasseville, 2004). Whereas others believe that using e-learning in higher education holds both advantages and disadvantages for teachers and learners (Valcke, 2004). However, HEIs are still facing obstacles in ICTs integration due to the lack of advanced development, technical and management support, awareness toward ICTs, and absence of programs and systems for implementing ICTs properly (Sife, Lwoga, & Sanga, 2007).

Colleges have been exploring different avenues regarding the utilization of innovation. There are still difficulties and concerns with respect to instructor's capacity to incorporate innovation into instructing and learning exercises, regardless of their endeavours, as declared by Oh and Russell. Buzhardt et al., noted some place that the marriage amongst training and innovation has frequently been rough. Sife et al. reported that the colleges still face a great deal of difficulties in attempted such a procedure like absence of fundamental way to deal with ICT execution, mindfulness and state of mind towards ICTs, regulatory help, specialized help, changing advanced education, self-improvement (Buzhardt & Heitzman-Powell, 2005).

Borrowed Models: Developing countries lack the models of development and borrow most of them from other developed countries in the West. Integrating new technologies within several environments and various types of resources lead to some difficulties. For instance, the majority of educational scholars of technology in China use the American models of technology settings although they have a large number of resources. This means that it is dispensable to handle all the essential factors that are necessary to develop and implement e-learning in education. Therefore, it is essential to be skilful in understanding these differences and be able to adapt them in the e-learning environments of each institution (Nawaz & Kundi, 2010).

Diversity of perceptions: In order to enhance technology-based learning programs within any educational institution, it is critical to recognize the ability of teachers to be aware of technology integration and how it influences the classroom (Zhao & LeAnna Bryant, 2006). There are several factors that have an impact on instructional practices such as the access to hardware, the experience of using technology-based instruction, the understanding of using motivational interaction, methodical teaching, creativity, faculty improvement at the intellectual level as well as demographic factors (Mehra & Mital, 2007) (Mehra & Mital, 2007). It is obvious that e-learning users extremely differ in their attitudes and perceptions; some are supporters while others are protestors. But many others are in between (Kundi & Nawaz, 2010).

Level of importance and relevance of ICTs integration depends on subject material. According to researches, teachers of humanities and social studies are less likely to use technology or computers in the classroom than mathematics and science teachers (Wims & Lawler, 2007). E-learning users decide the level and nature of integrating ICTs in higher education. Therefore, attitudes differ because the characteristics of individual, staff, and institutions differ (Mehra & Mital, 2007).

The role of educational technology as an instrument: ICTs integration should be based on theories of development and educational enhancement related to a technology approach, taking into consideration a large diversity of factors like political, social, cultural and economic ones. The problem is that development approaches focus on the technical dimension of the E-learning process and ignore others. Researchers still argue about whether education should be instrumental or technical. In other words, IT education seems to be ineffective because it focuses on the intellectual sides and ignores the job market demands (Ezer, 2006).

Some researchers support the theory which views technology as just an "instrument" ignoring its importance to instructional approaches and the relation between ICTs and work market. However, education serves society; the role of students is to relate their learning to their daily life and apply their educational knowledge properly within the community. Thus, students seem to be contributors to the society. They should be assisted to attain both personal and communal progress (Kundi & Nawaz, 2010) (Ezer, 2006).

Insufficient and inappropriate training: Teachers would never be able to integrate technology into curriculum and attain real progress and instructional advancement without adequate training of how to use ICTs in education and assist them to develop their skills and attitudes toward technology. Research reveals that sufficient and appropriate training is a main factor to achieve advanced technology integration (Zhao & LeAnna Bryant, 2006). Since teachers are important players in assessing the needs of students and classroom instruction, they need adequate training that help them to identify students' needs, plan to integrate technology into curriculum, and implement such instructional integration in their classrooms. Unlike IT programmers and designers, teachers are not in need neither to software application training nor to skill development to integrate advanced technology (Willis, 2006).

Problems of e-learning implementation: ICTs implementation is a complex process that involves three main steps: (a) select the most adequate type of ICT, (b) modify the new use of ICTs that fit with education, (c) alter implementation methods of teaching through ICTs (Nyvang, 2006). Challenges include being

unsatisfied with the new technological systems, incompatibility of new e-learning models with the existing activities, inadequate training assistance for users, and ignoring the complex process of technology by employees (Bondarouk, 2006). Integrating ICTs in higher education has impacts on students, teachers, administrators, and leaders; each of them should be adapted to the new technological practices. Therefore, working or learning within a network is not an easy task in higher education (Nyang, 2006).

5. IMPLEMENTATION OF AN ICT-BASED UNIVERSITY ECOSYSTEM

ICT-based university ecosystem is a complete university information system that will enhance the students' learning experience and will provide academic and administrative university management with accurate information analysis and statistical data to support advanced evidence-based decisions. Furthermore, the ICT-based university ecosystem will support measuring students' performance and course learning outcomes. It will contribute to academic planning and monitoring by providing relevant institutional information in various areas of concern. A prototype of such a system is being implemented and it is being used to rapidly develop sketches of particular processes or components. This is particularly useful in the case where the user requirements are not well understood or are difficult to formally determine (Hussein, Hammoud, Bazzi, & Haj-Ali, 2014). The ICT-based university ecosystem will include:

- 1. Course Digital Materials:** Shared materials prepared by instructor and shared to students as shown in Figure 2.

Figure 2. Course material sharing view

- 2. Online Course Syllabus:** This interface provides automation for producing the course syllabus according the pre-defined template. This includes course description, learning outcomes, covered topics, textbook and references and the breakdown of topics chronologically as shown in Figure 3.

Semester Timeline			
Number of Weeks:		Sessions/Week:	
Week	Session	Book Reference	Topic
Week 1	Session 1	Book w1 s1	Topic w1 s1
	Session 2	Book w1 s2	Topic w1 s2
Week 2	Session 1	Book w2 s1	Topic w2 s1
	Session 2	Book w2 s2	Topic w2 s2
Week 3	Session 1	Book w3 s1	Topic w3 s1
	Session 2	Book w3 s2	Topic w3 s2

Figure 3. Chronological breakdown of course topics

- 3. Course Assessments Associated with Course Activities:** Needed exam elements may be entered (**A** in the Figure 4); the summation of all elements must be equal to 100 (Here 70+30). For each element, outcomes are added and the sum of outcomes grades must be equal to element percentage (Here 40+30=70). When done, the user may click on “Add Exam Structure” to create the structure.

Information	
Course Code:	CSCI205
Course Title:	Computer Science Overview
Exam Title:	Attendance
Exam Percentage:	10.00
Final:	No

Exam Elements(100)			
Element Title	Percentage	Add	Remove
Part 1	70		A
Outcome	Grade	Add	Remove
Outcome2	40		B1
Outcome1	30		B2
Part 2	30		A
Outcome	Grade	Add	Remove
Outcome2	20		
Outcome3	10		

Add Exam Structure **Close**

Figure 4. Course assessment entry

- 4. Alignment of Course Learning Outcomes with Program Outcomes:** All course activities and their associated elements should map to the course learning outcomes which in turn should map to the program outcomes. The system ensures that this is done as depicted in Figure 5.

Information	
Major:	Computer Science (CSCI)
Course:	Arabic Language and Literature (ARAB200)
Type:	GER
Credits:	3

Outcome Relation					
Course Outcome	Major Outcome	Knowledge Domain	Level	Add	Remove
--Select Course Outcome	Computing Skills Solve Problems Interpretive Skills Ethics and Etiquette Work Effectively Addresses Problems	--Select Knowledge Domain	--Select Level		
--Select Course Outcome	Computing Skills Solve Problems Interpretive Skills Ethics and Etiquette Work Effectively Addresses Problems	--Select Knowledge Domain	--Select Level		

Update Relation **Close**

Information	
Major:	Computer Science (CSCI)
Course:	Introduction to Computers (CSCI200)
Type:	GER
Credits:	3

Outcome Relation			
Course Outcome	Major Outcome	Knowledge Domain	Level
Components of a computer system	Computing Skills Solve Problems Interpretive Skills Ethics and Etiquette Work Effectively Addresses Problems Produce Technical Document Effective Presentation	KNOWLEDGE	High
Hardware and Software components	Computing Skills Solve Problems Interpretive Skills Ethics and Etiquette Work Effectively Addresses Problems Produce Technical Document Effective Presentation	KNOWLEDGE	High
Computing machines	Computing Skills Solve Problems Interpretive Skills Ethics and Etiquette Work Effectively Addresses Problems Produce Technical Document Effective Presentation	KNOWLEDGE	High

Figure 5. Course assessment entry

- 5. Evidence-based decisions:** Figure 6 displays a graphical chart that will provide statistics on the course activities and how the students fared.

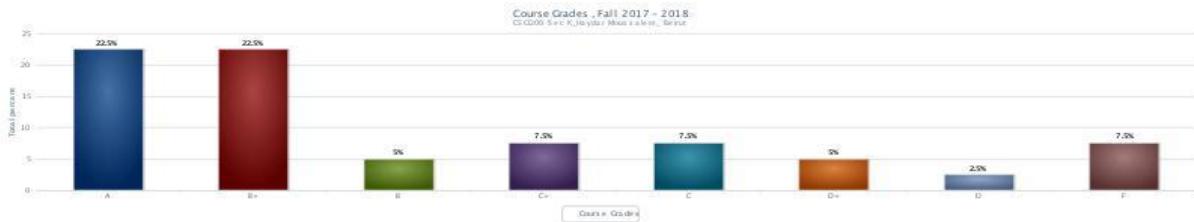


Figure 6. Letter grades distribution

Historical data over a certain interval of time which is settable by the user may also be displayed to view and analyze trends and make adjustments and decisions accordingly as shown in Figure 7.



Figure 7. Trend of letter grades

Similarly, the trends of results, grades percentages and curves, and course enrolment over the years may be displayed to monitor variations. The administration's approval of grades may also be displayed.

6. CONCLUSION

The ICT-based higher education ecosystem for enhancing students' learning and operational performance is very useful for multi-campuses universities, comprising a variety of useful functionalities that provide a clear, accurate and valid assessment for both instructors and stakeholders in key areas. Some functionality of the implemented system was presented including the ability of instructors to share resources with students, and course coordinators to build a dynamic course syllabus, the assessment measures alignment with course learning outcomes, and the alignment of course outcomes with program outcomes. Tools that can help in the decision making were also presented and briefly discussed.

REFERENCES

- Aaron,M., Dicks,D., Ives,C., and Montgomery, B. (2004). "Planning for Integrating Teaching Technologies," Canadian Journal of Learning and Technology, vol. 30, no. 2
- Bondarouk, T. V. (2006). "Action-oriented group learning in the implementation of information technologies: results from three case studies," European Journal of Information Systems, vol. 15, p. 42-53
- Buzhardt, J. and Heitzman-Powell, L. (2005). "Stop blaming the teachers: The role of usability testing in bridging the gap between educators and technology," Electronic Journal for the Integration of Technology in Education, vol. 4, no. 13
- Buzhardt, J. and Heitzman-Powell, L. (2005). "Stop blaming the teachers: The role of usability testing in bridging the gap between educators and technology," Electronic Journal for the Integration of Technology in Education, vol. 4, no. 13
- Cagiltay, N. E., Yildirim, S. and Aksu, M. (2006). "Students' Preferences on Web-Based Instruction: linear or non-linear," Journal of Educational Technology & Society, vol. 9, no. 3, pp. 122-136
- Carey, P. and Gleason, B. (2006). "Vision 2010: The Future of Higher Education Business and Learning Applications," Innovate Journal of Online Education, vol. 3, no. 1.

- Dinevski, D. and Kokol, D. P. (2005). " ICT and Lifelong Learning," European Journal of Open, Distance and E-Learning
- Ezer, J. (2006). " India and the USA: A Comparison through the Lens of Model IT Curricula," Journal of Information Technology Education, vol. 5
- Ezziane, Z. (2007). "Information Technology Literacy: Implications on Teaching and Learning," Journal of Educational Technology & Society, vol. 10, no. 3, pp. 175-191
- Graff, M., Davies, J. and McNorton, M. (2001). "Cognitive Style and Cross Cultural Differences in Internet Use and Computer Attitudes," European Journal of Open, Distance and E-Learning
- Hussein, B., Hammoud, M., Bazzi, H. and Haj-Ali, A. (2014). "PRISM-Process Reengineering Integrated Spiral Model: An Agile Approach to Business Process Reengineering (BPR)," International Journal of Business and Management, vol. 9, no. 10, pp. 134-142
- Kundi, G. and Nawaz, A. (2010) "From objectivism to social constructivism: The impacts of information and communication technologies (ICTs) on higher education," Journal of Science and Technology Education Research, vol. 1, no. 2, pp. 30-36
- Mehra, P. and Mital, M. (2007). "Integrating technology into the teaching-learning transaction: Pedagogical and technological perceptions of management faculty," International Journal of Education and Development using Information and Communication Technology, vol. 1, p. 3
- Nawaz, A. and Kundi, G. (2010). "Predictor of e-learning development and use practices in higher education institutions (HEIs) of NWFP, Pakistan," Journal of Science and Technology Education Research, vol. 1, no. 3, pp. 44-54
- Nawaz, A., Kundi, G. and Shah, B. (2007). "Metaphorical interpretations of information systems failure," Peshawar University Teachers Association Journal (PUTAJ), vol. 14, no. 15-25
- Nyvang, T. (2006). "Implementation of ICT in higher education as interacting activity systems," in The Fifth International Conference on Networked Learning 2006, Lancaster
- Sasseville, B. (2004). "Integrating Information and Communication Technology in the Classroom: A Comparative Discourse Analysis," Canadian Journal of Learning and Technology, vol. 30, no. 2
- Sife, A. S., Lwoga, E. T. and Sanga, C. (2007). "New technologies for teaching and learning: Challenges for higher learning institutions in developing countries," International Journal of Education and Development using ICT, vol. 3, no. 1
- Valcke, M. (2004). "ICT in higher education: An uncomfortable zone for institutes and their policies. In Beyond the comfort zone," in Proceedings of the 21st ASCILITE Conference
- Willis, J. (2006). "Creating a Working Model for Technology Integration Through a Lesson Planning WebQuest," Electronic Journal for the Integration of Technology in Education, pp. 25-33
- Wims, P. and Lawler, M. (2007). "Investing in ICTs in educational institutions in developing countries: An evaluation of their impact in Kenya," International Journal of Education and Development using ICT, vol. 3, no. 1
- Zhao, Y. and Bryant., F. L. (2006). "Can Teacher Technology Integration Training Alone Lead to High Levels of Technology Integration? A Qualitative Look at Teachers' Technology Integration after State Mandated Technology Training," Electronic Journal for the Integration of Technology in Education, vol. 5, pp. 53-62

MEASURING ON-CAMPUS STUDENT ENGAGEMENT IN THE CONTEXT OF BLENDED-MOOC

Fadiyah Almutairi¹ and Dr. Su White²

University of Southampton, UK

¹34 chamberlain road, Southampton, SO17 1PS

²Building 32, Room 3035, Web and Internet Science, ECS

ABSTRACT

This paper aims to develop a model of measuring student engagement in a blended-MOOC context. Massive Open Online Courses (MOOCs) are those that are delivered, usually by leading universities, with a promise to provide free high-quality education to an unlimited number of learners. They offer an opportunity for ‘blended’ course design, where instructors can integrate MOOC content within face to face educational activities and components. Three internationally recognized existing frameworks, the National Survey of Student Engagement (NSSE), UK Engagement Survey (UKES) and Student Engagement Questionnaire (SEQ) were selected as a theoretical framework. The model has nine indicators which provide information about distinct aspects of student engagement. These indicators were confirmed by a panel of 35 experts. A trial evaluation was conducted as a pilot study with 13 students who had taken a course in the blended MOOC format.

KEYWORDS

Blended learning, higher education, student engagement, MOOCs

1. INTRODUCTION

Massive Open Online Courses (MOOCs) are online courses predominantly produced by elite universities with the promise of providing free, high-quality education to an unlimited number of learners (Johnson et al. 2013; UK Universities 2013; Zhang 2013).

A recent development of MOOCs has been the integration of MOOC elements into campus courses in the form of blended learning to make the most of classroom time for activities such as discussions, hands-on activities or working in a group project (Chen 2013). This kind of blended learning has been termed as “distributed flip” (Caulfield et al. 2013), blended / hybrid model (Bruff et al. 2013; Griffiths et al. 2014; Holotescu et al. 2014) or blended MOOCs (LaMartina 2013; Israel 2015). MOOCs themselves have also evolved to blend the use of social networking tools in order to further stimulate engagement and interactions between learners.

The use of MOOC systems in blended contexts has the potential capacity to influence higher education in a variety of ways.

- For students, the systems can influence and shape or structure their approaches to learning and may stimulate class communication;
- For educators, MOOCs may assist the development and selection of online resources and change traditional teaching practices;
- For institutions and researchers, MOOCs can provide large data sets which can be analysed and used to investigate more deeply the processes of learning and learner behaviour.

If MOOC systems are changing teaching practices and campus learning environments, it is reasonable to assume that they are affecting the way students engage with their courses. In spite of this possibility, very little is known about how to measure student engagement in a context where MOOC materials are integrated. This study responds to the need to develop a model that can be used to explore student engagement in higher education within the blended-MOOC contexts.

1.1 Defining Student Engagement

A few decades ago, there was a focus on the concept of student engagement to identify what constitutes an effective learning environment (Coates 2007). A review of the literature on student engagement in higher education has shown that there is a number of student engagement definitions (Barkley 2009). In 1984, Astin (1984) defined student engagement as '*the amount of physical and psychological energy that the student devotes to the academic experience*'.

Then, Chickering and Gamson published 'The Seven Principles for Good Practice in Undergraduate Education'. These principles are used to improve undergraduate education. Chickering and Gamson found that effective educational practice includes: student-faculty contact; cooperation among students; active learning; prompt feedback; time on task; high expectations; and respect for diverse talents and ways of learning.

In 2000, the National Survey of Student Engagement (NSSE) was developed as a lens to examine the quality of the student learning experience (National Survey of Student Engagement 2007). The NSSE defines student engagement as '*the time and energy students devote to educationally sound activities inside and outside of the classroom, and the policies and practices that institutions use to induce students to take part in these activities*' (Kuh 2003). This definition of student engagement is grounded in several decades of prior studies, and specifically in four key research: 'quality of effort' concept (Pace 1984); 'theory of involvement' (Astin 1984); 'seven principles of good practice in undergraduate education' (Chickering & Gamson 1987) and (Pascarella & Terenzini 2005) causal model of learning and cognitive development.

From these studies, five benchmarks of effective educational practice have been identified, namely:

1. **Active and collaborative learning:** this benchmark is designed to measure the extent to which students interact and work with other students, and the extent of student participation in educational purposeful activities inside and outside of the classroom.
2. **Student interactions with faculty members:** This benchmark is developed to assess student and faculty interaction. This includes participation with teaching staff inside and outside the class, and the level of participation with them in regard to performance, assignments and career plan.
3. **Level of academic challenge:** This scale focuses on measuring students' academic effort and the institutional expectations of students.
4. **Enriching educational experiences:** This benchmark is designed to collect information about the number of complementary academic activities available to students. This includes joining in community service, participating in curricular and co-curricular activities.
5. **Supportive campus environment:** This scale reports the degree to which students perceive their institution is dedicated to student success.

In 2004, Fredricks et al. (2004) further recognised three main dimensions of student engagement. These are detailed as follows:

1. **Behavioural engagement:** students who are behaviourally engaged would comply with behavioural expectations and norms, such as asking questions in class, attending course-related events etc.
2. **Cognitive engagement:** students who are cognitively engaged would invest in learning and development, strive to go beyond the basic requirements, and enjoy the academic challenge, such number of assignments emphasizing higher order thinking skills, time spent on coursework requiring reflection, integration and synthesis of concepts, etc.
3. **Emotional (affective) engagement:** students who are emotionally engaged would experience a range of emotional reactions such as attitude, value, interest, belonging, enjoyment, etc.

1.2 Student Engagement Measures

Student self-reports are the most common method for measuring student engagement. Most of these self-reports are general and not related to specific subjects, although there are some examples of measures that measure student engagement in a specific discipline like math (Kong et al. 2003) or reading (Wigfield et al. 2008). Some of these reports can be used to measure student engagement at institutional level, while the others are designed for course level. As the aim of this study is to measure student engagement in blended-MOOC courses, the researcher review only self-reports that can be used at course level within higher education institutions.

Measuring student engagement within a particular course experience provides valuable feedback for assessing and improving students' investment in the learning process. The feedback obtained from such an experience could be used as a way of evaluating a particular course's structure, pedagogy and design.

Barkley (2009, p. 44) asserts that, 'whatever means [a] teacher [uses] to [assess] engagement in their classes, gathering appropriate feedback can help close the gap between what teachers think is happening in their classes and what students are actually experiencing.'

1. The National Survey of Student Engagement (NSSE) is a well-used instrument to study undergraduate student engagement at an institutional level. It was also used by a number of scholars (Robinson & Hullinger 2008; Delialioğlu 2012; Neumann & Hood 2009; Owston et al. 2013) to study student engagement at a course-level. The model has ten engagement indicators which grouped into four major themes see table abaxo.

Table 1. NSSE engagement indicators

Theme	Indicators	Description
Academic challenge	1) Higher-order learning	This indicator reflects mental activities such as memorising, applying, analysing, evaluating and creating new ideas/ ways of viewing things.
	2) Reflective & integrative learning	This indicator evaluates the level of making connections among concepts and between new ideas and previous learning, and viewing issues from new viewpoints.
	3) Learning strategies	This indicator assesses the extent to which students use some effective learning strategies such as identifying key information in readings, reviewing notes after class, and summarizing course material.
	4) Quantitative reasoning	This indicator measures the ability to use and understand numerical and statistical information in everyday life.
Learning with peers	5) Collaborative learning	This indicator collects information on how often students engage in collaborative work and discussions with peers.
	6) Discussions with diverse others	This indicator examines the extent to which students interact with and learn from others with different backgrounds and life experiences.
Experiences with faculty	7) Student-faculty interaction	This indicator explores how often students work/discuss/interact with their faculty.
	8) Effective teaching practices	This indicator captures students' perceptions of how much teaching staff uses organized instruction, clear explanations, and effective feedback.
Campus environment	9) Quality of interactions	This indicator asks students to rate the quality of their interactions with other students, academic advisors, faculty, student services staff, and other administrative staff.
	10) Supportive environment	This Indicator explores students' perceptions of how much an institution emphasizes services and activities that support students' learning and development.

2. The Student Course Engagement Questionnaire (SCEQ) was developed by Handelsman et al. (2005) to measure four types of engagement:
 - 1) Skill engagement includes levels of practicing some learning strategies and skills that promote academic success, such as note-taking. This type corresponds to the NSSE category of level of academic challenge.
 - 2) Emotional engagement includes level of affective involvement with course materials.
 - 3) Participation/interaction engagement includes levels of participation in class and interactions with instructors and other students. This type corresponds to the NSSE category of active and collaborative learning.
 - 4) Performance engagement includes students' levels of performance in the class, such as grades.

3. The Classroom Survey of Student Engagement (CLASSE) is an adaption of NSSE developed by Ouimet & Smallwood (2005) to assess student perceptions of engagement in a course. CLASSE is composed of two instruments:
 - 1) The student version of CLASSE which asks students how frequently they engage in various educational activities within a particular course.
 - 2) The faculty version of CLASSE which asks faculty of that course how important the various educational activities are in facilitating student success.

The both instruments have six indicators: 1) Engagement Activities, 2) Cognitive Skills, 3) Other Educational Practices, 4) Class Atmosphere Impressions, 5) Optional Course-specific 6) Demographic Items.

The two versions of CLASSE can be contrasted and examined “to identify important and valued educational practices that are occurring less frequently than desired or expected.”¹

4. The Student Engagement (SE) is a 14-item assessment which are adapted from the NSSE instrument for use at the course level (Ahlfeldt et al. 2005). The questions examine three aspects of student engagement: 1) Collaborative Learning; 2) Cognitive Development; 3) Personal Skills Development.
5. The UK Engagement Survey (UKES) is an adaption of NSSE to understand UK students' experience in nine key areas:
 - 1) Critical Thinking; which uses the same items of the NSSE Higher order learning indicator.
 - 2) Learning with Others; which asks the same questions of the NSSE collaborative learning indicator.
 - 3) Interacting with Staff; has six questions, some of these questions were taken from the NSSE student-staff interaction indicator.
 - 4) Reflecting and Connecting; which has the same questions of the NSSE reflective & integrative learning indicator.
 - 5) Course Challenge; includes two items which focus on students' responsibility for their own learning, and their sense of being challenged by the course.
 - 6) Engagement with Research; asks four questions about the emphasis in the course on students learning about the methods and results of research.
 - 7) Staff-Student Partnership; has three questions about the extent students contribute to a joint community of staff and student, and work with teaching staff in improving their course and evaluating teaching and assessment practices.
 - 8) Skills Development (Hard and Soft Skills); involves Twelve questions ask about students' development of a number of skills and abilities.
 - 9) How Students Spend Time; asks seven questions about the number of hours that students spend on different activities.

The UKES can be used at two levels; an institutional level and a course level.

6. The Student Engagement Questionnaire (SEQ) was developed by Coates (2006) to study student engagement in blended course designs and the influence of online learning management systems (LMSs) on students' involvement with their study. The model consists of two parts:
 - a. General Engagement: General engagement includes nine scales to measure more general forms of campus-based student engagement:
 - 1) Constructive Teaching (CT) scale which is about students' perceptions of whether teachers supported students' active exploration of course materials and construction of knowledge.
 - 2) Supportive Learning Environment (SLE) scale addresses students' perceptions of whether they feel they are members of a supportive community of learners at university.
 - 3) Teacher Approachability (TA) scale measures students' perceptions of teachers' interest in students.
 - 4) Academic Challenge (AC), addresses students' perceptions of whether students feel that feedback and expectations challenge them to learn.
 - 5) Active Learning (AL) scale measures the extent to which students set standards and use practices likely to produce learning.
 - 6) Student and Staff Interaction (SSI) scale measures the level of interaction with teachers.

¹http://nsse.indiana.edu/html/classroom_survey_of_student_engagement.cfm

- 7) Complementary Activities (CA) scale measures students' involvement in activities around the university campus.
- 8) Collaborative Work (CW scale) addresses students' participation in a range of collaborative work/activities.
- 9) Beyond-Class Collaboration (BCC) scale measures the extent to which students are required to and do work together with others outside classroom.
- b. Online Engagement: Coates (2006) believes that online learning, which is a part of blended learning, requires its own engagement scales:
 - 1) Online Engagement (OE) scale focuses on the extent to which students incorporate the use of online learning systems into university study.
 - 2) Online Active Learning (OAL) scale is about the degree to which students actively use online learning systems to improve their learning.
 - 3) Online Academic Relevance (OAR) scale measures the extent to which students use online learning systems to contextualise and enhance the relevance of their university study.
 - 4) Online Teaching (OT) scale captures students' perceptions of whether teachers use online learning systems in pedagogically effective ways.
 - 5) Online Collaboration (OC) scale assesses students' use of online systems in collaborative work with peers.
 - 6) Online Social Interaction (OSI) scale is about how students use online learning systems to interact with others.
 - 7) Online Contact with Staff (OCS) scale measures the extent to which students contact with teachers online.

1.3 Student Engagement within MOOCs

A number of individuals have made different categorizations and classifications of MOOC learners' engagement patterns. For example:

1. Phil Hill (2013), in the e-literate blog, classified MOOC participants as follows
 - a) 'Lurkers' - who enrol in the course just to observe or sample a few learning materials at most. Some of them may not even get beyond the registration, so they are basically behind the scenes.
 - b) 'Passive participants' - who typically watch videos, perhaps take exams, but tend to not participate in the course's activities or discussions.
 - c) 'Drop-in participants' - who actively participate in specific topics within the course, but do not attempt to finish the course.
 - d) 'Active participants' - who fully engage in the course and actively participate in course discussions and social tools, and complete the course tasks and activities.

This classification has been also used by Milligan, Littlejohn, & Margaryan (2013) in their study to describe the engagement patterns in cMOOCs. Stanford University has classified MOOC learners into four classes: 'Auditing, Completing, Disengaging, and Sampling' (Kizilcec et al. 2013). However, these archetypes are dynamic, and participants may change over time from one to the other over the duration of a course (Joseph & Nath 2013).

2. Hew (2014) selected three top-rated MOOCs in the disciplines of programming languages, literature, and art and design to investigate the factors necessary to promote student engagement in the MOOC context. He found there were five factors behind the popularity of these MOOCs. These include: '(1) problem-centric learning with clear expositions; (2) instructor accessibility and passion; (3) active learning; (4) peer interaction; and (5) using helpful course resources'.
3. Wintrup et al. (2015) explored student engagement in two MOOCs developed by the University of Southampton and delivered through the *FuturLearn* platform. Wintrup et al. (2015) adapted the UK Engagement Survey (UKES), which is largely derived from NSSE items. The results of the study indicate that 'both MOOCs succeeded in enabling significant proportions of participants to feel engaged in intellectual endeavours such as forming new understandings, making connections with previous knowledge and experience, and exploring knowledge actively, creatively and critically'.

Most of blended-MOOC studies focused on measuring student experiences in blended-MOOC course design in terms of student scores (Pérez-Sanagustín et al. 2016; Griffiths et al. 2014; Najafi et al. 2014; Freihat & Zamil 2014) and student satisfaction (Griffiths et al. 2014). Very little studies focused on measuring or developing a model to measure student engagement in the blended-MOOC setting.

2. A BLENDED-MOOC STUDENT ENGAGEMENT (BMSE) MODEL

Upon reviewing a number of student engagement measures the most established and adapted instruments NSSE, UKES and SEQ were selected as a theoretical framework.

The study made use of four NSSE engagement indicators to measure student engagement in general course engagement: Higher-order Learning; Reflective and Integrative Learning; Learning Strategies and Collaborative Learning. These indicators are relevant to the blended-MOOC course design, while the other indicators are about how students perceive their institution, not an experience of a particular course. This study did not use the student-faculty interaction indicator from NSSE because the NSSE items in this indicator overlook the interaction between student and staff which take place inside the classroom. However, this kind of interaction was appreciated by UKES. For example, UKES asked *how often have you asked questions in taught sessions or contributed to discussions about course material in other ways*. Therefore, the study made use of UKES to measure student-faculty interaction.

In order to explain potential impact of using MOOC, three indicators of student engagement in the MOOC system have been added. These are adapted from the Student Engagement Questionnaire (SEQ). SEQ is a model with 19 scales designed to measure student engagement, both in online LMSs and on campus. It has ‘high construct stability, face and content validity, and replicates well in different institutions and fields of study’ (Coates 2007). The proposed model adapts only five that are relevant to the MOOC context. The full proposed model has nine indicators to track campus-based students’ engagement with MOOC and general aspects of their course. The indicators are presented in Figure 1 and Table 2 in detail.

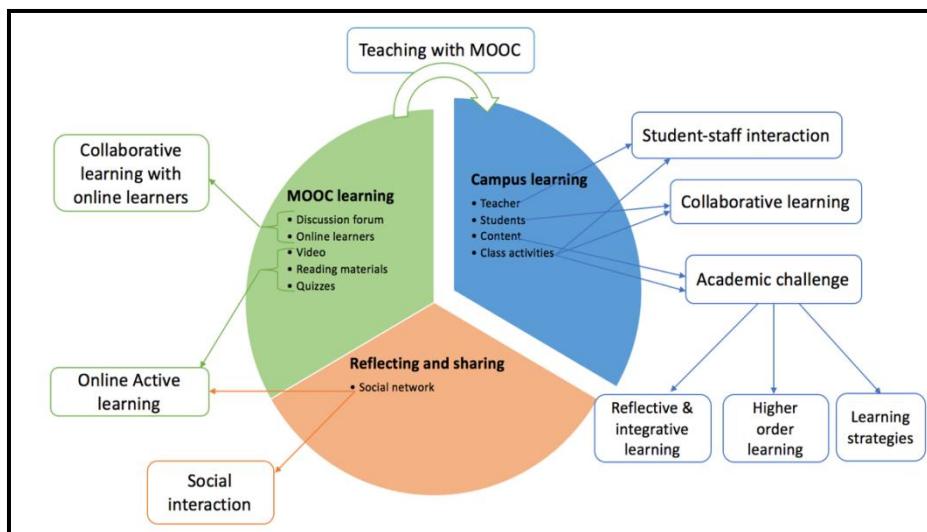


Figure 1. Student engagement in a blended-MOOC setting

Table 2. The Blended-MOOC Student Engagement (BMSE) model

Engagement indicators		Description
General engagement indicators	Academic Challenge	
G.1. Reflective and Integrative learning (source: NSSE)	G.1. Reflective and Integrative learning (source: NSSE)	Reflective and integrative learning: this indicator aims to examine the level of integration and connection of ideas and concepts.
	G.2. Higher order learning (source: NSSE)	Higher order learning: it captures the extent to which course materials and assessments emphasize challenging mental tasks such as application, analysis, judgment, and synthesis.
	G.3. Learning strategies (source: NSSE)	Learning strategies: it assesses the extent in which students performed some of effective learning strategies such as review/identify/summarize the course materials.

	<i>G.4. Collaborative learning (source: NSSE)</i>	Collaborative learning: this indicator examines the extent in which two or more students work together in academic activities.
	<i>G.5. Student-staff interaction (source: UKES)</i>	Student/staff interaction: this indicator explores the degree of students' contact and interaction with teachers.
MOOC engagement indicators	<i>M.1. MOOC Active Learning (adapted from: SEQ)</i>	MOOC Active learning: the items in this indicator collect information about the degree to which students actively use MOOC to enhance their learning.
	<i>M.2. MOOC Collaborative learning (adapted from: NSSE)</i>	MOOC Collaborative learning: this indicator captures the degree to which students use MOOC to do academic work with other online learners.
	<i>M.3. MOOC Social Interaction (adapted from: SEQ)</i>	MOOC Social Interaction: it aims to ask about how students use MOOC to experience a range of interactions and communications with others.
	<i>M.4. Teaching with MOOCs (adapted from: SEQ)</i>	Teaching with MOOCs: the questions of this indicator capture students' perceptions of whether instructors used MOOC in pedagogically effective ways.

3. RESEARCH METHODOLOGY

Two online questionnaires have been used to confirm the model.

1. The expert practitioner questionnaire was designed to collect experts' opinion on each indicator proposed from the literature. The questionnaire was divided into three parts. The first was to gather some general information about the participants. The second part asked nine 5-Likert scale questions, ranging from (Very important =5) to (Not at all important=1). The last section was two open-ended questions to let participants express themselves and provide information about whether there was uncertainty in any indicator, or whether any further indicators needed to be incorporated. The population of interest were instructors who had an experience in teaching in both settings: on-campus and MOOCs. Also, MOOC facilitators were interested, as facilitators who had experiences in teaching in campus setting and helping MOOC instructors by supporting MOOC learners through MOOC discussion and emails.
2. The student questionnaire was designed after confirming the proposed indicators by the experts. The questionnaire was distributed to and collected from students who had taken a course in a blended-MOOC format to evaluate and test the model.

3.1 Data Analysis Method

The expert questionnaire was analysed by using one-sample t-test to confirm the engagement indicators. If the mean value for a corresponding question of each indicator was significantly greater than 3.5. The reason of comparing with 3.5 is that this number is located between 'neutral' and 'important' on the five-point Likert scale.

Setting hypothesis

To test the null and alternative hypotheses, a confidence level was set with 95%, and the alpha α with 0.05. The null and alternative hypothesis were set as:

Null hypothesis: the mean value of the proposed indicator is equal to or lower than 3.5.

Alternative hypothesis: the mean value of the proposed indicator is higher than 3.5.

The proposed indicator is confirmed when the null hypothesis was rejected; i.e. p-value is less than 0.05.

4. RESULTS AND FINDINGS

The expert questionnaire survey was carried out over two months (from 10 January, 2015 till 15 March, 2015) with an involvement of 35 experts who are academics and educators/facilitators in the MOOC systems. As seen in Figure 2, the participants are from different backgrounds: 31.43% from Saudi Arabia; 28.57% of the respondents were from UK; 17.14% from Malaysia; 5.71% from Australia; and 2.86% from each of Algeria, France Mexico, Palestinian, Sweden and the United States.

As displayed in Figure 3, MOOC educators involved in this survey constituted 65.7% of the respondents, while 31.4% of respondents were facilitators and others constituted 2.9%. We only accepted academics who taught in the MOOC system. The 2.9% mentioned earlier were respondents who described themselves as both educators and facilitators. Some 31.4% of the sample had an experience of more than two years teaching in the MOOC system, and 31.4% had one to two years' experience and 37.1% had six months or less (see Figure 4).

In Table 3, it can be seen that there was agreement on the importance of all indicators. Interestingly, no participants registered an objection to the importance of any of the following student engagement indicators: reflective and Integrative Learning; higher order learning; student-staff interaction; and MOOC social interaction.

57.1% felt that reflective and integrative learning indicator is very important aspect of student engagement in the blended-MOOC context. However, 8.6% believed MOOC collaborative learning indicator is not important.

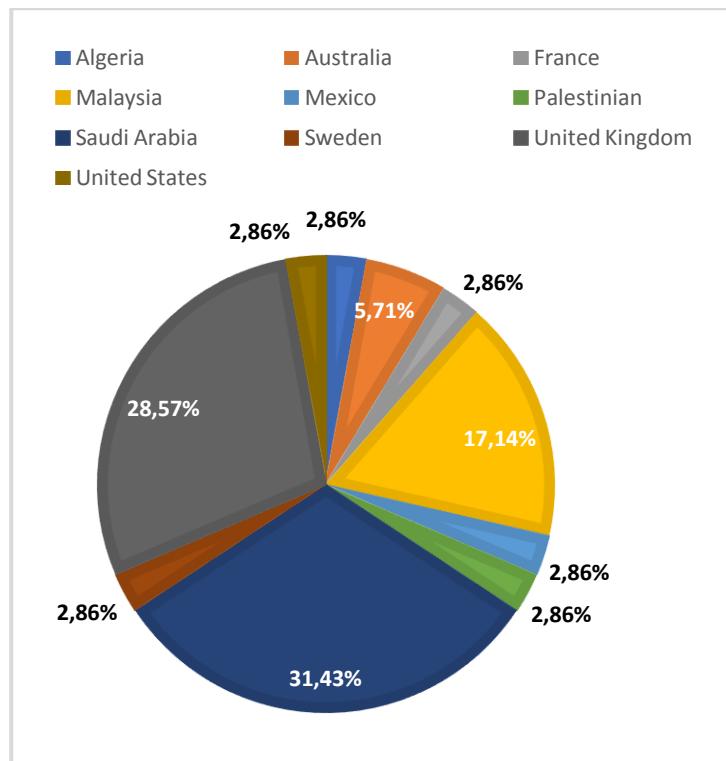


Figure 2. Participants by country

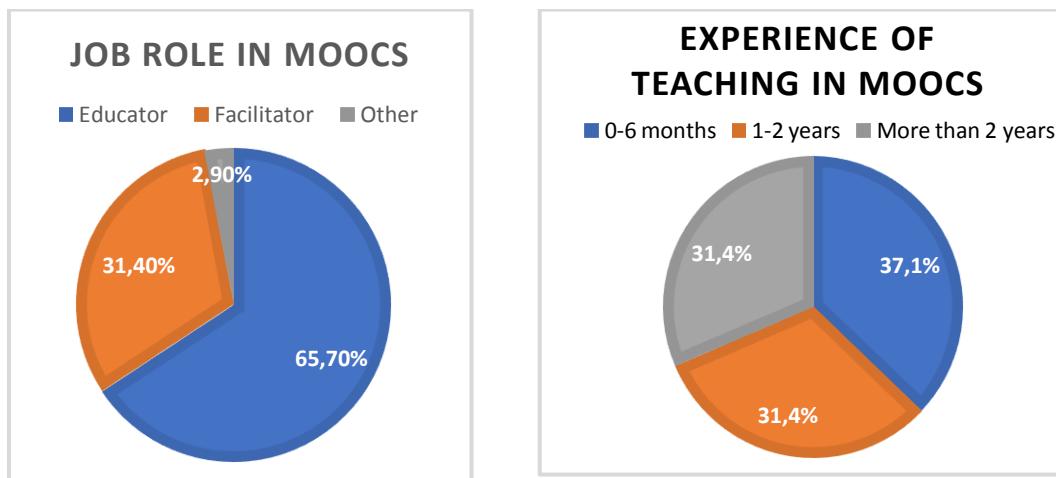


Figure 3. Participants by job role

Figure 4. Participants by experience

As shown in Table 3, the analysis of the results shows that all participants agreed on the importance of the proposed scales, as the mean value of each one was higher than the test value (3.5). The fact that all answers were significant, as p values for all indicators were less than 0.05, confirms that all proposed indicators were important to the blended MOOC framework.

Table 3. The results of expert questionnaire

Student engagement indicators	N	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Reflective and integrative learning	35	4.57	0.502	0.085	12.62	<.05
Higher order learning	35	4.4	0.604	0.102	8.81	<.05
Collaborative learning	35	4.26	0.852	0.144	5.52	<.05
Learning strategies	35	4.23	0.69	0.117	6.25	<.05
Student/staff interaction	35	4.37	0.646	0.109	7.89	<.05
MOOC active learning	35	4.2	0.759	0.128	5.45	<.05
MOOC collaborative learning	35	3.97	0.954	0.161	2.92	<.05
MOOC social interaction	35	4.2	0.719	0.122	5.75	<.05
Teaching with MOOCs	35	4.11	0.758	0.128	4.79	<.05

4.1 Pilot Study Results

The questionnaire was piloted with and collected from students who had taken either the digital marketing or the web science module in 2015/16 at the University of Southampton. The participants completed questions which were organized into two categories:

- Student engagement questions which were 37 questions classified into nine indicators: reflective and integrative learning; higher-order learning; learning strategies; collaborative learning; student-staff interaction; MOOC active learning, MOOC collaborative learning, MOOC social interaction; teaching with MOOC.
- One open question asking about their opinions about the questionnaire items/questions and design.

As seen in Table 4, the results comprise Web Science n = 7 and Digital Marketing n = 6, in total 13 participants. Another 15 attempted it, but did not complete the questionnaire.

Table 4. Student survey

Module	N	Time spent in the MOOC			
		none	about a quarter	about a half	all or nearly all
Digital marketing	6	0	2	2	2
iPhD Web Science	7	0	2	2	3
Total	13	0	4	4	5

4.2 Qualitative Data from Student Survey

Respondents agreed that the items were meaningful and clear. For example, one participant said: “the questions were clear to me. However, it would be more efficient maybe to keep the survey length a bit shorter”. Another student stated: “the questionnaire was easy to understand, and the course was really interesting”.

Small number of participants made comments on the course design. For example: One said: “The MOOC was not currently active when we were taking the module. We were added once it had finished and so there was little if any current discussions and support etc., in that sense this was one of the major disadvantages, also the MOOC was relied on too much with no real teaching only campus feedback sessions”.

Another participant commented: “I used the MOOC purely for educational materials, I was not interested in the social side of it, but that does not detract from my overall impression.”

4.3 Reliability and Validity of Student Questionnaire

The model was built on three widely used and well-validated models: NSSE, UKES and SEQ. The model has been reviewed by two experts prior to distribution in order to highlight any ambiguous, misleading or wrong questions.

The reliability of the student survey was assessed by considering Cronbach's Alpha. Reliability is generally considered to be satisfactory when values are greater than 0.70 (Hair et al. 2006), but this is sensitive to the number of items in the scale, so for short scales a lower value is acceptable, and conversely for longer scales.

Table 5 shows that the Cronbach's Alpha value of all apart from in two scales are equal to or more than 0.7. The two scales are slightly below the acceptable level, yet satisfactory for a short scale, thus the constructs were deemed to have adequate reliability.

Table 5. Cronbach's Alpha of each indicator

Indicator	Cronbach's Alpha	N of Items
Integrative & reflective learning	.864	6
Higher order learning	.700	5
Learning strategies	.513	3
Collaborative learning	.816	4
Student-staff interaction	.714	5
Active learning in MOOCs	.641	5
Collaborative learning in MOOCs	.700	3
MOOC social interaction	.859	2
Teaching with MOOCs	.823	3

5. CONCLUSION

Upon the review of student engagement literature, the well-established model NSSE was used along with other models UKES and SEQ to build up a model for measuring student engagement in the context of blended MOOC.

In this study, the indicator was accepted if the null hypothesis was rejected. Generally, a p-value is required to be .05 or less in order to reject the null hypothesis. The one-sample t-test (as seen in Table 3) shows that the p-value of all the proposed indicators is less than .05, so the result was found to be statistically significant and the proposed indicators may be used to measure student engagement in the blended-MOOC setting.

Moreover, the results of the expert questionnaire showed that the experts agreed on the importance of all the proposed indicators in the context of blended MOOC. With regard to the qualitative data, most comments were not indicator-related.

Moreover, the findings of the pilot study confirmed that:

- Questions in the model look clear and meaningful to most students, and
- The survey is deemed to have an adequate reliability in all indicators

REFERENCES

- Ahlfeldt, S., Mehta, S. & Sellnow, T., 2005. Measurement and analysis of student engagement in university classes where varying levels of PBL methods of instruction are in use. *Higher Education Research & Development*, 24(1), pp.5–20.
- Astin, A.W., 1984. Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, 25(4), pp.297–308.
- Barkley, E.F., 2009. *Student Engagement Techniques: A Handbook for College Faculty* 1st ed., New York, NY: John Wiley & Sons.
- Bruff, D. et al., 2013. Wrapping a MOOC: Student perceptions of an experiment in blended learning. *MERLOT Journal of Online Learning and Teaching*, 9(2), pp.187–199.
- Caulfield, M., Collier, A. & Halawa, S., 2013. Rethinking online community in MOOCs used for blended learning. *Educause Review Online*. Available at: <https://er.educause.edu/articles/2013/10/rethinking-online-community-in-moocs-used-for-blended-learning> [Accessed February 8, 2018].
- Chen, J., 2013. *Opportunities and challenges of MOOCs: perspectives from Asia*, Available at: <http://library.ifla.org/157/1/098-chen-en.pdf> [Accessed May 24, 2014].
- Chickering, A.W. & Gamson, Z.F., 1987. Seven principles for good practice in undergraduate education. *AAHE bulletin*, 39, pp.3–7.
- Coates, H., 2007. A model of online and general campus-based student engagement. *Assessment & Evaluation in Higher Education*, 32(2), pp.121–141.
- Coates, H., 2006. *Student engagement in campus-based and online education: University connections*, London; New York: Routledge.
- Delialioğlu, Ö., 2012. Student engagement in blended learning environments with lecture-based and problem-based instructional approaches. *Educational Technology and Society*, 15(3), pp.310–322.
- Fredricks, J.A., Blumenfeld, P.C. & Paris, A.H., 2004. School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), pp.59–109.
- Freihat, N. & Zamil, A.I.J. Al, 2014. The effect of integrating MOOC's on Saudi female students' listening achievement. *European Scientific Journal (ESJ)*, 10(34), pp.127–143.
- Griffiths, R. et al., 2014. *Interactive Online Learning on Campus: Testing MOOCs and Other Platforms in Hybrid Formats in the University System of Maryland*, Available at: http://www.sr.ithaka.org/wp-content/mig/reports/S-R_Interactive_Online_Learning_Campus_20140710.pdf [Accessed February 2, 2018].
- Hair, J. et al., 2006. *Multivariate data analysis* 6th ed., Harlow, Essex: Pearson.
- Handelsman, M.M. et al., 2005. A measure of college student course engagement. *The Journal of Educational Research*, 98(3), pp.184–192.
- Hew, K.F., 2014. Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCs. *British Journal of Educational Technology*, 47(2), pp.320–341. Available at: <http://www.ifres.ulg.ac.be/MOOC-blog/wp-content/uploads/2016/05/engagement-in-moocs-lu.pdf> [Accessed February 1, 2018].
- Hill, P., 2013. Emerging student patterns in MOOCs: A (revised) graphical view. *E-Literate*. Available at: <http://mfeldstein.com/emerging-student-patterns-in-moocs-a-revised-graphical-view/> [Accessed April 20, 2014].

- Holotescu, C. et al., 2014. Integrating moocs in blended courses. In *Proceedings of the International Scientific Conference of eLearning and Software for Education*. Bucharest, pp. 243–250.
- Israel, M.J., 2015. Effectiveness of integrating MOOCs in traditional classrooms for undergraduate students. *The International Review of Research in Open and Distributed Learning*, 16(5), pp.102–118.
- Johnson, L. et al., 2013. *NMC Horizon Report: 2013 Higher Education Edition*, Austin, Texas: The New Media Consortium. Available at: <https://www.nmc.org/system/files/pubs/1360189731/2013-horizon-report-HE.pdf> [Accessed June 11, 2014].
- Joseph, A. & Nath, B., 2013. Integration of Massive Open Online Education (MOOC) system with in-classroom interaction and assessment and accreditation: An extensive report from a pilot study. In *Proceedings of the international conference on e-learning, e-business, enterprise information systems, and e-Government (IEEE)*. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (worldComp), p. 105. Available at: <http://weblidi.info.unlp.edu.ar/WorldComp2013-Mirror/p2013/EEE3547.pdf> [Accessed May 1, 2014].
- Kizilcec, R.F., Piech, C. & Schneider, E., 2013. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge*. ACM, pp. 170–179. Available at: <http://rene.kizilcec.com/wp-content/uploads/2013/09/Kizilcec-Piech-Schneider-2013-Deconstructing-Disengagement-Analyzing-Learner-Subpopulations-in-Massive-Open-Online-Courses.pdf> [Accessed December 4, 2014].
- Kong, Q.-P., Wong, N.-Y. & Lam, C.-C., 2003. Student engagement in mathematics: Development of instrument and validation of construct. *Mathematics Education Research Journal*, 15(1), pp.4–21.
- Kuh, G.D., 2003. What we're learning about student engagement from NSSE: Benchmarks for effective educational practices. *Change: The Magazine of Higher Learning*, 35(2), pp.24–32. Available at: <http://www.tandfonline.com/doi/abs/10.1080/00091380309604090>.
- LaMartina, D., 2013. Blended MOOCs: The best of both worlds? *Campus Technology*, 21. Available at: <https://campustechology.com/articles/2013/08/21/blended-moocs-the-best-of-both-worlds.aspx?=&CT21>.
- Milligan, C., Littlejohn, A. & Margaryan, A., 2013. Patterns of engagement in connectivist MOOCs. *MERLOT Journal of Online Learning and Teaching*, 9(2), pp.149–159. Available at: <https://pdfs.semanticscholar.org/9b8f/37f6573c093b55588185dd1c486fa2465c1b.pdf> [Accessed February 2, 2018].
- Najafi, H., Evans, R. & Federico, C., 2014. MOOC integration into secondary school courses. *International Review of Research in Open and Distributed Learning*, 15(5), pp.306–322.
- National Survey of Student Engagement, 2007. *Experiences that matter: Enhancing student learning and success*, Bloomington. Available at: <https://files.eric.ed.gov/fulltext/ED512620.pdf> [Accessed August 2, 2016].
- Neumann, D.L. & Hood, M., 2009. The effects of using a wiki on student engagement and learning of report writing skills in a university statistics course. *Australasian Journal of Educational Technology*, 25(3), pp.382–398.
- Ouimet, J.A. & Smallwood, R.A., 2005. Assessment measures: CLASSE - The class-level survey of student engagement. *Assessment Update*, 17(6), pp.13–15.
- Owston, R., York, D. & Murtha, S., 2013. Student perceptions and achievement in a university blended learning strategic initiative. *The Internet and Higher Education*, 18, pp.38–46. Available at: <https://www.sciencedirect.com/science/article/pii/S1096751612000863> [Accessed February 13, 2018].
- Pace, C.R., 1984. *Measuring the quality of college student experiences: An account of the development and use of the college student experiences questionnaire*, Los Angeles, CA: Higher Education Research Ins. Available at: <https://files.eric.ed.gov/fulltext/ED255099.pdf> [Accessed February 1, 2018].
- Pascarella, E.T. & Terenzini, P.T., 2005. *How college affects students: A third decade of research*, San Francisco, CA: John Wiley & Sons.
- Pérez-Sanagustín, M. et al., 2016. Does taking a MOOC as a complement for remedial courses have an effect on my learning outcomes? A pilot study on calculus. In *European Conference on Technology Enhanced Learning, EC-TEL 2016, Lyon, France, September 13–16, 2016, Proceedings*. Springer, Cham, pp. 221–233.
- Robinson, C.C. & Hullinger, H., 2008. New benchmarks in higher education: Student engagement in online learning. *Journal of Education for Business*, 84(2), pp.101–109.
- UK Universities, 2013. *Massive open online courses (MOOCs) Higher education's digital moment?*, London, UK. Available at: <http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2013/massive-open-online-courses.pdf> [Accessed August 5, 2017].
- Wigfield, A. et al., 2008. Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, 45(5), pp.432–445.
- Wintrup, J., Wakefield, K. & Davis, H., 2015. *Engaged learning in MOOCs: a study using the UK Engagement Survey*, Available at: <https://www.heacademy.ac.uk/node/10346> [Accessed January 29, 2015].
- Zhang, Y., 2013. Benefiting from MOOC. In *A Paper presented in EdMedia: World Conference on Educational Media and Technology*. Association for the Advancement of Computing in Education (AACE), pp. 1372–1377.

TANGLE: A BLENDED EDUCATIVE SUITE FOR THE QUANTUM INFORMATION EDUCATION AND DEVELOPMENT OF RESEARCH SKILLS

Francisco Javier Delgado-Cepeda

Tecnológico de Monterrey, campus Estado de México

Carretera a Lago de Guadalupe km. 3.5, Atizapán, Estado de México, México. CP. 52926

ABSTRACT

Blended learning strategies are emerging as remarkable instruction resources not only for the formal curricula in higher education, they together are reporting value to disseminate certain specialized knowledge not included there but not less important for the professional development of students at the university. In such case quantum information and quantum computation areas are. In the last decade they have had an explosive development, thus involving a greater number of traditional disciplines, particularly those in engineering. Despite, these new areas still do not find an own place in the curricula for all possible involved programs. This work reports a blended learning initiative in those multidisciplinary fields as a learning suite containing courses oriented to the formal research, but departing from the corresponding scientific divulgation until research problems going on through the learning of the formal aspects needed in the middle. Suite is being constructed as multilevel approach to attract and to involve students since high school until graduate level. As an introductory research related with their usage in a blended context, some outcomes and their analysis are presented for an initial deployment during a two week nationwide workshop in the field.

KEYWORDS

Blended, MOOR, Engineering, Mathematics, Research

1. INTRODUCTION

In the knowledge society, some scientific developments for natural sciences (which during the 20th century were only advanced theories with a limited technological scope) in the 21st century are becoming a source of emerging technologies. Avant-garde scientific knowledge requires a mathematical and scientific contents often not included in the traditional engineering programs, only in those with a scientific profile. Such is the case for quantum information is mainly a physics discipline of emerging research which today is generating accelerated growth in applications and advanced technology. It involves multidisciplinary: physics, mathematics, computer science, automation, as well as electronic and materials engineering, all of them having important contributions there.

Currently, the lack of organized and supported research and education in Latin America for this area, still sets a gap between the attraction and the training of students into this area because learning resources have not been diversified and placed in simple terms. Books and notes are rarely written in Spanish, the native language of Latin American students. The aim of this work is to present an on-going educative initiative for quantum information education in the form of an educative suite of courses and resources taking advantage of the blended learning features. The development shows together the outcomes for an introductory deployment of the suite in its current stage, mainly presented in terms of achievements and opportunities.

2. TANGLE: THE SUITE IN THE BLENDED LEARNING CONTEXT

The areas of quantum information and computation are showing clear evidences of their importance: two Nobel prizes (2012 and 2016), the creation of first systems to establish quantum computation, the feasibility of establishing quantum information processing systems in terms of technology, and the growing world demand for human resources in science and engineering to develop them.

2.1 STEM Movement and Quantum Information Education

The above disciplines are professional activities do not existing yet properly, but expected to be highly valued in the immediate future. Thus, adequate educational resources are necessary to conduct students into them as a professional life possibility. The project in process is aligned with Science, Technology, Engineering and Mathematics (STEM) initiatives and the highlighted goals of the United States National Research Council, the National Science Foundation (NSF, 1999) and the European Commission (nIU, 2016; UE, 2016), as a fundamental area for the technologically advanced societies.

Quantum information processing group (QIPG) is a research group working in the last disciplines in Mexico sustaining collaborations with researches of United States, Italy, Canada and China. The group has a genuine interest in the human resources education in the discipline as a promising area of research and development involving master and doctoral students to engage them in international research projects. Also, the participation of some high school and undergraduate students in the group is oriented to learn and to develop basic associated research concluding in a scientific publication.

The current concern in QIPG is the lack of representation in Latin America of quantum information research, not displaying it as a high professional development area. Thus, it boosts the motivation on training resources hosted on an educative suite (called Tangle), particularly for Latin America (Wade, 2014) by deploying features and goodness of blended learning.

2.2 Tangle: A Suite based on Blended Learning and Scientific Collaboration

In the contemporary education trends, Massive Online Open Courses (MOOC's) (Black, 2014) and Mobile Online Open Researches (MOOR's) (Hosler, 2014) are approximations of blended learning (Gleeson, 2014) combining resources to develop concrete knowledge or research (Weisskopf, 2015): tutorials, adaptive materials, screencasts, computer codes repositories, learning activities, discussion blogs, online classes, research projects, etc. Although MOOC's sometimes are criticized by its student retention, MOOR's have an intrinsic motivation, either in his massive version (Hosler, 2014) or in the small groups version (Goral, 2014; Delgado, 2017) because it provides an emotional motivation in terms of future personal development.

In other educational approaches, QIPG has developed certain experience in the implementation of MOOC's, MOOR's, Small Private Online Research (SPOR) and flexible evaluation based on blended learning with an emphasis on research (Delgado, 2016a; 2016b). In addition, through the quantum information and quantum computation summer schools organized, students from Latin American countries (Colombia, Bolivia, Peru and Ecuador) have had punctual approaches with QIPG, making plausible the attraction to the own local programs. Thus, the contribution in progress depicted in this paper proposes a set of courses including materials, applications and collaboration spaces under MOOC, SPOR and MOOR modalities integrated in an educational suite with a multi-level training format (covering levels from High School, Undergraduate and Graduate).

For in site students (those enrolled in the institution), the suite could be used through several courses included in the curricula of each level (Table 1). For the external students (national or international students from the three levels), the suit becomes an open resource seeking to attract them into the official programs of the institution and particularly to be part of QIPG (some of them in open summer research stays). The following development of article first depicts how the blended learning approach contributes to the last goal. After, we show some technical results for an introductory research around the suite impact on an initial group of students using it to support the ongoing design and the value of blended learning approach on a quantitative basis.

3. TANGLE SUITE DESIGN AND INTRODUCTORY RESEARCH

The main goal of the educative proposal is the full construction for a series of courses and resources grouped in an educative suite for the areas of quantum computation and information, transversal for the three levels of Education, and crossing through the scientific divulgation, formal education until research collaboration. This series of courses and resources are differentiated and they are configurable to meet: a) the attraction of students at all levels, b) the acquisition of skills necessary for the quantum information area, and c) the

settlement of collaboration networks on research problems (a potential laboratory for learning and research). This section depicts the Tangle's design together followed by an initial deployment and its outcomes in the SPOR format by using the first Tangle's module.

Table 1. Associated courses to introduce and/or to include Tangle resources (partially)

Associated courses	Level	Semester	Program(s)
Experimental sciences IV	High school	6º	International Baccalaureate
Quantum Information Processing	Undergraduate	7º	Computer Sciences and Physical Engineering
Topics of modality in scientific research	Undergraduate	7º and 8º	Electronic, Mechanic and Mechatronics engineering
Introduction to Quantum Computation	Graduate	2º	Nanotechnology graduate programs
Introduction to Spintronic	Graduate	3º	Nanotechnology graduate programs

3.1 Tangle: Future Overall Design Expectation

Tangle's contents include a series of learning modules as MOOC's and MOOR's for the key areas of information and computation in Spanish. Modules are reconfigurable for several alternative depth levels and profiles: scientific, engineering and computational. Different methodologies in each module is according to the audience and the contents in agreement with Connectivism (Siemens, 2005): active learning, blended learning, and learning based on research. Figure 1 depicts the general structure and plan for the suite, currently only the first module is active.

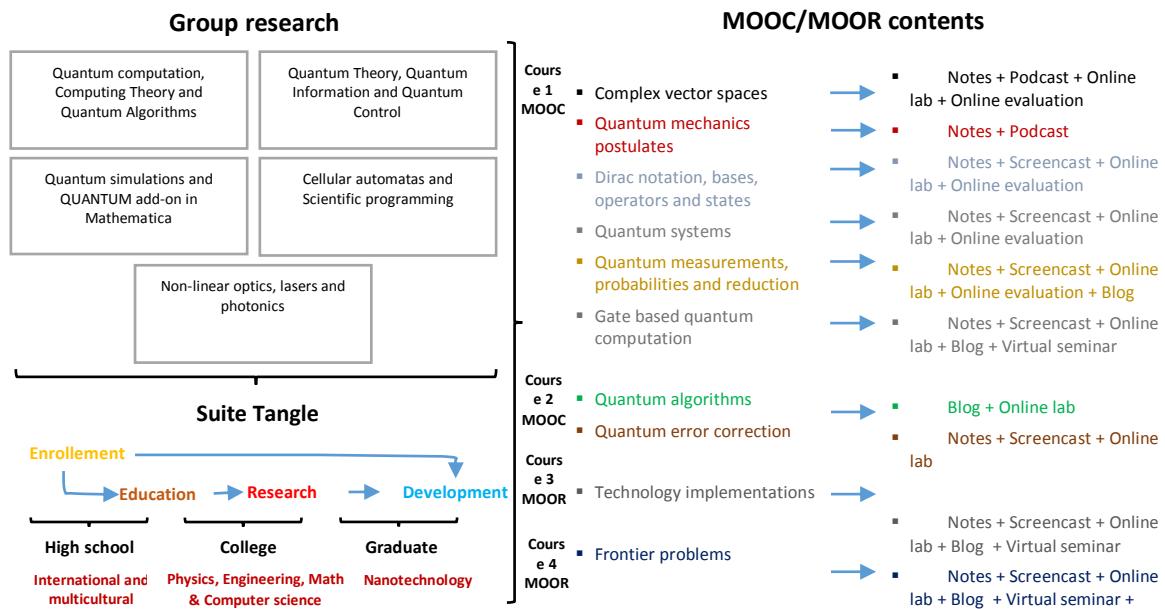


Figure 1. QIPG expertise, educative plan courses and their resources, as well as the multilevel scope

There are an incipient or null historical record in the academic impact of quantum information education as emerging area because it normally transcends the classical curricula. Therefore, future educative research objectives around of the suite includes: a) rising of user's demographic statistics (AAa), b) learning analytics to assess the level of contents coverage (AAb), performance (AAD) and captivation (AAC) of participants through the suite's courses and levels. The relations between course components as didactical elements and planned analytics are shown in the Figure 2.

3.2 Structure of the Blended Workshop

Initial design of Tangle (main part of the first module) was implemented with 25 undergraduate students of Computer Science, Physics and Engineering who rolled out in a first approximation in the employment of the Tangle suite for an immersive experience of two weeks in the form of a nationwide blended learning experience during the activity called “Semana I” (Delgado, 2018). It is a yearly mandatory activity for the students at Tecnológico de Monterrey attending a creative, social or research task organized in the institution. For a week, a face-to-face (mainly) challenge of intensive learning is conducted. During 2017 the QIPG proposed there (for interested students in science and engineering) a workshop in quantum information being supported by the Tangle’s resources during an additional previous week -online activities during the first week of immersion at distance-. Objective was to develop basic knowledge of this discipline to be able to propose tentative related future research culminating in a scientific publication of his own in subsequent spaces (research modality or graduate studies). This activity was an insight approximation for the real implementation of Tangle. The following section of the report establishes the main results related to this experience to assess the roadmap of Tangle development.

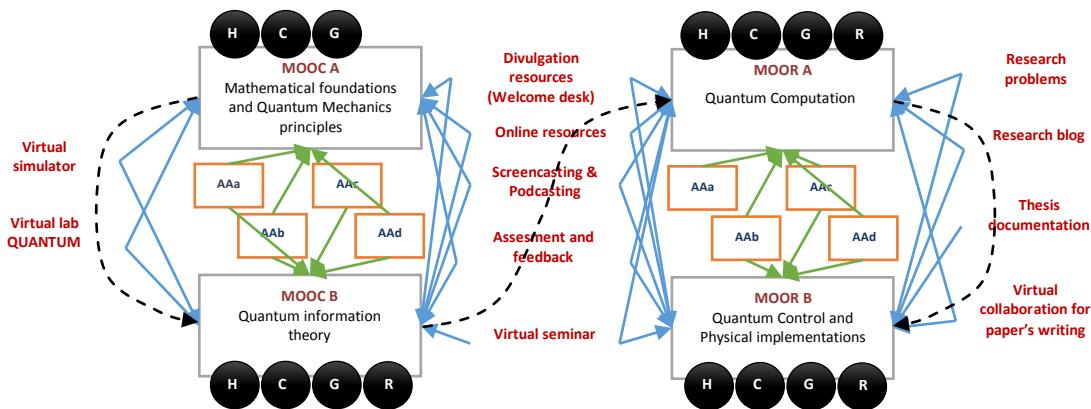


Figure 2. Learning elements for each course, target public (H: High School, C: Undergraduate, G: Graduate, R: Research) and associated immersed analytics (AAa, AAb, AAC, AAd)

3.3 Objectives for an Introductory Research

The workshop was conducted during two weeks. The first one at distance covering a series of materials introduced in the first module of Tangle ranging from divulgation until elementary lectures introducing the concepts and the mathematics involved. Each student should to cover those materials at his own pace during that week. Contents being effectively covered were tracked for each student, despite it was not an element of formal evaluation for the “Semana I” activity. This informal following will be called as the “Online grade” for each student, and it assumed as requisite for the face-to-face section of the workshop.

Workshop was developed (in the face-to-face week) in México City. The students registered for the workshop came from seven states of México, some of them from cities far away around one thousand kilometres. The students enrolled were ranging semesters from 2nd to 9th. There, students attend some key lectures before to the hands on with algorithms, programming and research development. Through the workshop several aspects were evaluated for each student through an observation guide and specific rubrics in the activities): participation (effective work in terms of intervention, support to their team) and leadership (directing some development in their teams). This component in the student’s evaluation is labelled in this work as the “Face-to-face grade”. Finally, as the goal was to construct a future research proposal, this component was measured for each team through the report and the innovation in the proposal (a discrete scale was defined, giving the 100% grade for new problems, 75% grade for complex problems but commonly studied in the literature, and 50% grade for problems do not reaching some of these standards), together with the individual performance in the public proposal presentation. Thus, a differentiated grade was normally

assigned to each student (this component will be called the “Quality final product grade”). Together, face-to-face grade and quality final product grade are combined into the “No-online grade” (the only grade officially reported for this activity).

Some concerns in this experience is to study as the online component of learning could conditioning the no-online one, particularly that related with the final product quality performance. This approach will become useful for the whole experience through Tangle. Thus, we set the following two objectives in this research work:

- a) To analyse how the online learning activities are related with the demonstrated performance in the face-to-face instruction
- b) To track which became the online learning coverage distribution on the student population and its relation with the no-online evaluation (the overall evaluation in the workshop, despite it could not be the grading proposal for Tangle where online evaluation it is expected to be considered)

4. OUTCOMES AND DISCUSSION

By tracking the effective individual coverage of materials in the online component of the workshop and then tracking the individual performance during the face-to-face workshop in the terms depicted before, the grades for each student were constructed. Also, the independent grade related with the research proposal and its presentation was tracked. Figure 4 shows the outcomes (with the 0-100 grade scale proportionally converted into a 0-1 grade scale for simplicity) a dispersion graph comparing the Online performance against the face-to-face performance for each student (each black dot). Radius of each black dot is proportional to the final product quality grade (in the range of 0.5 to 1). The linear regression depicted exhibits a relative good correlation with a Pearson coefficient of $R^2=0.74$. There are not conclusive visual evidence that the final quality product is compromised by the online materials coverage, mainly due to the social knowledge exchange in the teams during the workshop.

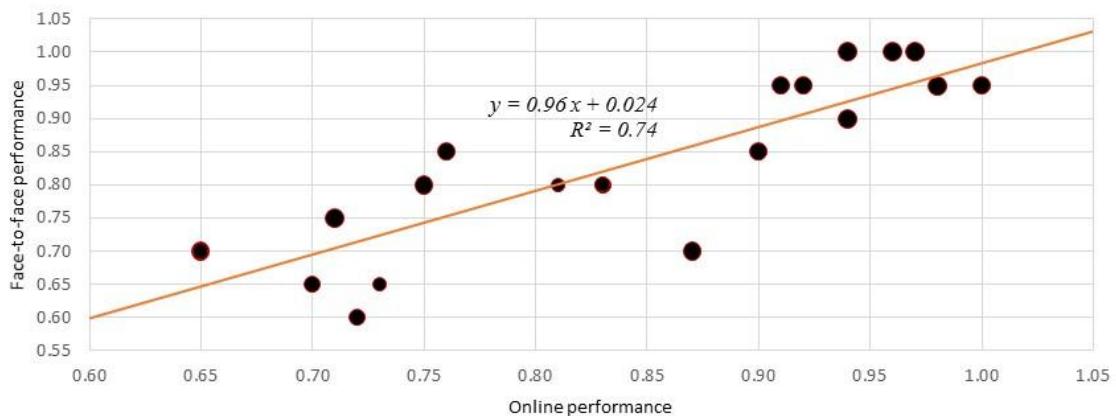


Figure 3. Online and Face-to-face performance vs. Final product quality

In order to have more conclusive outcomes between the differences established on the face-to-face performance and quality final product performance by the online component, we split the group in two subgroups: one of them having online performances larger than 0.7 and another with lower performances. Then, we run one-factor ANOVA tests with $p=0.05$ between both groups for each one of the first performances. This analysis discriminates each correlation by separating both no-online performance components by contrast with Delgado (2018). Results are reported in the Table 2. The first test gives statistical evidence for the impact of online performance on the face-to-face performance between both subgroups ($p=0.0002<0.05$). While, effectively the second test does not exhibit statistical difference between the quality final product performance between two groups ($p=0.3060>0.05$), as previously suggested the Figure 3.

In addition, the Figure 4 shows a statistical distribution (in percentage) for the online performance in the population. This distribution (in only six classes due to the reduced size of population) reveals two peaks in extreme behaviours for the coverage of material. While many students covered the online materials, other subgroup appears in the low range of coverage. In this graph, the average for the overall no-online performance in each distribution class is shown (upper line). This line reveals how a low range coverage seems to compromise the performance in the face-to-face and goal components of the workshop despite the expected team scaffolding.

Table 2. ANOVA tests to discriminate the difference between low and high online performance on other performances

ANOVA for Face-to-face performance based on Online performance difference						
Source	S.S.	d.f.	M.S.	F _{calc}	p-value	F _{critical}
Between	0.167	1	0.167	19.705	0.0002	4.381
Whitin	0.161	23	0.008			
Total	0.328	24				

ANOVA for Quality final product performance based on Online performance difference						
Source	S.S.	d.f.	M.S.	F _{calc}	p-value	F _{critical}
Between	0.010	1	0.010	1.106	0.3060	4.381
Whitin	0.176	23	0.009			
Total	0.186	24				

Through the previous outcomes, it becomes clear the use of online materials became useful to improve the knowledge and performance during the face-to-face section of workshop. Despite the orientation of workshop was not completely similar to the correspondent course planned as a MOOC, materials and activities there prepare to students since a divulgation level until an operative level integrating the physical concepts and mathematics involved in the basic quantum information theory. With them, students activated their previous knowledge in vector spaces, complex numbers and state diagrams arranging them as the basic structure to understand the basics of quantum information processing theory. Then, the face-to-face interaction in the workshop let to present other crucial aspects as quantum mechanics principles, quantum evolution, and other advanced topics, which despite they are included in the MOOC, we preferred presenting them during the second week interaction.

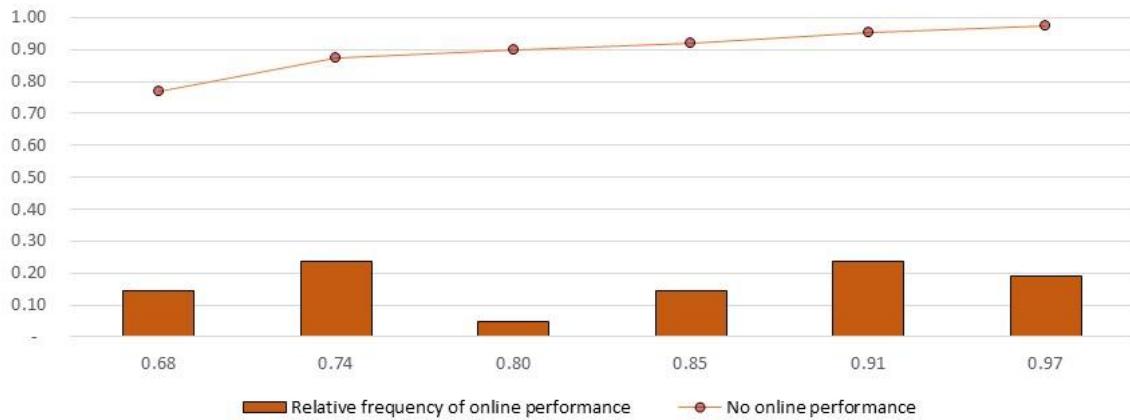


Figure 4. Relative frequency of online performance vs. No-online performance (Delgado, 2018)

A notable aspect revealed in the analysis of the previous outcomes concerns to the statistical independence between the performance demonstrated in the final presentation of the research project and the online performance, an issue possibly explained due to the learning achieved in the social interaction generated in the face-to-face section. This observation could support the belief about the valuable supremacy of the blended learning on the pure online one, because definitely the dominion exhibited in the proposal presentation for each student was reinforced in the teamwork developed.

5. CONCLUSIONS

Contemporary schemes of instruction and education should be running as a better approach in terms of effective development and flexibility. Blended schemes as the depicted here should be disposable for new generations. It becomes clear that new generations are learning in a different way, even if it is due to the extent of information through the new media or because they grew in a completely different world's conditions, or both. Soon, several courses should be evolved into MOOC, MOOR or SPOR schemes, changing the traditional higher education arena as it has been known. Transversal and multilevel Education (as here has been proposed) will be more necessary and common as the closer past (note that in the old history of Education it was a more frequent practice).

Skills and professional competencies require a continuous practice and several steps before to be owned, superseding any course extent. Traditional evaluation is limited to the contents' dominion, with uncertainty and not at the skilled level that each student need. The evaluation goes beyond the content knowledge, in the arena of skills and of competencies development, there is a really complex task for teachers (Schwartz, 2014) and researchers developing new knowledge and technology in terms to prepare new generations to stand in for the research work.

ACKNOWLEDGEMENT

Support to this research through NOVUS 2017 initiative is acknowledged to Tecnológico de Monterrey.

REFERENCES

- Black, D. 2014. MOOCs and the Move Toward Competency-Based Education. Retrieved from: <http://moocs.com/index.php/moocs-and-the-move-toward-competency-based-education/>
- Delgado, F. 2016. Follow-up on a Numerical Methods Course Using m-learning and Proposal for an Evolution to Research Based Learning. *Proceedings of The Eight International Conference on Mobile, Hybrid, and on-line Learning 2016*.
- Delgado, F. 2016. Upgrading to Numerical Methods Course into New Mobile Technologies for Mathematical Education: An Approach Based on Flexibility and Skill Development. In *Mobile and Blended Learning Innovations for Improved Learning Outcomes* (David Parsons, ed.). IGI Publishing: London.
- Delgado, F. 2017. Small Private Online Research: A proposal for a Numerical Methods course based on Technology use and Blended Learning. *Proceedings of Mobile Learning 2017*.
- Delgado, F. 2018. A learning challenge for a multidisciplinary quantum information and quantum processing workshop during "Semana I". *Proceedings of INTED 2018*.
- Gleeson, T. 2014. Co-Teaching a Blended Class across Universities. Inside Higher Ed. Retrieved from: <http://www.insidehighered.com/blogs/higher-ed-beta/co-teaching-blended-class-across-universities>
- Goral, T. 2014. SPOC's may provide what MOOC's can't. University business. Retrieved from: <http://www.universitybusiness.com/article/spocs-may-provide-what-moocs-can%E2%80%99t>
- Hosler, A. 2014. Massive Open Online Research: The MOOC Evolves into the MOOR. Retrieved from: <http://www.emergingedtech.com/2014/01/massive-open-online-research-the-mooc-evolves-into-the-moor/EmergingEdTech>
- nIU. 2016. Quantum Europe 2016: A New Era of Technology. European Commission. Retrieved from: http://gdriqfa.unice.fr/IMG/pdf/quantumeurope2016_report.pdf
- NSF. 1999. Quantum Information Science: An Emerging Field of Interdisciplinary Research and Education in Science and Engineering. Retrieved from: <https://www.nsf.gov/pubs/2000/nsf00101/nsf00101.htm>
- Siemens, G. 2005. Connectivism: A Learning Theory for the Digital Age. *International Journal of Instructional Technology and Distance Learning*, vol. 2, no. 1, pp. 3-10.
- UE. 2016. Quantum Manifesto: A new era of technology. European Union. Retrieved from: <http://horizon2020.mon.bg/en?h=downloadFile&fileId=86>
- Weisskopf, V. 2015. CERN: A Model for International Research. CERN.
- Wade, G. 2014. New model needed to close widening education gap. Retrieved from: <http://www.universityworldnews.com/article.php?story=20141121091229767>

ENHANCING READING SKILLS THROUGH ADAPTIVE E-LEARNING

Julie Villessèche^{1,2}, Christophe Quaireau^{2,3}, Olivier Le Bohec^{2,3}, Jérémie Nogues³, Sandrine Oriez¹, Anne-Laure Besnard¹, Fanny De La Haye^{3,4}, Yvonnick Noël^{2,3} and Karine Lavandier³

¹ Ace, Univ Rennes, F-35000 Rennes, France

² LP3C, Univ Rennes, F-3500 Rennes, France

³ Platform TACIT, Univ Rennes, F-3500 Rennes, France

⁴ Université de Bretagne Occidentale, Brest, France

ABSTRACT

E-learning is part of instructional design and has opened a whole world of new possibilities in terms of learning and teaching. Our project is an adaptive e-learning platform, developed to enhance skills from primary school to university learners. Two purposes converge here: a pedagogical one – offering new possibilities, especially in terms of teaching scenarios (blended learning); and a research one – confirming the effectiveness of an adaptive e-learning tool in the case of individualized cross-disciplinary competences such as comprehension of implicit information (French). The case study presented here concerns primary-school learners using the Implicit module over the 2016-2017 academic year. This article gives a first positive answer to the effectiveness of such a tool in this specific context.

KEYWORDS

Adaptive environment, E-learning tools, Blended learning, Reading comprehension, Primary-school reading, Web-based tutoring

1. INTRODUCTION

With the rapid development and diffusion of the Internet and its related technologies, physical transformations in the classrooms have occurred, alongside changes in the ways of teaching and learning. Being able to search for information on the Internet or being able to search for a word in a dictionary are both valid tasks today from a pedagogical point of view. E-learning is part of instructional design as Merrill et al. (1966: p. 2) defined it: “*a technology for the development of learning experiences and environments which promote the acquisition of specific knowledge and skill by students*”. And it became part of a pragmatic research question: can it be a supportive environment which helps in solving a teaching/learning problematic: individualization? Every teacher faces the problem of heterogeneity in the classroom: personalizing the learning process – i.e. taking into account individual competences – is time-consuming and difficult to realize when dealing with tight-scheduled programs.

There is a large amount of studies investigating individualized systems and proving their superiority in terms of learners’ performance compared to traditional teaching systems or “one-size-fits-all” approaches (Vandewaele et al., 2011: p. 119). Those individualized systems have different forms: student guides, tests, personal corrective feedback, tutoring, etc. But the most basic idea is that it is difficult to create personalized training: it is time-consuming, it is also more demanding in terms of staff, and thus, in terms of budget as well.

However, the reason researchers kept investigating this area is the effectiveness of individualized systems of training. Bloom (1984) identified what he called the “2 sigma problem”: “*The most striking of the findings is that under the best learning conditions we can devise (tutoring), the average student is 2 sigma above the average control student taught under conventional group methods of instruction*” (p. 4). From there, researchers investigated other parameters influencing learning conditions – parameters different from the teacher-learner relationship: influence and role of peers, tools (complementary computer learning courses, for example). The aim was to reduce the cost of tutoring while attaining the same results.

Slowly, e-learning imposed itself as a solution, and was found particularly fitting to put in place adaptive teaching individualized approaches. Bangert, Kulik & Kylik (1983) showed that “*computer-supported instruction at the secondary level*” was particularly effective as it kept “*youngsters interested and actively responding while guiding them easily from one level of difficulty to another. The result [appeared] to be better learning*” (p. 153). What this study does not highlight is the reason of the effectiveness of computer-supported learning. One of the advantages of tutoring is that it insures a one-to-one configuration (one tutor – one learner) and a consequence of that is the possibility to adapt the pace of the teaching process to the pace of the learner. It is the same for adaptive learning environments: the idea is to match with the pace and level of competence of the learner, giving them the possibility to follow courses at an accelerated or extended pace.

What is exactly adaptive e-learning? Paramythis and Loidl-Reisinger (2004, p. 182) consider a learning environment adaptive “*if it is capable of: monitoring the activities of its users; interpreting these on the basis of domain-specific models; inferring user requirements and preferences out of the interpreted activities, appropriately representing these in associated models; and, finally, acting upon the available knowledge on its users and the subject matter at hand, to dynamically facilitate the learning process*”. Hence, according to them, one of the first aspects of adaptive learning is *monitoring*. Adaptive means being able to perform this task and so, to estimate the competences of the learners and influence their learning strategies in order to enhance their performance. Putting in place an external monitoring environment is necessary as learners tend to overestimate their competences when monitoring themselves (Koriat & Bjork, 2005).

When dealing with adaptive environments, as they fit the individualization principles, which have mainly been proved to be effective in terms of learning, there is no real thought given to the learning content available to the learners. Here, our project, which was first launched in 2012, tackles the following question: To what extent is adaptive e-learning effective when considering precise – but cross-disciplinary – competences such as comprehension, vocabulary or grammar? This article is a first step to answer this question. Firstly, a state of the project is presented, while questioning the relationship of adaptive and online environments. Secondly, through a case study concerning comprehension, we try to elucidate the question of the effectiveness of an adaptive e-learning tool when targeting a cross-disciplinary competence. In closing, we present some further developments and possible directions that can be addressed while training learners to develop a particular competence through an adaptive e-learning environment.

1.1 Creating an Adaptive Online Environment

What is exactly e-learning? According to Plantec (2002), “*E-learning simply means using ICT for the various tasks associated to education or training: advertising, administrative registering, document production, synchronous or asynchronous communication between teacher and learner, assessment. This leads to renewed pedagogical approaches which are possibly independent in time and location and to tailor-made programs involving networks of institutions or companies*”. As this definition shows, e-learning is a vast domain of varied pedagogical practice and tools existing in both teaching and research fields. But, in recent years, one of the e-learning branches which has heightened a renewed interest in both pedagogical practices’ and research interests’ sides has been adaptive e-learning. As we underlined a few lines earlier, e-learning has drawn the interest of both teachers and researchers because of the possibilities it offers in terms of individualization.

At the beginning of the 21st century, a lot was written and done about online environments. However, not so much was done to create tools which would be entirely adaptable to individuals and which would, at the same time, allow teachers to put in place blended learning, especially in France. Indeed, the idea, regarding this project, is not to replace the teacher-learner relationship by a machine-learner interaction; we want to enhance the possibilities given to teachers to individualize their programs through adaptive e-learning.

In 2001, Coomey and Stephenson (2001) highlighted the fact that, though online activities were more and more used in the classroom, they had not found, at the time, “*any definitive evidence of the overall effectiveness of ‘e-learning’ compared with more conventional methods. This is not to say that this medium is ineffective but rather to say that there is little systematic and empirical work to show evidence of its evaluation*” (Mehanna, 2004, p. 280). This is what is to be tested here: the use of a medium in the classroom.

As what is at stake for e-learning is its effectiveness, this means a requirement for online tested projects, with a tailor-style ambition: the idea that it is possible to have multiple scenarios to fit the needs of multiple

individuals. “*Different students have different preferred ways to learn. Some may understand quickly through images, others may prefer texts and readings. Some may deal well with theories, others may learn through experiments and examples. By gaining insights into different learning styles, it offers means to design and provide interventions tailored to individual needs*” (Truong, 2016: p. 1185).

Though Truong’s introduction gives an insight into what an online user-model-based system could be, it confines the adaptive part to the ways to learn: images, texts, readings, etc. It does not tackle the real stakes of an adaptive system: individual differences cannot only be taken into account through learning strategies, and cannot only be based on subject preferences. It also has to be based on subject performance, which is the core of an adaptive system. The effectiveness of such a system has been established in several studies, such as Di Giacomo et al. (2016), Yang et al. (2014), etc.

Vandewaeter et al. (2011: p. 119) identifies three types of adaptive learning: (1) Macro-adaptive instruction; (2) Aptitude-treatment interaction; (3) Micro-adaptive instruction. Macro-adaptive instruction corresponds notably to mastery learning: the learner goes at their own pace, in the sense that they have additional instructions or additional support until they achieve mastery on the test. Aptitude-treatment interaction is a more comprehensive system of instruction in the sense that it tries to take into account individual aptitudes and characteristics: for example, one of its principles is that anxious students prefer highly structured systems of learning. Thus, “*Snow (1980) defines three levels of control, complete independence, partial control within a given task scenario, and fixed tasks with control of space. Several studies have shown that the success of different levels of learner control is strongly dependent on the students' aptitudes, e.g. it is better to limit the control for students with low-prior knowledge*” (Mödritscher et all, 2004).

The last type, micro-adaptive instruction, is the one we are interested in here, as it fits adaptive e-learning environment. Indeed, micro-adaptive instruction “*diagnoses learner's specific learning needs during instruction and subsequently provides appropriate instructional prescription for these needs*” (Vandewaetere, 2011: p. 119). This means that this type of model is dynamic and can fit into a classroom environment, enhancing the possibilities of teacher-learner interaction, as it creates a tutoring system, which can be handled by the teacher, thus putting in place a blended learning system (Kakosimos, 2015). On one hand, as the system provides adaptive content to the learners, the teachers can provide feedback individually or to groups of students, and on the other hand, the system creates the content environment which fits the abilities of each learner without any external intervention.

This typology helps to orientate the research about adaptive learning in two directions. On one side, there is the approach: what type of system of instruction? On the other side, there is the content: what type of exercises? At which rate? What is the main entry for the learner: An entirely adaptive system? A mixed system with a test to evaluate the learners’ level of competence and then adapted exercises with regular check tests to re-evaluate the learners’ level? In any case the goals remain the same: to be able to identify the needs of the learners and to provide them with the content adapted to their needs.

1.2 Need to measure Performance

It is on the basis of these thoughts that the platform TACIT (Testing Adaptatif des Compétences Individuelles Transversales) launched its first module for the comprehension of implicit information in 2012. It is an online platform accessible from any browser. There are two different parts in this platform: one planned to be used by speech therapists, while the other one is intended for the classroom (so far, primary and secondary school – one project is in progress for universities).

The platform is a type of micro-adaptive instruction system, as defined earlier (Vandewaetere, 2011). It allows the learning process to take three forms: evaluation, individual or group training (autonomy), and tutored training. As any micro-adaptive instruction environment, the first requisite on the platform for any learner is to take a first evaluation, which places them on a scale from A (easy) to J (difficult). Once this is done, the training they will be given will correspond to their level, which can be regularly re-evaluated. Individual or group training corresponds to a full-autonomy training, while tutored training is interrupted by the feedback given by the teacher after each exercise.

This platform was used to retrieve data about primary-school learners, who used one of the two modules which are currently available online: Implicit (the other module is Vocabulary). Both modules are in French. They match two different goals: a pedagogical one – helping learners with comprehension difficulties

without letting down those who are more advanced, and a research one – verifying the effectiveness of an adaptive e-learning tool, which was created to help learners enhance their competences in comprehension.

So far, the platform is used in France and abroad. Since September 2017, every primary or secondary school which has registered has had access to both modules. The case study which follows concerns a pool of pupils from the academic year 2016-2017, all working on the Implicit module.

Table 1. Number of users (schools) per year on the platform TACIT

Year	2013	2014	2015	2016	2017
Number of schools	154	422	647	629	992

2. ENHANCING COMPREHENSION THROUGH ADAPTIVE E-LEARNING: A CASE STUDY

Comprehension is a challenge for both pupils (learning) and teachers (teaching); comprehension is a fundamental and cross-disciplinary competence (Potocki, 2013): pupils need this competence in order to develop their capacities in the other domains and teachers need to be able to adapt to the individual differences – i.e. individual competence – of their pupils. One child out of five has difficulties in comprehension (Daussin, Keskpaik & Rocher, 2011; Cnesco, 2016). And those difficulties affect their adulthood: 21,6% of adults in France have performances equal or inferior to the most elementary level in reading (OECD, 2013). In this report, the cognitive skills of adults are directly related to a certain number of facts, such as their level of education, job, wage, etc.

Two different levels of comprehension can be found in a text and readers need to master both of them: understanding the explicit and the implicit information. So, to understand a text and thus, build a coherent mental representation of the situation in the text (Graesser & Clark, 1985; Kintsch & van Dijk, 1978; Perfetti, 1999; Trabasso, Secco, & van den Broek, 1984; van Dijk & Kintsch, 1983), a reader needs to be able to:

- Establish connections between the literal or explicit pieces of information
- Produce the missing information or inferences, i.e. deducing certain elements through details of the text or through ensuring consistency between a text and their general knowledge.

To be a proficient reader, you have to master both of them. But what differentiates a high level from a low level learner in comprehension is not the ability to answer the literal questions about a text, but the inferential questions. Indeed, the pupils who understand the less are the ones who produce few inferences (Cain & Oakhill, 1999). So, in order to help those with a low level in comprehension, teachers need to help them establish the connections, the coherence between the different sources of information available in the text, and enable them to express the implicit information in an explicit form.

2.1 Comprehension and E-learning

Comprehension is a complex task, which requires multiple competences. We have already quoted two of them: as a reminder, understanding the literal information, and producing inferences, – inferences which can be deduced from a text, but which, from a literal point of view, are missing. Of course, these are not the only competences at stake in a comprehension activity.

Ecale & Magnan (2015) show that there are also two types of processes which enable a learner to understand a text: those which concern the phonological and orthographic identification of the words, and those which imply a syntactic and semantic knowledge. Thus, the comprehension of a text cannot be limited to the understanding of the explicit and implicit parts; it has to be defined through the multiple competences needed to perform the task: decoding and identifying the words, vocabulary, morphology, syntax, semantics and the construction of inferences (Bianco et al., 2014). In this study, we have chosen to focus on the implicit task, i.e. the production of inferences, as it is the discriminative factor in terms of comprehension (see above).

The idea that a focus on the implicit task considerably improves the level of comprehension has also been developed in more detailed studies – some directly concerning French (Rémond, 2007; Emin, 2003). Those studies also show that implicit tasks are complex from a content point of view, but they are also complex from a learner's competence point of view. Rémond (2007: p. 19-20) shows that learners can be split into four groups regarding their results in implicit tasks:

- Level 1. 90% of pupils “know how to find explicit information and are able to make simple inferences”
- Level 2. 60% are also “able to make inferences and simple interpretations from information within different parts of the text”
- Level 3. 26% are even “able to make inferences about characteristics of characters and places from the text and know how to justify them, through personal knowledge and experiences. They understand simple metaphors”
- Level 4. 9% also “know how to interpret characters’ intentions, feelings, behaviours from the text, and are able to understand concepts such as find out the theme”.

This study gives us some insight into one of the difficulties linked to the question of inferences: there is an important diversity in the learners’ level and in the inferences. In her results, Rémond (2007) identifies several types clearly: place, time, feelings, etc. On top of those factual inferences, Quaireau et al. (2016) define eight other types of inferences.

2.1.1 Types of Inferences

“Inferential language requires children to use their language skills to infer or abstract information by inferencing or analysing, as occurs when a teacher asks a child to predict what a book might be about (e.g., ‘What do you think will happen in this story?’)” (Zucker et al., 2010: p. 66). This means that children have to work out the answer from hints and clues – more or less explicitly expressed – from the text: so, the text has to be read carefully, and personal knowledge and experiences may be needed in order to draw conclusions from those hints and clues.

Table 2. Inference classification used on the platform TACIT

<i>Types of inferences</i>	<i>Definition</i>	<i>Example</i>
<i>Grammatical – anaphora</i>	Those inferences allow the reader to link a word used as replacement or substitute (such as a pronoun) to its referent (Lefebvre et al., 2012, p. 11).	A car crashed into a tree, which was an old oak. Given its state, it will never run again. What does “it” represent? [The car]
<i>Grammatical – linguistic marks</i>	Those inferences allow the reader to identify information from signs of gender, number, plural and singular.	I saw Jenny and Tom last week. She is a doctor now. Who is a doctor? [Jenny]
<i>Syntactical</i>	Those inferences allow the reader to deduce information through the place of the punctuation.	“Panda. Large black-and-white bear-like mammal, native to China. Eats, shoots and leaves” (Truss, 2003) What does the panda eat? [The text does not enable us to answer]
<i>Hypothetical / conditional</i>	There are elements in the text allowing the reader to make a hypothesis, or which condition the answer to the question.	When I grow up, I’ll build a castle with four towers. How many towers does my castle have? [None]
<i>Numerical</i>	The answer can be deduced from calculation.	There are ten turns in this game. I’ve won four, and lost three. If I win another one, I’ll be sure to win the game. How many turns are needed to win? [five]
<i>Global semantics</i>	The answer can be deduced from the interaction between the meaning of the words and general knowledge or ability.	In Jane’s family, women have been teachers for three generations. What was the job of Jane’s grandmother? [teacher]
<i>Local semantics</i>	The answer can be deduced from the hints and clues which can be derived from the meanings of a few keywords.	After the race, Gillian got a gold medal. Why did Gillian get a medal? [Because she won the race]
<i>Spatiotemporal</i>	The inferences concern elements regarding spatial or temporal information.	The day after tomorrow is the beginning of the week. Which day are we? [Saturday]

As specified by Hall (2015), “*Researchers have proposed a number of inference taxonomies (...) and consensus as to a definitive taxonomy has not emerged*”. We sum up the inference classification used on the platform TACIT in table 2.

2.1.2 A Need for Tools

Most of the e-learning environments existing in French on the comprehension matter do not concentrate on the question of implicit information. But, as underlined before, the learners who understand the less when reading a text are also the ones who produce fewer inferences (Cain & Oakhill, 1999). However, if a tool is created, it has to meet two pedagogical imperatives: helping those who produce fewer inferences, and encouraging those who already produce more inferences to go further.

In order to take into account those individual differences, an adaptive environment needs to propose further help for those who need it. Producing inferences implies being able to link the clues and hints from the text: this means that the attention of the low-level learners can be directed towards those clues in order to help them produce the inference. The two solutions used on this platform come from the research on attention (Posner, 1990). There are two types of attention: (1) exogenous attention, which is related to salient events present in the environment (on the platform, we propose to use words highlighted in blue); (2) endogenous attention, which appears when a person decides to pay attention to a part of their environment, according to a pre-determined goal (on the platform, we propose preliminary questions to help answer the main question).

2.2 Methodology

Each of the projects which are going to be presented in this article started with the same procedure, which is the creation of exercises. For the module on implicit comprehension, we started with the creation of 1000 exercises. This number was considered high enough to cover all the possible forms of inferences. The first 200 were submitted to pupils as fill-in-the-blank exercises, with only one instruction: to produce an answer. The purpose was to collect wrong answers, which would be produced by pupils and which, thus, would be consistent with error strategies they usually have in mind. Those answers were then selected as deflectors for the exercises, whose final form was multiple choice questions. Each exercise consists of a text, – different lengths were selected, from one sentence to twenty lines – followed by an inferential question accompanied by four possible answers – the right answer, two deflectors and “I don’t know”. Once those strategies were established, the 800 exercises remaining were entirely written by the researchers engaged in the project. The 1000 exercises were then presented to 2300 pupils, aged from 7 to 14 – age-groups which correspond to 4 years of primary school and 4 years of secondary school (collège). The aim was to create a scale of difficulty, according to the IRM (Item Response Model), which enabled to create several evaluations, which assess learners’ competences in a similar way.

The IRM are statistical models which allow researchers to represent both the learners’ level of competence and the level of difficulty of the exercises on the same continuum. Though they are not very much used in applied psychometrics in France, they enable researchers to create adaptive pedagogical systems, where the exercises are selected according to the child’s level of competence. The model which was used here is the Rasch model: “*the examiner can use the target person’s performance on such a variable segment of self-chosen items to estimate his ability [...], the process is self-tailoring. As the target person takes the test he finds for himself the items in the test booklet of difficulty best for him*” (Wright, Douglas, 1975: p. 2). The model was first tested using RLRsim-package (Exact Restricted Likelihood Ratio Tests for Mixed and Additive Models) in the R environment.

Table 3. Distribution of pupils by school level and gender in the present study

	<i>Primary school levels</i>				
	CE1	CE2	CM1	CM2	Total
<i>Female</i>	432	1414	2209	2719	6774
<i>Male</i>	444	1469	2219	2934	7066
<i>Total</i>	876	2883	4428	5653	13840

Once the exercises had been created and their difficulty established, a study was put in place for pupils in primary school on the academic year 2016-2017 (2nd to 5th year). The children's repartition was as follows:

We tried to model the variance of results for the last evaluation the learners took during the academic year. We used a stepwise Bayesian linear regression. The method chosen was a stepwise in which the choice of predictive variables is carried out by an automatic procedure. More precisely, the stepwise selection is the forward stepwise selection. To determine the most probable model, we used the BIC (Bayes Information Criterion) and we chose a definite number of independent variables, which had a presumed link to the dependent variable (i.e. the last evaluation results). As we used a Bayesian linear regression, we started by standardizing the predictors before calculating the terms of the interactions.

We chose the starting level of learners, the number of training exercises, and the length of the trainings (calculated in days) in order to check if the observed progress is really linked to the use of the platform. As our analysis focused on a relatively substantial sample ($N=13840$), we also integrated complementary variables. Indeed, Cohen (1992) and Hair et al. (2006) showed that the number of observations directly determine the maximum number of variables accepted by a statistical model. The more observations there are, the easier it is to include new variables in the model. The other two potentially confounding variables we included and controlled were the gender and the school level.

2.3 Results

First, the regression shows that the most probable model, for our data, is the one which accepts the random effects (tested by simulation with RLRsim). The model also indicates that the last evaluation results are predicted by:

- The first evaluation results
- The length of the training period
- The school level
- The gender
- The number of training exercises the learners took
- An interaction between the first evaluation results and the length of the training period
- An interaction between the first evaluation results and the gender
- An interaction between the first evaluation results and the school level

This model explains 48% of the variance (Residual standard error: 0.69 on 13777 degrees of freedom; Multiple R-squared: 0.4858, Adjusted R-squared: 0.4854; F-statistic: 1085 on 12 and 13777 DF, p-value: < 2.2e-16). The key elements are the effect of the training period and the different interactions mentioned above.

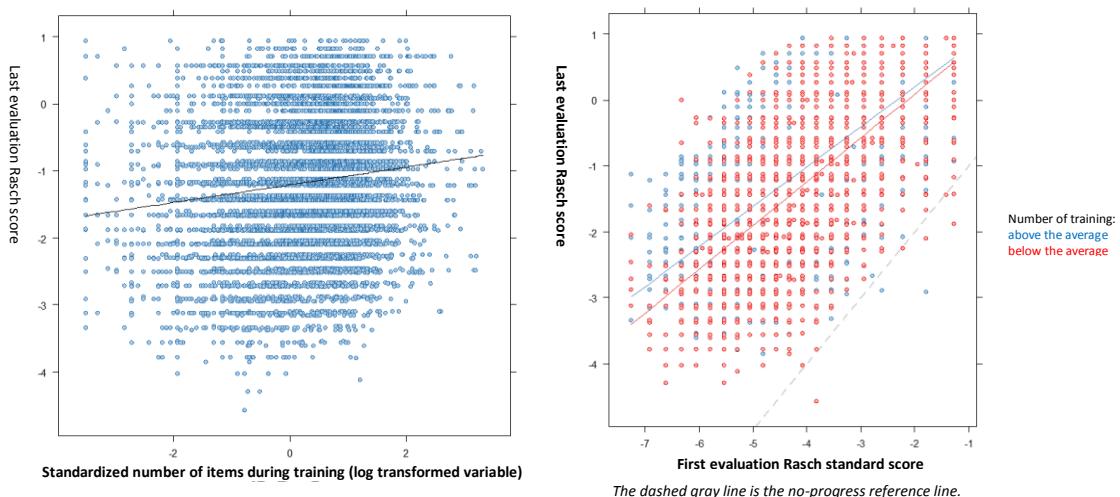


Figure 1. Effect of the training period length (left); Interaction between the first evaluation results and the training period length (right)

In Figure 1, we can see that the more exercises the learners do, the higher the last evaluation results were. This result shows that the observed progress is definitely linked to the use of the platform and confirms our assumption – the effectiveness of the platform. The second graph underlines the fact that the low level learners are the ones who make the most progress, and especially when they trained a lot during the academic year.

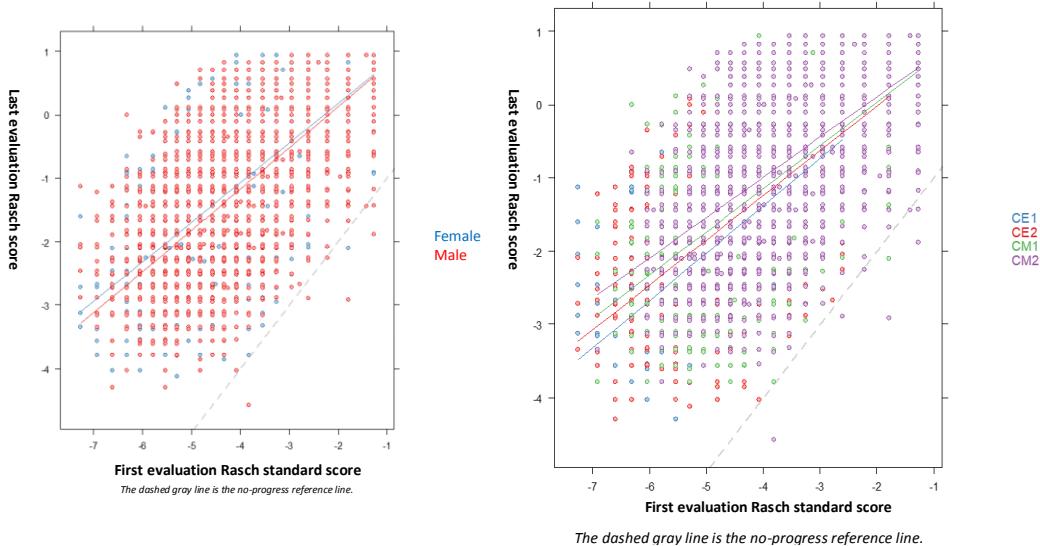


Figure 2. Interaction between the first evaluation results and the gender (left); Interaction between the first evaluation results and the school level (right)

The third graph (Figure 2) shows again that the low level learners are the ones who make the most progress, and this progress is more salient for low level female learners.

Again, the last figure shows that the learners who initially had the lowest level are the ones who make the most progress and the higher the school level, the more salient the results are.

3. CONCLUSION: E-LEARNING ENVIRONMENT AND EFFECTIVENESS

Adaptive e-learning is an effective solution regarding comprehension for primary-school learners. The feedback we have from teachers is that the platform blends easily into the lessons, enhancing the game aspect without losing sight of the objective: enhancing the learners' performances in reading comprehension.

What remains to be studied is:

- Is the platform efficient for all school levels? For the moment, only data from primary school learners have been retrieved and used.
- Is the platform efficient for every competence? Another module remains to be studied: Vocabulary.
- Is adaptive e-learning efficient for full-autonomy session? A module in English Grammar has been created and should be implemented by the end of 2018. The targeted learners are students in the universities, and the objective is to allow them to help them reach the B2 level (See Common European Framework of Reference for Languages) they are supposed to have when enrolling in university.

As Wijekumar et al. (2013: p. 366) pointed out, “*Even though scholarly journals are full of technology-based solutions that appear to produce large effect sizes in small studies, many do not reach their forecasted potential in large scale randomized controlled trials*”. So far, the platform has reached an equilibrium between researchers’, teachers’ and learners’ expectations. The implemented adaptive environment has been shown to be effective for enhancing reading comprehension performance for primary school learners.

REFERENCES

- Bangert, RL., Kulik, JA. and Kulik, C-L., 1983. Individualized Systems of Instruction in Secondary Schools. *In Review of Educational Research Summer*, Vol. 53, No. 2, pp. 143-158.
- Bianco, M., et al., 2014. Reading Comprehension development: Presentation of a special issue. *In L'Année Psychologique*, Vol. 114, No. 4, pp. 613-621.
- Bloom, B., 1984. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *In Educational Researcher*, Vol. 13, No. 6, pp. 4-16.
- Cain, K. and Oakhill, JV., 1999. Inference Making Ability and its Relation to Comprehension Failure in Young Children. *In Reading and Writing*, Vol. 11, No. 5-6, pp. 489-503.
- Cohen, J., 1992. A power Primer. *In Psychological Bulletin*, Vol. 112, No. 1, pp. 155-159.
- Coomey, M. and Stephenson, J., 2001. Online Learning: it is all about dialogue, involvement, support and control – according to the research. *In Stephenson, J. (ed.). Teaching and learning online: pedagogies for new technologies*. Kogan Page, London.
- Daussin, J-M., Kespaik, S. and Rocher, T., 2011. L'Evolution du Nombre d'Elèves en Difficulté face à l'Ecrit Depuis une Dizaine d'Années. *In France, Portrait social*, INSEE.
- Di Giacomo, D. et al., 2016. The Silent Reading Supported by Adaptive Learning Technology: Influence in the Children Outcomes. *In Computers in Human Behavior*, Vol. 55, pp. 1125-1130.
- Ecalle, J. and Magnan, A., 2015. *L'Apprentissage de la Lecture et ses Difficultés*. Dunod, Paris.
- Graesser, AC. and Clark, L., 1985. *Structures and Procedures of Implicit Knowledge*. Ablex, Norwood (NJ).
- Hair, J. F. Jr., et al., 2006. *Multivariate Data Analysis* (6th ed.). Pearson/Prentice Hall, Upper Saddle River, NJ.
- Hall, C. (2015). Inference Instruction for Struggling Readers: A Synthesis of Intervention Research. *In Educational Psychology Review*. Vol. 28, No. 1, pp. 1-22.
- Kakosimos, KE., 2015. Example of a Micro-Adaptive Instruction Methodology for the Improvement of Flipped-Classrooms and Adaptive-Learning Based on Advanced Blended-Learning Tools. *In Education for Chemical Engineers*, Vol. 12, pp. 1-11.
- Kintsch, W. and Van Dijk, TA., 1978. Toward a Model of Text Comprehension and Production. *In Psychological Review*, Vol. 85, No. 3, pp. 363-394.
- Koriat, A. and Bjork, R. A., 2005. Illusions of Competence in Monitoring One's Knowledge during Study. *In Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 31, No. 2, pp. 187-194.
- Lefebvre, P. et al., 2012. Analyse Conceptuelle de la Compréhension Inférentielle en Petite Enfance à partir d'une Recension des Modèles Théoriques. *In Revue des Sciences de l'éducation*, Vol. 38, No. 3, pp. 533-553.
- Mehanna, WNM., 2004. E-Pedagogy: The Pedagogies of Learning. *In ALT-J*, Vol. 12, No. 3, pp. 279-293.
- Merrill, MD. et al., 1966. Reclaiming Instructional Design. *In Educational Technology*, Vol. 36, pp. 5-7.
- Mödritscher, F. et al., 2004. The Past, the Present and the Future of Adaptive E-learning. *In Proceedings of International Conference of Interactive Computer Aided Learning (ICL)*, Villach, Austria.
- Cnesco (Conseil National d'évaluation du Système Scolaire), 2016. Note d'actualité. Available online: http://www.cnesco.fr/wp-content/uploads/2016/12/161206_Note_PISA.pdf [Last check: 02/05/2018].
- OECD (Organisation for Economic Co-operation and Development), 2013. *Education at a Glance 2013: OECD indicators*. OECD Publishing, <http://dx.doi.org/10.1787/eag-2013-en> [Last check: 02/05/2018].
- Paramythios, A. and Loidl-Reisinger S., 2004. Adaptive Learning Environments and e-Learning Standards. *In Electronic Journal on e-Learning*, Vol. 2, No. 1, pp. 181-194.
- Perfetti, CA., 1999. Cognitive Research and the Misconceptions of Reading Education. *In Oakhill, J. et al. (ed.). Reading development and the teaching of reading: A psychological perspective*. Blackwell Science, Oxford, England.
- Plantec, JY., 2002. e-Learning: Stakes and Tendencies. INSA Toulouse, France. 3rd International Workshop on Electromagnetic Compatibility of Integrated Circuits, Nov 2002, Toulouse, France. pp. 65-68.
- Potocki, A. et al., 2013. Effects of Computer-Assisted Comprehension Training in Less Skilled Comprehenders in Second Grade: a one-year follow-up study. *In Computers & Education*, Vol. 63, pp. 131-140.
- Quaireau et al., 2016. TACIT-Ortho : un logiciel en ligne pour aider à comprendre l'implicite des textes. *In Les Entretiens d'Orthophonie 2016, Les Entretiens de Bichat*, pp. 1-7.
- Rémond, M., 2007. La Lecture des Elèves Français au Collège : un état des lieux. *In La lecture au début du collège*. ONL-MEN, Paris.
- Trabasso, T., Secco, T. and Van den Broek, P., 1984. Causal Cohesion and Story Coherence. *In Learning and comprehension of text*. N.J. Lawrence Erlbaum Associates, Hillsdale.

- Truong, HM., 2016. Integrating Learning Styles and Adaptive e-Learning System: current developments, problems and opportunities. *In Computers in Human Behavior*, Vol. 5, pp. 1185-1193.
- Van Dijk, TA. and Kintsch, W., 1983. Strategies of Discourse Comprehension. Academic, New-York.
- Vandewaetere, M. et al., 2011. The Contribution of Learner Characteristics in the Development of Computer-Based Adaptive Learning Environments. *In Computers in Human Behavior*, Vol. 27, pp. 118-130.
- Wijekumar, K. et al., 2013. High-Fidelity Implementation of Web-Based Intelligent Tutoring System Improves Fourth and Fifth Grades Content Area Reading Comprehension. *In Journal of Educational Psychology*, Vol. 109, No. 6, pp. 741-760.
- Wright, BD. and Douglas, GA., 1975. Best Test Design and Self-Tailored Testing. *In Research Memorandum*, N0. 19, pp. 1-55.
- Yang, Y-T., et al., 2014. An Online Adaptive Learning Environment for Critical-Thinking-Infused English Literacy instruction. *In British Journal of Educational technology*, Vol. 45, No. 4, pp. 723-747.
- Zucker, TA. et al., 2010. Preschool Teachers' Literal and Inferential Questions and Children's Responses During Whole-Class Shared Reading. *In Early Childhood research Quaterly*, Vol. 25, pp. 65-83.

IMPLEMENTATION AND OUTCOMES OF A UNIVERSITY COURSE IN STATISTICS FOR BUSINESS UNDER THE FIT BLENDED MODEL

Faustino Yescas-Martínez and Francisco Javier Delgado-Cepeda

Tecnológico de Monterrey, campus Estado de México

Carretera a Lago de Guadalupe km. 3.5, Atizapán, Estado de México, Mexico. CP. 52926

ABSTRACT

FIT courses are an approach to distance learning oriented to flexibility, fast-track learning, and directed interactivity. This model has been recently developed in our institution to develop certain basic skills included in some key but non-core courses in several university undergraduate programs. FIT courses normally are taught as interactive broadcasting involving lecturing, problem-solving and continuous evaluation. Students have a ubiquitous presence several times in the week and they are enrolled in disperse far locations through the country. By depicting its methodology centred on tools and educative resources to face the challenge of understanding and applying the statistics' concepts and constructions, this work reports the outcomes of the teaching and learning experiences in terms of the daily log and the quantitative learning to be gained. We show the analysis for the course evaluation outcomes as a result of the continuous assessment in the learning tasks.

KEYWORDS

Blended, Distance learning, Statistics, Business education

1. INTRODUCTION

Probability and statistics are sciences having a wide application in areas related to science, technology, business and society. As a consequence, they appear in most programs of Higher Education as a mandatory course. However, the theoretical framework of these disciplines is not always accessible to understand them at the first approach and to come into the immediate application. Normally, time is necessary to apprehend this knowledge.

For Business programs, the course "Statistics for decision-making" has been included in the FIT (Flexible, Innovative and Technology) methodology. There, the communication technology is fundamental to teach the class sessions, to promote the student participation, and to deliver advice and evaluation. The approach emphasizes the use of concrete custom-made materials (lectures, homework, notes, video, activities, etc.) serving as a guide along the entire course and observing their application immediately after of each theoretical explanation. Course was supported and managed by a System Management Learning (SML) allowing creation and transmission of content, to track the student progress and the class involvement.

This paper depicts the experience and implementation outcomes for a university blended course in statistics for business. By reporting the learning performance outcomes obtained during the teaching experience under the blended FIT methodology, it is exposed the challenge for understanding and applying statistics. The course emphasizes the applied aspect of the Core Activity (CA), followed along the entire course and serving as a class guide in the course.

2. DEVELOPMENT

In the current section, we depict the course methodology as an evolution coming from distance and blended learning and also the objectives behind this research.

2.1 Theoretical Framework

Experience presented lies on the distance education concept. It has a long history in Education, emerging as alternative space for people who cannot attend a classroom. In the first correspondence course in 1728, Caleb Philips published an advertisement in the Boston Bulletin where asks about potential students for a short writing course. He declares you will receive weekly lessons in home to be instructed. Academic degrees online are available from 1858, being London University the first to offer them through distance learning. Despite, the first official institution of distance education was founded in 1921, University of Wisconsin-Extension. In the Internet era, in 1990, the first SML was provided by SoftArc, allowing to students interaction and advising for learning materials.

E-learning has been developed several years ago in Higher Education. The approach has gain terrain in the mind of such traditional educative structure due to the unavoidable advancement of technology, the increasing demand of qualified professionals in the terms dictated by industry, and the impossibility to deal with a growing population trying to access this educative level (Laurillard, 2005; 2008). Online resources have been integrated in parallel with the Internet development. It began to be a valuable space as complement for the traditional classes. Thus, those resources were the first competing with the exclusive face-to-face class to arise the pure online version of Education, a more similar version to the old distance learning based on lectures and reading materials (Mendoza-Gonzalez, 2016). With the advancement of the Internet speed, materials such as video and audio began to be included there, inclusively real time broadcasting, particularly since the mobile revolution around the period between 2011 to 2013 (Johnson et al, 2011). Those facts boosted the concept of Massive Online Open Courses (MOOC's) as a possible solution to the massive demand of Higher Education courses combined with flexibility (Hollands and Tirthali, 2014). The outcomes in this area have not been sufficiently satisfactory, particularly in terms of retention and effectiveness (Wood, Baker and Hooper, 2004; Garrison and Vaughan, 2008). In this arena, blended learning (in several modalities) has arisen as the most promising approach combining the flexibility of online Education together with the advising being offered in the face-to-face Education (Lothridge, Fox and Fynan, 2013).

Statistics education has been evolving from some theoretical approaches into more applied ones just as computer technology has also evolved. The use of several computer systems specialized in statistics (it includes from the basic functions in Excel until more specialized software as WinSPC, Statgraphics, SPSS, SAS and others) became more common in the last two decades to carried out the class in a computer lab. It made the class more dynamical providing an added value to the face-to-face class (Moore et al, 1995). Last practice also made statistics more valuable in terms of professional development because software become useful in science, industry and business to made decisions (Fonseca and Correia, 2009), together it has made statistics more affordable for students (Song and Slate, 2006). Through the time, statistics courses have been highly valuable as an applied knowledge necessary for the most of higher education programs, particularly those related with industry and business. It has generated the interest of blended approaches to teach statistics (Baharun and Porter, 2009). The effectivity of blended approaches to teach statistics has been reported already (Kakish, Pollacia and Heinz, 2012), setting a similar educative context for the approach reported in this work: diverse mathematical student levels of knowledge and a growing population of attendees.

2.2 The PATH Model and the FIT Methodology

FIT methodology emerges as initiative within the new educative model of Tecnológico de Monterrey (Tec 21) to support the Bachelor in Business Administration program (BBA) under the PATH model (an experience allowing students to explore different business related programs before the definitive program choosing). The new scheme aims to enhance the skills for present generations into the future challenges. In this model, there are three types of courses: mandatory, elective and specialized, distributed through the three stages of curriculum.

In the first stage, students explore and build their experience on a range of courses available in the first semesters. There, the statistics FIT course brings rational and critical thinking to students. In the second stage, each student decides how to approach to the professional path and finally the third stage provides specialization in the career of his choice.

Technology component for communication is primary for the course methodology. Through it, students actively participating in the sessions and teacher gives advice staying in touch throughout the course.

Technological platforms used are basically three: Zoom (to broadcast the class, advisory and interviews from mobile devices), Canvas (the mobile supported SML where course materials, tasks, evaluation and feedback are located), and Remind (a mobile app to be in touch, teacher and students). FIT methodology exhibits novel aspects in the course delivery. The outstanding one is how the content is structured on the digital SML platform (notes, tasks, presentations, videos, evaluations, etc.).

2.3 Design Elements for the Course

FIT methodology for statistics course is based on the strategy called applied learning, which is performed by a series of projects developed through didactic techniques as collaborative learning and case method. In addition, some engagement techniques as storytelling (the use of stories as part of an educational strategy) and roleplay (where each student assumes a role under a defined situation).

Course development is given through a core activity. It is the learning centre in the form of an applied problem to be solved during the semester. The core activity is divided in three stages. As support, each student has the course content, class twice per week at distance carried out within a strategy of engagement called Business Intelligent Meeting. In it, the teacher plays the role as a banking institution director and students as consultants in statistics of an employee trying to climb positions in the institution. This history and the roles played therein make it important that student should be properly trained in the course. At the same time, each student goes on support activities, task allocation and training quizzes.

As part of the didactic model, the student's evaluation is both formative and summative. It is formative because each student is allowed to take the quizzes several times, taking the highest grade obtained for the evaluation score. Then, it is summative because each one brings finally a component for the final overall evaluation. Evaluation tools include: rubrics, observation guides and quizzes. Activities to deliver are added to the Canvas platform, where an entire evaluation is displayed, on which each activity has a specifically score of the total course score. Feedback takes place within the same platform, where the teacher assesses, feeds and assigns the corresponding rating. Students have access to their entire grades permanently. The platform reminds the delivery of activities via e-mail.

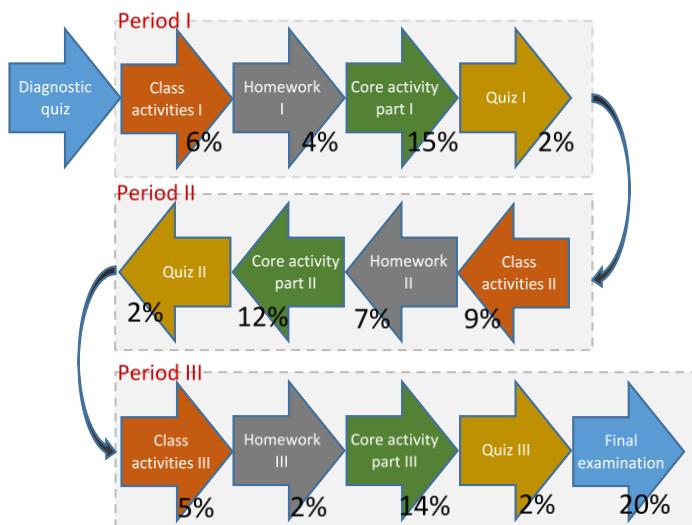


Figure 1. Activities through three periods of FIT course with their percentage value in the overall evaluation

Within the course, an individual portfolio is continuously constructed for each student. In it, the student places their learning evidence classified in competencies, disciplinary and cross-disciplinary skills. The evidences ranging from the solution of support activities, consultations, collaborative works, etc. Canvas platform allows the teacher to have an individual monitoring for each student in terms of performance and evaluation, helping him in the planning of the continuous feedback. At the end of the course a standardized online final exam is carried out. The test is designed and applied by a team of teachers at distance. The value of this evaluation is only 20% of the final overall course score. Figure 1 shows the course development in terms of activities and their percentage value in the overall evaluation. This scheme repeats three times in the semester through periods. Despite, CA is the same through the entire course as the final goal of learning.

2.4 Objectives and Research Scope

Compared with traditional courses in statistics in the same institution, they are face-to-face, they are also managed by a SML (Blackboard in this case), and they have three examinations through the semesters. Then, several possible research questions arises around the current course under a different methodology. The first one is about the final learning outcome in terms of effective traditional knowledge to be appropriated, particularly due to the practice to repeat the quizzes until to get the best outcomes. How this methodology is compared with the common course approach in terms of learning impact and skill development and how this practice is reflected on the concepts and skills finally being acquired? In a more basic and concrete aspect, how is the effective learning compared in this course with the traditional one in terms of attention, attendance, participation and teacher advice?

In agreement with the last concerns, the research objectives in the current report have been fix as:

- a) To track the individual performance for the overall content coverage
- b) To assess the average of skill development in terms of statistics dominion
- c) To relate the last aspects with the activities developed during the course
- d) To assess effectiveness of FIT methodology with the traditional model in terms of general performance

In the following section, based on the students grading and other observations through the course, we develop the corresponding analysis for the last aspects. Despite the initial deployment for this course was limited to a small group of 18 students, results are considered as an insight research about the potential goodness of FIT methodology in this technical course contributing to the entire PATH model.

3. OUTCOMES, ANALYSIS AND DISCUSSION

In order to analyse the outcomes in terms of the objectives previously stated, the whole grades obtained was tracked activity by activity, together with the outcomes for students in the same course during the last year but under the traditional model of instruction, particularly those in the final examination, a test with the same contents applied in the FIT course. In the following subsections we develop some analysis to get some insight into the goals of the research objectives.

3.1 Individual Performance compared between Activities and Final Evaluation

The first analysis is based on the individual grades of students comparing the overall set of grades for the semester activities. A representative figure has been calculated weighting each grade activity (with exception of the final examination grade) in agreement to Figure 1. Then it has been normalized in a 0-100 scale.

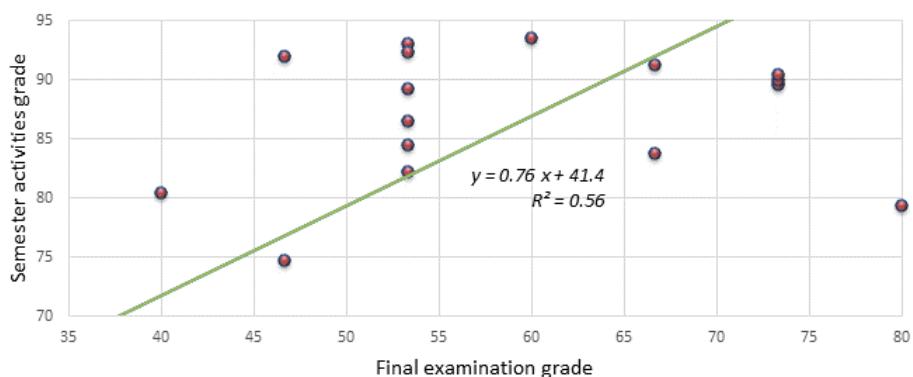


Figure 2. Dispersion plot showing the grades for the entire weighted activities of semester versus the final examination

Thus, the Figure 2 shows the dispersion outcomes for each one of the 18 students in the FIT model group, comparing semester activities grade with the final examination grade. The linear fit for the data is reported on the graph. Correlation between two grades is low. In addition, linear fit shows the gap between semester

activities performance demonstrated through activities in relation with the grade in the representative final examination involving the different topics in the course. This aspect is presented inclusively in the traditional statistics course without FIT methodology as will be analysed below.

3.2 Average Performance Compared across Activities and Main Themes

To assess the overall performance for the FIT group in the statistics course, we get the average and the dispersion of grades activity by activity across the course. Then, Figure 3a reports synthetically the outcomes showing schematically the standard deviation for each average. On the right (Figure 3b), another index has been reported. Calculating for each student and then getting the group average, the ratio between the proportion of wise choice in the multiple option final examination with respect to the correspondent grade (in 0-100 scale, by section involved) in the core activity across the course has been obtained.

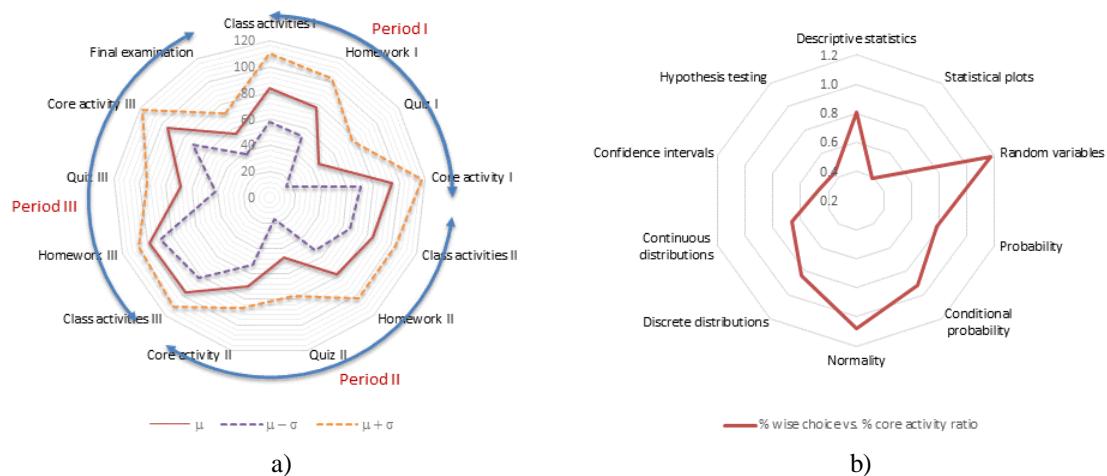


Figure 3. a) Activity course grades' averages, and b) ratio of wise choice and CA percentage by contents

Figure 3a reveals good performance in the homework and class activities, but it exhibits that the quizzes have normally a low evaluation despite the politics of flexibility under they are applied, a possible correlation with the final examination outcomes despite this last evaluation have only one attempt allowed. Notably, the dispersion for the quiz grades in general is wide, but this feature is not present for the final examination, possibly due to the one attempt politics. In this sense, quizzes flexibility does not work properly in the FIT approach. While, Figure 3b exhibits the weakest topics in the course, reflecting how a concept or statistical procedure could be well-developed in applications but it is not sufficient clear under a prompted request in an examination. There, Statistical plots and Hypothesis testing appear as the weaker knowledge after the FIT course.

3.3 Tests Comparing Activities Course Performance with Final Examination

Finally, two more statistical test have been carried out. The first one was an one-factor ANOVA test (5%) being get by dividing the entire FIT group of students in three subgroups as function of their semester activity grade (less than 80, between 80 and 90 inclusive, and more than 90). Test then analyses the comparison between averages of grades obtained in the final examination among three subgroups. The final finding does not show a meaningful difference among the final examination grade average for the subgroups ($p=0.41>0.05$). This test corroborates the finding in the Figure 2, there is an apparent unplugging between the skills developed and demonstrated during the course and the way as they are requested in the final examination.

Table 1. One-factor ANOVA test for the final examination outcomes based on groups of semester activities performance

Source	S.S.	F.D.	M.S.	F_{calc}	p-value	$F_{critical}$
Between	610.0	2	305.0	0.95	0.41	3.68
Within	4793.8	15	319.6			
Total	5403.7	17				

Supplementary Student t -test and F -test (5%) for the comparison of averages and variances in the final examination grades between the FIT group (18 students) and the historical group of students in traditional courses of statistics by the same teacher during the last year was performed (41 students). For the averages difference, by taking $H_0: \mu_{FIT}=\mu_{hist}$ versus $H_1: \mu_{FIT}<\mu_{hist}$, we get $t_{calc}=-0.07>-1.67=t_{critical}$. Then, there are not meaningful difference between both averages. For the variances, considering $H_0: \sigma_{FIT}=\sigma_{hist}$ versus $H_1: \sigma_{FIT}<\sigma_{hist}$, then $f_{calc}=0.77>0.45=f_{critical}$. Thus, no meaningful difference between variances is present. As consequence, any coarse differences in the two groups of students for the final examination grades is found. Table 2 reports the inputs for the last tests. Thus, outcomes for these tests show a comparable behaviour in the learning and evaluation in both schemes (FIT and traditional) despite the broadcasting is substituting the face-to-face class. Nevertheless, apparently the FIT methodology is not compromising the outcomes achieved by the traditional model, it is contributing some facility to cover this class at distance in places where a specialist in statistics is not always available.

Table 2. Input data for the test comparing final examination grades between FIT and traditional students

Group	n	μ	σ
FIT group	18	55.2	17.8
Historical group	41	55.41	20.1

4. CONCLUSIONS

The FIT methodology objective was defined to deliver and to expand specialized courses in places where some academics are absent in order to implement the PATH model. In addition, the design was conceived to deliver practical courses in key topics, it is the case of the statistics course in terms of rational thinking abilities. In the development of this research, it becomes evident FIT methodology attempts to give certain flexibility in the learning process, this aspect has not been adequately conducted in terms of student discipline and self-improvement.

Concretely, for the most traditional aspects in the course, as the continuous topical quizzes, together with the classical overall evaluation component (final examination), a clear improved dominion does not seem be pursued by the students. Results shown in Figure 2 reveals a gap between the application tasks with respect to the results in the concrete knowledge based final examination. It is barely appreciated in the dispersion outcomes in the quizzes (Figure 3). A finer discrimination of this issue is appreciated in Figure 4 for specific topics apparently denoting a lower performance in the final examination.

Despite it does not seem to exist learning differences between the FIT methodology and the traditional course for statistics, at least in the demonstrated dominion during the final examination (the only aspect with some possible degree of comparison between two learning models). Still, the skill development in statistics boosted by the FIT methodology should be studied more deeply as a possible outstanding aspect between two learning models. Other remarkable aspect is around of blended learning experience and distance learning. This experience becomes valuable because is closer to the training offered by some companies or industries. Thus, the experience in the development of abilities to be trained using similar techniques provides to students a possible professional value. FIT methodology at least matches the milestones of traditional learning, but in the future, some learning elements in the FIT methodology could be present in the professional training. Therefore, in order it becomes outstanding, it requires to be improved in terms of the student development, particularly in self-discipline and excellence.

In the last sense, the tangible elements to be improved in the FIT course of statistics are: a) to improve the statistical dominion outstandingly to the traditional learning course, b) to get a better demonstrable skill development in statistics for business in terms of using of technology or other applied issues as a some kind of certification, and c) to develop a superior grown in terms of learning discipline and self-regulation to reach the best learning outcomes. Last facts will become the most remarkable gain in this approach close to the future educative models.

ACKNOWLEDGEMENT

Economical support to present this work is acknowledged to the Physics and Mathematics department of Estado de México campus of Tecnológico de Monterrey. Support for the development of this research through NOVUS 2017 initiative is acknowledged to Tecnológico de Monterrey.

REFERENCES

- Baharun, N. and Porter, A. 2009. Teaching statistics using a blended approach: Integrating technology-based resources. Proceedings asilite Auckland 2009.
- Fonseca, M. and Correia, T. 2009. An increasing demand of basic biostatistics courses following the Bologna principles: a flexible blended learning approach. Proceedings of 22nd annual conference of the consortium of higher education researchers.
- Garrison, D., and Vaughan, N. 2008. Blended learning in higher education: Framework, principles, and guidelines. San Francisco, CA: Jossey-Bass.
- Hollands, F. and Tirthali, D. 2014. MOOCs: Expectations and Reality. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University.
- Kakish, K., Pollacia, L. and Heinz, A. 2012. Analysis of the Effectiveness of Traditional Versus Hybrid Student Performance for an Elementary Statistics Course. International Journal for the Scholarship of Teaching and Learning 6 (2), 25.
- Johnson, L., Smith, R., Willis, H., Levine, A., and Haywood, K. 2011. The 2011 Horizon Report. Austin: The New Media Consortium.
- Laurillard, D. 2005. E-Learning in Higher Education in P. Ashwin (Ed.) Changing Higher Education: The Development of Learning and Teaching. London, Routledge-Falmer.
- Laurillard, D. 2008. Digital technologies and their role in achieving our ambitions for education, Professorial lecture. London, Institute of Education, University of London.
- Lothridge, K., Fox, J., and Fynan, E. 2013. Blended learning: efficient, timely, and cost effective. Journal for Forensic Sciences, vol. 45, no. 4, pp. 407-416.
- Mendoza-Gonzalez, R. 2016. User-Centered Design Strategies for Massive Open Online Courses (MOOCs). IGI Global: USA.
- Moore, D., Cobb, G., Garfield, J. and Meeker, W. 1995. Statistics education fin de siecle. The American Statistician 49(3), 250-260.
- Song, H. and Slate, J. 2006. Taking statistics doesn't have to be scary: Keeping the heart rate down. Essays in Education, 17.
- Wood, R., Baker, J. and Hopper, D. 2004. Hybrid structures: Faculty use and perception of web based courseware as a supplement to face-to-face instruction. The Internet & Higher Education vol. 7, no. 4, pp. 281-297.

SOCIAL MEDIA PLATFORMS: CO-CREATING CODES OF CONDUCT

Jenny Eppard¹, Awil Hashi², Sandra Baroudi¹ and Marlieke Rodjen Helder¹

¹Zayed University, United Arab Emirates

²Ednastics

ABSTRACT

The way in which participants on social media sites interact is often dictated by rules that are either preset or develop progressive through participant collaboration. In order to explore this topic, interactions of an online community were observed for over one and a half years which spanned two related but physically different blogs. Bloggers were later interviewed to gauge their thoughts on codes of conduct and how they are created. The aim was to use purposeful sampling to the extent available in this community. Observations in the beginning were detached but at a later stage transformed to participant observations. Interview questions were semi-structured but in general they started with grand, general questions and became increasingly narrow. This study showed that allowing various stakeholders in the blogging community to co-create rules of engagement for both owners and participants provide for a richer experience with minimal outside interference.

KEYWORDS

Online Communities, Bloggers, Blogging Rules, Social Media Rules, Online Community Rules

1. INTRODUCTION

Social networking sites are evolving and becoming staple in the lives of many individuals. In these virtual spaces, individuals with a common interest meet and create a community where participants can freely express their opinion on any topic. Outwardly, this freedom appears to be liberating and outside of the constraints of normal society that are free from codes of conduct. However, as this paper will show, sites need to be monitored to some extent to ensure that well-intentioned members can use the blog without interference. Therefore, the purpose of this paper is to: 1) investigate the type of rules that guide interactions in online communities, as well as 2) creating guidelines for maintaining a sustainable blogging community.

2. LITERATURE REVIEW

2.1 Definition of Online Communities

There are numerous definitions of online/virtual communities. Probably the most noted definition is the one provided by Rheingold (1994):

...virtual communities are cultural aggregations that emerge when enough people bump into each other often enough in cyberspace. A virtual community is a group of people who may or may not meet one another face-to-face, and who exchange words and ideas through the mediation of computer bulletin boards and networks (p.57-58).

Preece (2001) defined online communities as “any virtual social space where people come together to get and give information or support, to learn, or to find company. The community can be local, national, international, small or large” (p.3). According to Blanchard (2008), virtual communities are groups of people who interact through the use of technologies and who developed a sense of belonging together. They feel they are sharing a common identity and influence each other. Drawing upon the foregoing definitions, online

communities can be generally defined as ‘virtual spaces where individuals with a common interest meet, using the Internet as a medium instead of gathering in person in an environment where there is some degree of familiarity’.

2.2 Rules Within Online Communities

Rules are the codes of conduct that different groups adhere to in online communities. They can either be formal (written) or informal (unwritten) (De Souza & Preece, 2004). Online organizations such as Facebook, YouTube and Twitter have rules that participants are obligated to follow when producing comments. These rules are usually posted under the title of standards, policies, terms of service or safety regulations and cover topics such as threat of harm, nudity, depiction of dangerous acts, cyberbullying, impersonation, child endangerment, violent imagery and spam. Benoit, Bilstein, Hogreve, and Sichtmann (2016) found that active members know the types of unwritten rules in an online community.

Arshad and Chuhalev (2007) developed a list of rules that are usually applied in online communities. Their work was based on previous research on the rules in online communities (Graham, 2003; Ostrom, 1990; Kollock, 1998; Godwin, 1994; Hsu and Lu, 2005; Johnson, 2001; Gallant, Boone and Heap, 2007; Wang & Fesenmaier, 2004). They presented the following as rules in online communities: putting a process in place that allows community to articulate norms and define rules; updating/revising rules in response to changing norms; identifying who decides in the end on the rules of articulation; taking feedback into account in formulation of rules; matching governing rules and local needs; participating in the modification of rules; putting a monitoring system in place; providing community members with low cost conflict resolution mechanisms; allowing members to settle their own disputes; providing institutional memory; providing incentives to users to remain online and discourage inactive users; providing incentives to promote users participation; forming special interest groups; including rules that are not a burden to the community; allowing for fine grained moderation and finally specifying a system of graduated sanctions.

Rules in online civic life were investigated in the seminal work by Fisher, Margolis and Resnick (1996). Their main findings regarding these rules and their functioning were 1) online communities lack a formal method of applying consequences to violations, 2) the rules were difficult to locate and understand, 3) the members preferred for the posts to be germane to the topic and 4) participants remain true to their intentions and not manipulate their reasons for interacting in the community. Rules in online communities are highly important as they constitute the policies of the community. These rules present a framework for users that defines the nature of their interaction and online activities (De Souza & Preece, 2004). Social media environments that share similar characteristics have the potential to be highly interactive and collaborative providing a basis for informal user-driven learning.

As stated before, it is important to investigate the rules that guide interaction in the community and create guidelines that could maintain a sustainable blogging community. Hence, the following three questions guided this study:

1. Which types of rules guide interaction in the community?
2. What happens when these rules are not adhered to?
3. How could an online community be sustained through informal or formal codes of conduct formation?

3. METHODOLOGY

3.1 Setting

Social media platforms represented by blogs is the focus of this study. The setting included an initial blog, an interim blog and an alternative blog. The initial and interim blogs were newsfeeds under the umbrella of a popular news organization. The blogs included news updated from professional journalists and comments from citizen bloggers. Within this system the community moved from an initial blog to the interim blog with a related topic. Both of these blogs were created and administered by the same organization. Finally, the group moved to an alternative blog that was created, monitored and regulated by the community themselves instead of by professional journalists or an umbrella organization.



Figure 1. Migration pattern of the blogging community

3.2 Participants

The demographics of the population under consideration contained bloggers who partook in a news-related blogging system that covered three related blogs with many of the same members. The bloggers participated by commenting on the blog or the comments made by other members. The sample size for the interviews was ten informants. The interviewees represented various members of the blog each with their own individual role. They were contacted by the researcher through their external, personal accounts. Participants were from different cultural backgrounds (mostly from North America and Europe). Before giving their consent to participate, they were assured that the data would be processed anonymously.

3.3 Data Collection

In the current study, a case study was designed using qualitative research techniques. Employing data triangulation to enhance the validity, data was collected through document analysis (web page analysis), detached observations, participant observations and interviews. Observations, both detached and participant, of the participants' posts and online exchange took place at randomized timings. During the interviews, in-depth questions were used and the question topics were focused on the goals of this project. Much of this data was derived from subject-to-subject interaction and through researcher-to-informant dialogue. Questions went from main questions to follow-up questions (more specific) to probes (most specific). From these questions, more specific questions were addressed (mini-tour questions). The interviews focused on motives for joining the blog and what their views were on learning through participation. The interviews took place at the end of the data collection period and concurrent with it.

A research diary was kept throughout the study documenting all of these occurrences. It was included as part of the data collection and analysis as the diary is an integral part of qualitative research since the researcher - like a quantitative test or survey - is considered an instrument (Spradley, 1979).

3.4 Data Analysis

Taxonomic analysis (Spradley, 1979) was used as a method to analyze the collected data. The data was first organized based on categories using focused coding (Emerson et al., 2005). Thereafter, the categories were clustered into classes. These include such topics as rules created by citizen journalists, rules created by the umbrella organization, keywords, infractions (such as identify theft, spamming), notification and consequences.

4. RESULTS

4.1 Formal Rules

Rules on the initial and interim blogs came explicitly from the staff at the umbrella news organization and implicitly from the blogging community. The news organization had specific rules that bloggers should follow although they were rarely enforced. Bloggers were theoretically not allowed to repeat posts as one blogger explained; "I don't think it is wise 2 use the name & the ava of the activists. Could backfire on u. Just sayin..". This post was repeated more than three times with different user names. Another explicit rule on the initial blog prohibited cursing, which was circumvented by using symbols. Additionally, bloggers could not

upload videos, images, and links and were limited to texts or workarounds. Many of them complained about this restriction as other websites do offer this possibility.

Officially, bloggers could use flags to identify other members who broke the rules. These flags were informing the blog administrators and other bloggers about the violation of the terms and conditions. Despite that these infractions were noted by the bloggers, they were not consistently addressed by the moderators or administrators of the blog. Many of the bloggers seemed to want the initial and interim blog's staff to actively moderate the site but due to the lack of official moderation, they moved to their own alternative blogging site in which they feel to have more control over the rules.

Upon entering the alternative blog, the rules were listed in the center of the page. There was a general belief among the informants that the moderators of the alternative blog (volunteer positions) were more capable to manage bloggers who disrupted the blog than the moderators managing the initial and interim blogs. This is due to many reasons of which the administrators of the alternative blog permitted links and uploads and did not limit the blog content. Additionally, the bloggers who used the alternative blog felt more relaxed because they were able to determine who can develop the rules and were encouraged to provide feedback. Hence, the blog was moderated at all times as participants volunteered to take shifts in monitoring the comments.

Examples of the rules were as follows: keep to the topic, no cursing, no spamming, no personal attacks, no jokes about religion, no racist attacks, respect the moderators (alternative blog, December 2014). There was a bit of humor embedded in the list but the gist of the rules was to drive bloggers to be respectful and courteous to each other. In addition, the rules were updated to suit the needs of the community at any given time.

4.2 Unofficial Rules

The bloggers had their own implicit set of rules, even on the initial and interim blogs, that evolved and amended over time by the participants themselves. For example, the bloggers were to refrain from intense virtual arguments. Debating was acceptable but it should not turn into a full-scale argument:

Claire: The main implicit "umbrella" rule is [sic]not to let one-to-one debate getting out of control. This can be clearly seen when a debate is too long, gets re- petitive, gets too harsh, or when it attacks the others personally [sic] (interview, June 12, 2012).

Adam: One shouldn't get into personal feuds (interview, June 11, 2012).

Felix: Do not insult others. cool down the emotions in many cases (interview, June 6, 2012).

Since that there were not any meaningful mechanisms in place for solving disputes on either site, most of the bloggers tended to have a one-sided view on the initial blog, the interim blog and the alternative blog. Authors who were opposing the general opinion were often called trolls and ridiculed and therefore did not post as often as the other group.

The community used to address conflicts and misunderstandings by encouraging credible materials and sources instead of having comments that are based on individuals' emotions or rumors. Adam explained "... it's understood that one should cite sources, and provide links when possible" (interview, June 11, 2012). Hence, if someone posted an unreliable comment, other members would complain and ask for a reference to ensure the credibility of all comments.

An additional rule imposed by the bloggers was to stay on-topic. Adam stated that: "whenever there are discussions about particular topics, bloggers should stay on-topic and avoid mentioning irrelevant news from other parts of the world. My sense has always been to label such news as "off topic" at the top" (interview, June 11, 2012).

Online and physical security of the bloggers was a concern. There were several incidences of hacking, threats and even one physical attack. Some of the bloggers imposed a rule on excluding any personal information from the comments. Therefore, the moderators on the alternative blog deleted email addresses existing on the commentary space of the public blog.

As a consequence, informal rules that bloggers encouraged others to follow on the initial and interim blogs turned into formal rules on the alternative blog. Cal wrote the following about the importance of confidentiality on the blog:

Implicit rules-confidentiality of other bloggers (some of whom are reporting from affected nations- where we had direct info from people using sat phones at huge risk) (interview, June 10, 2012).

Personal security became an increasingly important issue as the blog(s) continued and the community gained more attention and participation. Therefore, the majority of the bloggers refrained in revealing their identity or any information about their associations. This was not the case of the interim blog, where the pro-popular opinion of bloggers was flagged and their identities were hacked into, changed, or deleted. Hence, it was apparent that the flagging system was not working and that an active moderator was needed in order for the blog to function properly.

As a result, various unofficial rules were placed to manage bloggers' participation. For example, there are specific rules that bloggers should adhere to whenever they use the site as a conduit for information provided on other social media sites. In addition, there are several rules such as do not steal identity and do not repeat posts multiple times (spamming), that prohibit bloggers in disrupting the activity on the blog and the frequency of individual posts. Adding "off-topic" as a heading is also used to control bloggers comments about critical topics such as politics or religion.

Despite all these efforts, there was a lack of a proper structured system of sanctions to be enforced. As mentioned above, flagging was used to identify an inappropriate comment to moderators to investigate in and then take the decision in either warning the blogger or prohibiting him/her from participating. Flagging was a useful method to monitor bloggers comments or posts, but unfortunately it was not always consistent which led to major consequences.

4.3 The System Falls Apart

The interim site broke down due to lack of moderation and an infiltration of another blogging system aiming to disrupt the existent blog for undisclosed reasons. To this end, the community shifted to an alternative site that was monitored by volunteers mostly from the blogging citizenry in the previous systems. This new system seemed to flourish when effective monitors on blogs were appointed. In addition, participants were given more freedom in providing feedback to the monitors. They were also invited to create rules as long as these new rules did not put participants at risk.

4.4 Virtual Wars

During its lifespan, the interim blog was periodically under attack. These attacks took the form of spamming and identity theft. Jessi, a frequent poster, described one of the attacks: "Just a warning We are being slowly, insidiously invaded at this moment here...look around you" (Interim blog," January 28, 2012). During this particular attack the spammers were reposting the same message frequently and using the usernames of long-term members. The bloggers rectified this issue by ensuring active moderation on the alternative blog. As a result, the alternative blog was not so inundated by spammers, in contrast to its predecessor:

Me: Are you guys... able to stop spammers more than on the other site?

Felix: Blog admin is our friend (-: (interview, June 6, 2012)

The administrators on the alternative blog consisted of both moderators and bloggers who regularly monitored the blogs traffic.

The spam attacks outwardly appeared to be a concerted attack from a professional spammer. As a result of the spamming, certain functions on the blog stopped working or were limited in their utility such as the "like" button, used to indicate agreement with a post. Several bloggers suggested various techniques to fool the spammers. For instance, one blogger recommended that everyone take an avatar with an anonymous mask instead of using a personal avatar:

Craig: Can I make a serious suggestion? With all this avatar stealing, why don't we all adopt the anonymous mask? That way you have to focus on the name, rather than on the "face"? Will make the avatar thieves very obvious!

Lee: why are real posters being moderated when trying to counter endless spam from trollbots, and yet the very same trollbots continue their spamming and identity theft? Why?

Karen: Stephen, you go to Edit Profile under the tab above. And then in the "Accounts" tab in that you can change posting name (Interim Blog, January 28, 2012)

The author knew that this would not stop the spamming but hoped that it would curtail spams by confusing the perpetrator(s). To carry this out, users must be aware of the technical side of the blog. For those who were not tech-savvy or familiar with the intricacies of the blogging environment this was a time-consuming process.

As the interim blog became more popular, it attracted spammers, who effectively disrupted the blogs without much interference from blogging moderators:

Sean: ‘When the blog became the frontline...., along with the escalation in the ‘real’ world went to work on disruption in a military manner of efficiency... The spam worked effectively and the blog supplier did not have the electronic capacity to control it. You have to remember that as a real-time outfit, there are real time results. The group coal- leased with like-minded people like me, some activists of long standing, some not [sic] (interview, June 18, 2012).

The staff at the interim blog did not provide the manpower nor the technical support to stop the spammers. This was a ‘propaganda war’, as the blogger wrote, and it was a war that the spammers apparently were winning on the interim blog which shut down shortly after this post was written.

4.5 Migration

The evolutionary nature of online communities was reflected when there was a migration in search of better management and autonomy in line with communal goals. It was evident that online communities tend to form with common overarching goals arising from a need to exchange ideas, support and learn from each other. In line with constructivist theory, bloggers wanted to co-create rules of engagement and to advance their communal goals. Created and managed by a news organization, the initial and the interim blogs served as learning venues for bloggers but in the end lacked any real input from the citizen bloggers.

Bloggers abandoned the interim blog because it was administered by the umbrella of a news organization which proved to be weak in moderation, enforcing rules and sanctions, responding to feedback, and providing a forum for unrestricted discussions. Utilizing knowledge gained from previous experiences, bloggers created their own blog with co-created rules and an active moderation system.

5. CONCLUSION

The purpose of this paper was to investigate on the type of rules that guide interaction in the online communities as well as creating guidelines that could maintain a sustainable blogging community. Ten bloggers participated in the current case study. Multiple qualitative data collection methods were used. Data was analyzed through taxonomic analysis to identify categories and themes. Findings showed that some set of rules have destabilized the site to some degree. However, adherence to others held the site together. Rules such as those touted by the citizen bloggers – no fighting and staying on topic, no repetition of posts, no cursing – were followed by most of the bloggers. Rules, such as not including links to other websites in the posts were not followed consistently. This specific rule created controversy on the blog as many felt it was unfair and at least a contributing factor into why the alternative site was created. The lack of adherence to rules such as no identity theft and no repetition of posts ultimately led to the demise of the initial blog which was never reestablished after it closed in May, 2012.

Additionally, it was found out that when bloggers participated in producing rules and giving feedback, this helped with creating a monitoring system which proved to be effective for the longevity of a blogging system. Based on this study, a visible monitoring system with a clear understanding of how the rules are created and who takes responsibility for these rules may have allowed the blogging community to continue on the site. Hence, the data collected in this study deemed necessary for a sustainable blogging community based on the following conditions:

1. Active and noticeable moderation
2. Free but responsible comments
3. Flexible and fluid rules
4. Creation of rules by bloggers
5. Clear sanctions

These findings indicate that sustainable online communities can be maintained by allowing flexible, but continuous and observable moderation that adheres to the rules and boundaries created by the community. All the aforementioned findings are similar to the rules presented by Arshad and Chuhalev (2007) and Kavada (2015) that are usually applied in online communities. Kavada, in particular, found that rules should be clearly defined early on and that topics should not deviate from the established purpose. In a related study, virtual communities were found to follow socially constructed norms that closely mirrors real life contexts which is the essence of our findings (Kaplan & Haenlein, 2010).

This study also found that the community adopted both formal and informal rules that the participants had to follow. In alignment with Benoit et al (2016), these numerous types of unwritten rules in online communities became written rules after the final migration to another blogging system. Additionally, as the need for consistent management of the community became apparent, informal rules moved into the realm of formal rules. On the initial blog, the rules were not always clear. On the final blog the rules were clearly stated on the introduction page in the final blog as were the consequences for violating the rules. This shows the importance of clarity regarding the code of conduct and how it guided the bloggers.

Another finding was that formal standards set by professional organizations focused primarily on safety and legal issues. However, the informal standards developed by citizens highlighted other issues as well – such as staying on topic and cultural sensitivity.

6. LIMITATIONS

The major limitation of this study was the lack of input from different types of stakeholders. While we interviewed and followed participants who varied in their activity on the blogs, we were unable to interact with staff at the umbrella news organization or the lurkers on the blogs. Knowing their opinions of, understanding of or preference for the rules on the blogs would have been beneficial for the quality of our findings. While the outcomes of the study are not generalizable, the results are transferable and could be applicable to other similar contexts, situations and populations.

7. FUTURE RESEARCH AND FINAL THOUGHTS

This study was significant in highlighting the necessity of active and consistent moderation in online communities. Participants need to be part of the development of rules and to feel empowered through their opinions and through their feedback. Future research should focus on how to further improve moderation systems for the administrators, as well as investigating new ways to encourage bloggers to adhere to the rules. In addition, to test reliability of the findings, other types of blogging communities should be researched as well.

Another avenue for research could be an in-depth qualitative study into the nuanced interpretations of cultural sensitivity and staying on topic. It would be helpful to better understand what is exactly meant by cultural sensitivity and appropriate topics. Furthermore, future research could focus on investigating how strictly these sensitive topics are observed when compared to defamatory remarks made on social media sites.

Using blogs for expressing free opinions and at the same time ensuring that rules which have been developed collaboratively is not easy to accomplish. Blogging communities, such as the one in this study, could be an outlet for responsible and constructive discussions. However, to achieve this, mechanisms are needed to offer: 1) resolutions to conflicts, 2) quick and thoughtful feedback, 3) a clear indication of the rules and 4) a transparent system of sanctions.

REFERENCES

- Arshad, J., Cuhalev, J. (2007). Rules in online communities: development of heuristic guidelines. Report for the Department of Computer and Systems Sciences Stockholm University / Royal Institute of Technology.
- Benoit, S., Bilstein, N., Hogreve, J., & Sichtmann, C. (2016). Explaining social exchanges in information-based online communities (IBOCs). *Journal of Service Management*, 27(4), 460-480.

- Blanchard, A. (2008). Definition, Antecedents, and Outcomes of Successful Virtual Communities. Retrieved September 10, 2011 from http://www.psych.uncc.edu/alblanch/Blanchard_VC_ENC.pdf.
- De Souza, C. S., & Preece, J. (2004). A framework for analyzing and understanding online communities. *Interacting with Computers*, 16(3), 579-610.
- Fisher, B., Margolis, M., & Resnick, D. (1996). Breaking ground on the virtual frontier: Surveying civic life on the Internet. *The American Sociologist*, 27(1), 11-29.
- Gallant, L. M., G. M. Boone, and A. Heap (2007). Five heuristics for designing and evaluating web-based communities. *First monday* 12.
- Godwin, M. (1994). Nine principles for making virtual communities work. *Wired*, 2(6), 72-73.
- Graham, C. R. (2003). A model of norm development for computer-mediated teamwork. *Small Group Research*, 34(3), 322-352.
- Hsu, C.-L. and H.-P. Lu (2005). Consumer behavior in online game communities: A motivational factor perspective. *Computers in Human Behavior* (23), 1642-1659.
- Johnson, C. M. (2001). A survey of current research on online communities of practice. *Internet and Higher Education* 4, 45-60.
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*, 53(1), 59-68.
- Kavada, A. (2015). Creating the collective: social media, the Occupy Movement and its constitution as a collective actor. *Information, Communication & Society*, 18(8), 872-886.
- Kollock, P. (1998). Design principles for online communities. In *Harvard Conference on the Internet and Society*, pp. 58-60.
- Ostrom, E. (1990). *Governing the Commons*. Cambridge University Press.
- Preece, J. (2001). Sociability and usability in online communities: Determining and measuring success. *Behavior and Information Technology*, 20(5), 347-356.
- Rheingold, H. (1994). A slice of life in my virtual community. In L.M. Harasim (Ed.) *Global Networks: Computers and International Communication* (pgs. 57- 80). Cambridge, MA: MIT Press.
- Social Networking Fact Sheet. (2013, December 27). Retrieved December 16, 2014, from <http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/>
- Spradley, J. P. (1979). *The ethnographic interview*. Holt, Ri.
- Wang, Y. and D. R. Fesenmaier (2004). Modeling participation in an online travel community. *Journal of Travel Research* 42.

DESIGNING AN IOT-ENABLED GAMIFIED APP FOR EMPLOYEE ENERGY CONSERVATION AT THE WORKPLACE

Dimosthenis Kotsopoulos¹, Cleopatra Bardaki¹, Thanasis G. Papaioannou²,
Stavros Lounis¹ and Katerina Pramatari¹

¹*ELTRUN e-Business Research Center, Department of Management Science and Technology, Athens University of Economics and Business, Athens, Greece*

²*Services, Technologies and Economics (STEcon) Laboratory, Department of Computer Science, Athens University of Economics and Business, Athens, Greece*

ABSTRACT

Energy conservation in workplace environments is recognized as both important and impactful towards reducing worldwide CO₂ emissions and protecting the environment. However, the focus of existing research has not primarily been on the employees' energy consumption behavior, albeit its potential impact on workplace energy efficiency. Aiming to affect employee energy behavior change towards more conscious consumption patterns, we have identified gamification, a contemporary instrument that, when carefully utilized can lead to employee behavior change. Furthermore, we have followed a User Centered Design (UCD) approach towards designing a gamified IoT-enabled mobile app that provides energy consumption related feedback to employees at their workplace. We present the characteristics of the designed app, as well as explain the rationale behind their choice. Usability results derived from employees in 3 workplaces at 3 different EU countries indicate that through our approach we have designed a gamified solution that is very well perceived and engaging to our target audience; scoring high in both usability, as well as user experience aspects.

KEYWORDS

Gamification, Energy Conservation, Employee Motivation, Workplace, User Engagement, Incentives

1. INTRODUCTION

Electricity is the preferred energy source in the commercial sector, accounting for 53% of the energy demand in 2012 and expected to reach 62% in 2040, with an average yearly growth of 1.6% – the fastest among all sectors (Conti *et al.*, 2016). Moreover, the buildings sector accounts for 20% of the total delivered energy consumed worldwide (Lülf and Hahn, 2013). The need for energy conservation hence arises and has also been stressed through the Paris Agreement on climate change, as well as the Montreal Protocol on ozone depletion –targeting energy efficiency in buildings as a means to achieve broader environmental protection goals (UN News Centre, 2016; IEA EEfD, 2017). However, energy production worldwide has more than doubled since the early 1970s, along with CO₂ emissions from fuel combustion (IEA, 2017). Buildings in specific are responsible for 40% of the total energy consumption in the EU and 36% of CO₂ emissions, while one third of this demand can be attributed to non-residential buildings (European Commission, 2017).

Energy efficiency – specifically in companies – is additionally important as it can lead to improvements in worker comfort, product quality, overall flexibility and productivity, as well as reductions in maintenance cost, risk, production time and waste (IEA EEfD, 2017). Furthermore, improving the operational efficiency of buildings by using real-time data could lower total energy consumption, while the largest potential savings are in heating, cooling and lighting, which together represented more than 60% of total final energy demand in buildings in 2015 (IEA Digitalization & Energy Working Group, 2017). With technological innovation creating new opportunities for energy efficiency and digitalization increasingly becoming impactful on the energy sector (IEA EEfD, 2017), the total public energy R&D budget in IEA countries has more than quadrupled, from 4% to 21% since the 1970s (IEA Energy Data Centre, 2017). The growing application of

ICT (and IoT technologies) in energy systems that could cut energy use in buildings between 2017 and 2040 by as much as 10%, utilising real-time data to improve operational efficiency (IEA Digitalization & Energy Working Group, 2017).

Further to technological means, the behavior of buildings' occupants is considered an important factor in the consumption of energy, as it can add – or save – a third to a building's designed energy performance (Nguyen and Aiello, 2013), while the amount of energy consumed in EU countries was over 5.600 kWh per employee on average in 2015 (ODYSSEE, 2015). However, limited literature exists on the behavioral aspects of energy conservation at work (Scherbaum *et al.*, 2008), and employees' energy use at an individual behavioral level of analysis (Bansal and Gao, 2006), while little is known about how organizational context affects employee energy-saving behaviors (Lo *et al.*, 2012). Notably, employee behavior has also been found to (adversely) affect the effectiveness of technical measures implemented to save energy at work (Lo, Peters and Kok, 2012). Gamification – defined in its most widely accepted definition as “the use of game design elements in non-game contexts” (Deterding *et al.*, 2011) – has been employed towards increasing occupants' motivation for energy conservation (Reeves *et al.*, 2012; Knol & De Vries, 2011; Brewer *et al.*, 2013; Geelen *et al.*, 2012; Orland *et al.*, 2014; Bourazeri & Pitt, 2013), with reductions in energy consumption in the range of 3-6% recorded and more than 10% achievable, as reported in a comprehensive review of relevant published studies (Grossberg *et al.*, 2015).

Inspired by the above mentioned facts, we decided to design a gamified mobile application for energy conservation at the workplace, by following a User Centered Design (UCD) approach. Our research is conducted in three different workplaces in different EU countries: a museum in Luxembourg, a public agency in Spain and a municipal service in Greece. Next, we present related literature, briefly discuss our findings from the user requirements analysis we conducted and present the design characteristics of our gamified approach towards conserving energy at the workplace through IoT-enabled gamification. Ultimately, we illustrate design guidelines for a personalised gamified application that dynamically provides feedback to employees towards saving energy at work, based on informed design choices.

2. BACKGROUND

2.1 The Human Factor in Workplace Energy Conservation

With digital technologies rapidly changing, questions arise about how technology, behavior and policy will evolve over time and how they will impact energy systems in the future (IEA Digitalization & Energy Working Group, 2017). The human factor alone in energy conservation has been studied, from a wide range of disciplinary perspectives, since the oil shocks of the 1970s (Stephenson *et al.*, 2010). It has been found that energy conservation through behavioral change should be considered alongside the deployment of technological improvements aimed at reducing energy consumption (Delmas *et al.*, 2013), while energy in commercial buildings is mostly consumed through heating and cooling systems, lights, refrigerators, computers, and other equipment (Conti *et al.*, 2016). However, the role of occupants' behavior has in general been largely overlooked in energy consumption analysis so far, even though it has been known to also significantly affect the successfulness of technology-based energy efficiency interventions (Lutzenhiser, 1993). More importantly, a limited number of studies on energy conservation in workplaces are available, compared to households, with few examining individual employees' energy consumption behavior and none including inter-organisational comparisons (Lo *et al.*, 2012). Altruistic motives (e.g. supporting the organization in energy and monetary savings, contributing to environmental protection, complying with peer expectations) have been found to be more salient towards engaging employees to conserve energy at the workplace, due to the fact that no personal monetary gains are normally expected (Matthies *et al.*, 2011). Therefore, promising means include educating employees in energy conservation actions, altering organisational procedures and norms, and utilising feedback to increase employees' awareness of their own behaviour and its consequences (Lo *et al.*, 2012). Additionally, behavioural interventions employing feedback have led to 5-15% in energy savings through direct and 0-10% through indirect feedback (Darby, 2006). At the same time, a meta-analysis of energy conservation experiments conducted between 1975-2012, revealed that the use of feedback has led to 7.4% reductions on average and, in contrast, monetary incentives to a relative increase in energy usage by the participants (Delmas *et al.*, 2013). Finally, tailored information was found to be more effective towards energy behaviour change (Matthies *et al.*, 2011).

2.2 Workplace Gamification towards Energy Conservation

Gamification, “a relatively new instrument in the orchestra of motivation” (Kotsopoulos *et al.*, 2016), can lead to behavioural change by utilising positive emotional feedback, break existing habits and update them with new ones by continuously setting appropriate stimuli (Blohm and Leimeister, 2013). Furthermore, it can be used to increase employee participation, improve performance and compliance in specific goals (Seaborn and Fels, 2015), and enhance employee satisfaction (Robson *et al.*, 2015). Additionally, by utilising gamification, traditional organizational processes can be turned into fun, game-like experiences (Robson *et al.*, 2014), leading to behavioural change, increased and sustained employee motivation, engagement and productivity within an enterprise (Webb, 2013; Pickard, 2015). Examples of organizations that are using gamification to motivate employees towards varying targets at the workplace include the U.K.’s Department for Work and Pensions (Burke, 2014), Deloitte (Huang and Soman, 2013) and IBM (Erenli, 2013). More importantly, the engagement mechanisms used in popular games can also be leveraged to promote real-world energy saving behaviours (Reeves *et al.*, 2012). Examples of energy efficiency games that have already been deployed in workplace environments include “Cool Choices”, “WeSpire”, “EcoInomy” and “Carbon4Square” (Grossberg *et al.*, 2015) – WeSpire has led to over 5 million positive actions in 45 countries (WeSpire, 2017), while Cool Choices has helped over seven thousand employees, in organizations across multiple industries, to increase their energy savings through almost 260.000 energy saving actions (Cool Choices, 2017).

However, the effectiveness of gamification relies on leveraging the psychology of motivation to encourage players to play (Ashridge, 2014). Furthermore, in a utilitarian setting, engagement by gamification can depend on the motivations of users and the nature of the gamified system (Hamari *et al.*, 2014). Therefore, understanding the individuals that are involved in a gamified experience is fundamental (Robson *et al.*, 2015) and a user-centred approach should be followed, characterized by a focus on the needs and desires of end-users, in the design of gamified systems (Seaborn and Fels, 2015). Hence, gamification must also be designed to match the target users’ individual characteristics and preferences in game mechanics, towards increasing motivation for certain behaviours (Uskov & Sekar, 2015; Werbach & Hunter, 2012).

2.3 User Centered Design in Gamification

Creating a game that establishes immediate and continued motivation to continue playing over long periods of time is a very complex issue. The practice of creating engaging, efficient user experiences is called User-Centered Design (UCD) and it entails taking the user into account during product development (Garrett, 2011). Furthermore, it is a design approach, widely considered the key to product usefulness and usability, based on the active involvement of users towards improving the understanding of user and task requirements (Mao *et al.*, 2005). More importantly, UCD methods have been employed by commercial game companies, such as Ubisoft Entertainment, Electronic Arts and Microsoft, during their game development process, to make their products more pleasurable and enjoyable (Pagulayan *et al.*, 2002). Employing UCD in gamification has in fact been noted as so important that, “meaningful gamification” has been defined as “the integration of user-centered game design elements into non-game contexts” (Nicholson, 2012). In organizational settings it entails putting the needs and goals of the users over the needs of the organization, and is expected to result in longer-term and deeper engagement between participants, non-game activities, and supporting organizations (Nicholson, 2012).

The most successful mobile games fit efficiently and effectively into their users’ lifestyle, don’t require prolonged concentration, allow busy players to pause the game as needed, are challenging, but don’t require special experience or knowledge to be successful at the start of the game (Pagulayan *et al.*, 2002). Furthermore, a scoring system should bear a connection to the underlying activity (energy-saving in our case) to make it meaningful for the user (Nicholson, 2012). Interestingly, the most common way game designers employ to balance out the competition – allow less skilled players to compete effectively with more skilled players – is to design team games (Pagulayan *et al.*, 2002). Additionally, by providing multiple ways to achieve within the gamification system, users can select those that are meaningful to them while, by making each game element meaningful in a different way, we can increase the chances that we tend to each user’s needs in the game (Nicholson, 2012).

3. USER REQUIREMENTS ANALYSIS

3.1 Approach

We adopted a UCD approach to elicit the requirements of employees via a multi-step process. As a first step, we visited our prospective users' work environments, towards observing their everyday working routines, as well as the extant opportunities for energy-saving therein. We examined the electrical infrastructure and devices, as well as their impact on energy consumption, and parameters that may affect future gameplay scenarios. Additionally, the daily work schedule and work characteristics (i.e. sedentary / on the move / in front of PC) of the employees were recorded. The working hours in each facility were noted, so as to derive when the gamified app should be providing content to the end users. Furthermore, we surveyed building characteristics that are relevant to our application, such as the orientation of each office space relative to the sun (to derive lighting conditions throughout the day and prepare appropriate lighting fixtures-operating feedback). We also noted shared electrical equipment (printers, coffee makers, etc.) and shared spaces vs. individually used offices. As per the electrical infrastructure, a thorough survey was made, to record the characteristics of the facility, so as to prepare the deployment of IoT infrastructure that will monitor energy consumption on a near-personal basis (Papaioannou *et al.*, 2018).

Having concluded our survey of the premises, we proceeded to interview a representative sample (more than 10% of the employees at each workspace), towards eliciting their own personal needs and preferences in a game that would be designed to motivate them to conserve energy at work. We kept notes during the interview process, covering aspects of the employees' preferences and preconditions for participating in our energy saving initiative. Furthermore, we explored their personal conception of their daily schedule, how they conducted their work, as well as what kinds of energy-saving opportunities existed within their work environment. We delineate our findings briefly in the next sections, while a more detailed account of the UCD process we followed, the user requirements analysis phase and the insight we gained towards designing our gamified solution can be found in (Kotsopoulos *et al.*, 2017; Lounis *et al.*, 2017).

3.2 Energy-Saving Opportunities and Game Design Insight

Consistent with existing literature (Nguyen and Aiello, 2013), we identified the following opportunities for energy saving in the workplaces we surveyed, both through interviews, as well as on-site visits: (i) turning the lights off when leaving the workplace or ambient light suffices (e.g. near the windows on sunny days), (ii) turning the printers off when unused, (iii) turning PC screens off when unneeded, (iii) turning PCs off when leaving the workplace, or away for prolonged periods of time (e.g. during lunch, meetings or on-site technical visits), (iv) turning lights off in commonly used areas, when they are vacant or during afterhours, (v) operating kitchen area equipment in an energy efficient way (e.g. kettle, coffee maker), (vi) operating thermostats optimally, so that the temperature remains within acceptable levels and energy is consumed for heating/cooling optimally, (vii) making sure that windows are kept closed while the rooms are heated or cooled/air-conditioned, so that conditioned air is not freely dissipated to the environment with the relative energy consumed lost in the process, and (viii) using the stairs instead of the elevator when ascending/descending floors at work. These energy-saving opportunities were translated into corresponding in-game challenges and packaged into fixed timeframe bundles, as described in the next sections, to further support long-term game use and engagement, as well as increase energy-saving motivation.

As per the app design itself we found that, in our pilot sites, both personal and collective actions, individual and team play should be considered, while – since only a few of the employees play games – a basic game play could be more effective in energy behaviour change. The interviewees in their majority (61.5%) opted for a team-based game scenario of pursuing goals vs a single player scenario (38.5%). At the same time, 65% believed that rewards are not necessary for being energy efficient because they feel that it is their responsibility to be energy conscious. Furthermore, we noted the participants' need for utilizing real energy consumption data in the gameplay. More specifically, we found that incentives should be allocated on – and be proportionate to – actions that can be validated by the system. At the same time, in team-based scenarios, they should be allocated according to the contribution of each participant / team-member.

4. THE GAME DESIGN

Based on the UCD approach we followed in the requirements analysis phase, we designed a game to fit our participants' needs and preferences. The game is designed to provide feedback to the users towards adhering to energy efficient behaviours. The adherence to these behaviours is corroborated through a specially designed IoT-enabled architecture and platform described in (Papaioannou *et al.*, 2018).

4.1 The Game Concept and “Persona”

The core concept of the app revolves around a virtual living and evolving main “Persona” in the form of a tree, that represents the individual and team-based effects of the energy consumption behaviour of the users on all the energy consuming devices in their vicinity of operation. The function of this visualisation scheme is to motivate the user to actively and continuously take part in the challenges provided by the app towards energy conservation.

In accordance to the users' preferences, the gamified app facilitates team play. A growing tree with vivid elements represents the user's teams' activity, as well as achievements within the gamified app, growing according to each team's performance. Apart from growing, the features of the tree also become more detailed and enriched and vivid elements in the form of birds that occasionally reside on the tree according to a team's performance. Animations are used to make the concept more vivid, attractive and motivating for the end users. The tree grows and is enriched with vivid elements as a result of challenges taken up and completed by users and their teams. The higher the team score, the more the team tree grows, while the more challenges completed by the team, the more the vivid elements (birds) that reside on it. To preserve the scalability of the visualisation concept, more advanced vivid elements (birds) can replace less advanced, so that the team tree may not become cluttered.

The game interface is designed to also combine textual information with the tree concept, so that team score/rank and personal position/rank in the game can be viewable at all times on the top part of the interface. Below the tree resides a scrollable area, visible when the user swipes up on the screen, containing information on ranking and scores, in the form of a leaderboard, as well as recently accepted, won/lost challenges. An additional view exists containing a comprehensive achievements history per team and user, statistics regarding challenges, team positions and badges earned. The time horizon for the visualisation concept is four to six months, to enable users to reflect on their own behaviour, and form the basis for long-term energy behaviour change. By the end of this period, a team can have achieved a fully grown, blooming tree, enriched with birds (badges), by systematically completing in-game challenges. Screenshots from the app, illustrating the visual representation of the forementioned game features in the app UI can be found in Figure 1.



Figure 1. Screenshots of the app interface. As the user advances in the game, the tree “persona” grows in stages. The list of active challenges are visible through the UI at any given point in time

4.2 In-Game Team Formation

Based on the insight we gathered, we derived that the gamified app should facilitate game play. Teams can be formed for the purpose of the game with the following criteria: (i) *Geographical*: Employees working in the same shared office space belong to the same team. In cases where a number of individual offices/rooms exist next to each-other, employees in different rooms with similar functions can belong to the same team. Finally, the employees may belong to a team competing vs other workplaces/buildings. (ii) *Role-oriented*: Employees of the same department / with similar work descriptions may be grouped in a team. (iii) *Device-oriented*: In cases where energy-consuming devices (such as printers / air-conditioners, lights, windows) are shared by employees, these users may be teamed together.

Employees may concurrently belong to the following teams: (a) *The pilot site team*: This is a global pilot site team, to facilitate cross-organizational teams, multi-national company settings, etc. (b) *Team within the workplace*: This is the main team scheme in the game and employees are foremost considered to belong to this team. As teams that work together in their everyday working life operate on same offices, act upon common energy using devices and tend to have similar energy needs this team formation approach is considered the main team formation approach. (c) *Sub-teams according to device usage*: By reversing the focal point of energy consumption from the employees to the energy-consuming devices, team formation may also be conducted around the commonly shared devices.

All points accrued will be reflected on both the participants' "*Team within the workplace*" and individual scoreboards, available for viewing interchangeably from the users. Comparison between Teams is made based on their position in the team leaderboard.

4.3 Types of In-Game Challenges

Whenever an energy saving action is available, a visual notification informs the users of its availability. Each time the challenge is in turn performed by a player, points are credited both to their personal, as well as their teams' profile. Two main categories of in-game challenges can be accepted in the app. *Personal challenges*, like turning ones' PC off when leaving from work and *Team challenges*, like turning the workspace lights off when leaving for the afternoon. Team challenges although taken-up by individual participants, are enacted on behalf of all the team members in the team. The outcome of all challenges is mirrored both on the personal and team progress in the game, while the points accrued by individual actions will also perpetually be added to their teams' scores. Finally, all suggested actions are verified by the system when users swipe on specially located NFC tabs while performing them.

The initial set of main challenges that run each day over the course of the game and around which the game revolves are:

(i) ***Morning challenge***: Each morning, the employees compete in a challenge, through which the workplace's energy consumption is initialised. This challenge includes the following sub-challenges: "*Check in Challenge*": A base challenge for the employees to declare that they have arrived and want to start playing., "*Elevator up*" challenge: Upon entering the building, the employees opt for using the stairs, instead of the elevator, to reach their office., "*Lights on*" challenge: The minimum lights that are needed in their workspace are lit by each team in the morning., "*Coffee on*" challenge: Each team prepares a unique load of coffee to share in the morning., "*Kettle*" challenge: Each team uses the kettle once, based on their common need in hot water (i.e. for tea) in the morning., "*Equipment on*" challenge: Upon arriving at their desk, the employees turn on any office equipment they need for the day (PCs, printers, etc), at the same time swiping their tags.

(ii) ***Daily challenge***: During the day, the employees are prompted to perform specific energy saving actions, depending on environmental variables (temperature, luminosity), as well as occupancy (leaving the office) based on the available sensors for real time monitoring and/or the work schedules and division routine: "*Temperature*" challenge: Whenever the temperature is higher than the optimum set level of temperature comfort in the winter – or lower than optimum in the summer – the users are prompted to adjust the thermostat of their office accordingly. The optimum temperature levels shall be set in the beginning of the game, based on the World Health Organisation (WHO) optimum temperature suggestions for workplaces in the summer and winter., "*Illumination*" challenge: According to the lighting conditions within each space, the users may be prompted to turn the lights off (when there is ample ambient light, based on luminosity

sensor readings)., “*Windows*” challenge: The users are prompted to close the windows when the air conditioners are on, or to close the air conditioners when the windows are open (based on contact sensors)., “*Away*” challenge: Whenever employees are away from their office, they are prompted to switch off any unnecessary devices (employee presence at their office corroborated by NFC Beacons).

(iii) *Evening / Leaving office challenge*: The employees compete in a challenge, through which the workplace’s energy consumption is minimized for the night. Sub-challenges included: “*Elevator down*” challenge: Upon leaving the building, the employees opt for using the stairs, instead of the elevator, to leave their office., “*Lights off*” challenge: The lights are switched off in the workspace by each team, by the last team member leaving the office., “*Coffee off*” challenge: Each team turns off their coffee maker before leaving the office in the evening., “*Equipment off*” challenge: Before leaving their desk, employees turn off any equipment that isn’t needed afterhours (PCs, printers, etc).

The forementioned daily challenges are also bundled into larger timeframe challenges, to increase user engagement. Hence, weekly and monthly challenge versions exist for all the different challenges and sub-challenges presented. For example, the employees try to adhere to the same daily energy saving challenges on all days of a week, to attain the corresponding “weekly challenge”. Similarly, at the end of each month, the repeated adherence to actions will be appraised towards winning “monthly challenges”.

4.4 Individual and Team Accomplishments / Badges

At the end of each time period (e.g. Week / Month), the team that performed best in the game shall be given the title “*Energy week/month Champion*”. This title is based on the actions performed within each calendar day/week/month and not taking into account the position of teams in the leaderboard. Similarly, at the end of each week/month, the person that has performed best in the game, within their team, will receive the title “*Energy Week/Month Captain*”. Additionally, at the end of each time period, the person that performed second best within each team shall receive the title “*Energy Week/Month Deputy*”. These personal titles shall also be given, based on performance within the week/month and not overall game position.

A star system has also been designed, to visually represent repeated performance in the game. If the position of Energy Champion/Deputy is earned consecutively, a star will be won by the player: (a) *Silver*: for earning the position of Energy Deputy 2 weeks in a row, (b) *Gold*: for earning the position of Energy Captain 2 weeks in a row. Additional badges in the form of birds can be won by teams winning the team challenges. As an example, a team winning the “*Daily Morning Challenge*” will see a bird (badge) arriving on their team tree. By winning the “*Weekly Morning Challenge*”, the bird (badge) will be replaced by an upgraded version (larger / more vividly animated). The same scheme stands for all types of fixed-time team challenges that lead to earning bird badges. However, to make the type of reward visually different, different types of birds will be won for performing different challenges.

4.5 Personal Progression in the Game

To enable the experience of personal progression in the game, players receive an “*energy saving rank*”, based on their overall collection of points during the game. Five ranks have been defined, corresponding to different point thresholds, while a small figure of an avatar, indicative of each rank will be visible on the top part of the game interface. The ranks available, denoting personal experience, as well as activity in the game, are: (i) “*Apprentice Saver*” (0-49 points accrued in the game): Having just enrolled in the game, the player has not performed any actions. He is by definition given the rank of “apprentice”, (ii) “*Junior Saver*” (50-99 points): As the game progresses and the player earns the first basic bundle of points, he becomes more experienced in the game and upgrades to the rank of “junior”, (iii) “*Saver*” (100-199 points): Similarly, when a normal level of actions have been performed, the rank of “saver” is attained., (iv) “*Advanced Saver*” (200-499 points): When the player has reached a threshold of points that corresponds to a relatively large number of actions, they may attain the “advance saver” status., and (v) “*Expert Saver*” (500+ points): If the player has spent enough time performing actions in the game, he may reach this status, indicative of truly conscious energy savers. Only the very active players may reach this rank at the end of the game.

4.6 Team Progression in the Game

The teams' progression in the game is mainly visible via the tree persona. The tree grows and is enriched with bird badges according to the energy efficient actions performed and challenges won by the team members. To enrich the experience, team members will also receive feedback regarding their team tree such as, for example, "Your tree has a new bird visitor, because your team remembered to turn off the lights when leaving the office all the days last week". That way, the more active and successful the team members are in the game (according to challenges accepted and completed) the more their team tree is enriched by vivid elements such as birds, and the more points are added to the team score, the more the tree grows. Screenshots of the app illustrating the badges, personal and team progression in the game can be found in Figure 2.

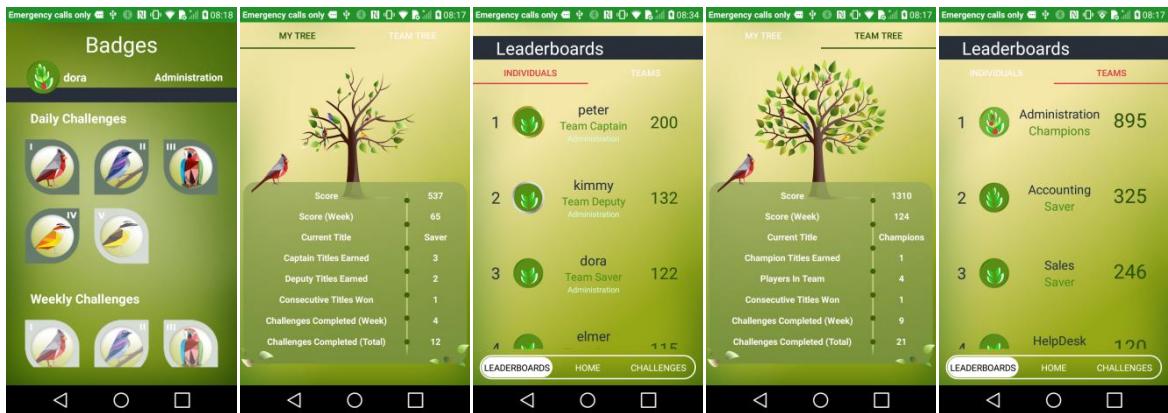


Figure 2. Screenshots of the app interface. The user can review their earned badges, personal and team status, position in the team and personal leaderboards, as well as status of the tree "personas"

5. DISCUSSION

5.1 Player Journey

Based on the characteristics of the game delineated, a user that enrolls in the game can follow different journeys within the game play sphere. Two main paths exist simultaneously: The game can be played both individually and in teams. The basis for progression in the game is gathering points for successful adherence to suggested actions/challenges, also packaged into specific timeframe challenges (daily/weekly/monthly). As a team member, the player can enjoy contributing to the team tree growing, flourishing, and gaining birds (ornamental badges) based on the successful completion of challenges by the respective team members. Furthermore, as both a single player and a team member, the player can experience empowerment by gaining a good position in the respective team and team leaderboards, as well as gaining personal roles within the team and experiencing a feeling of relatedness when gaining team roles. The game logic structure diagram, delineating the different player journeys available within the game is presented in Figure 3.

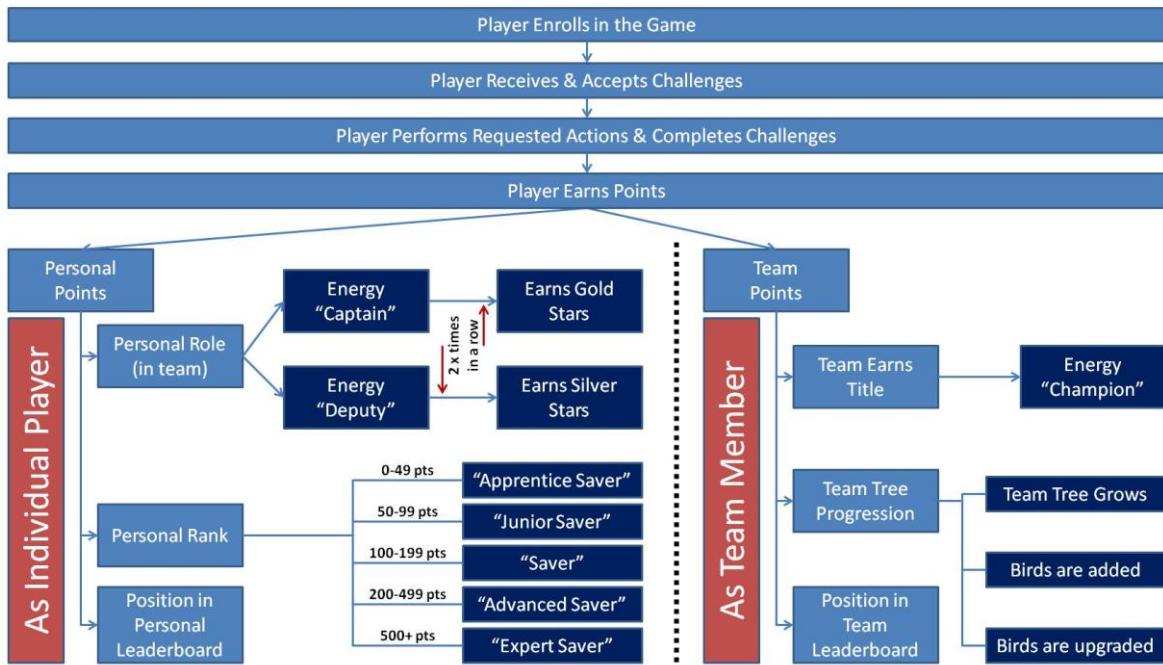


Figure 3. The Player Journey in the Game - Individual and Team Play Scenarios

5.2 Usability Evaluation

To validate the usability and appeal of our designed app to its target audience, we conducted usability tests with users from all three pilot sites. After presenting the aims and scope of the developed solution to the participants and a short video tutorial of the app, we asked them to playtest the mobile app for 10 minutes, on Android mobile phones with the app pre-installed. Consequently, the participants were asked to answer on questionnaires assessing usability and user experience. We employed the System Usability Scale (SUS), using an acceptable threshold of 68 as indicated by (Bangor *et al.*, 2009), as well as the User Experience Questionnaire (UEQ) (Laugwitz *et al.*, 2008), that assesses the users' experience in terms of: (i) Perspicuity (Clear, Easy to Learn, Easy, Understandable), (ii) Novelty (Leading edge, Creative, Innovative, Inventive), (iii) Dependability (Motivating, Exciting, Valuable, Interesting), (iv) Stimulation (Supportive, Meets Expectations, Predictable, Secure) and (v) Efficiency (Efficient, Fast, Organised, Practical). Finally the participants were asked to provide their own additional comments.

We recorded results from N=16 employees across the three sites (10 male and 6 female) with an average age of 41.75 years old. Across all sites, the average SUS score was 75.31, well above the acceptable threshold of 68. Furthermore, the vast majority (75%) of the participants rated usability above the threshold. Regarding the users' experience, based on results from deploying the UEQ, participants on average rated their experience (on a scale from 1- 7) very highly on Perspicuity (6.02), as well as sufficiently highly in Efficiency (4.81), Dependability (5.03), Stimulation (5.22) and Novelty (5.44). The complete results recorded on the SUS and UEQ questionnaires on all three sites, along with the corresponding average scores per site and total average scores can be found on Table 1.

Table1. System Usability Scale (SUS) and User Experience Questionnaire (UEQ) scores recorded in our sample on three different sites (Greece, Spain and Luxembourg)

Site	Age	Gender		SUS	UEQ			
		Male	Female		Perspicuity	Efficiency	Dependability	Stimulation
Greece	47	+		75	7	2.25	3.5	5.25
	34	+		85	7	6.5	6.5	5.5
	40		+	87.5	7	4.5	6.25	6.5
	43		+	95	7	6.75	6.5	6.75
	34		+	60	5.5	3.5	5.75	4.75
Average	39.6	2	3	80.5	6.7	4.7	5.7	5.65
Spain	44	+		77.5	6	5.25	4.75	4.75
	37		+	75	7	4.75	5	4.75
	55	+		57.5	5.25	4.5	4.5	4.25
	44	+		75	6.25	3.75	5	5.5
	29	+		80	6	6	5	5.75
Average	41.8	4	1	73	6.1	4.85	4.85	5.05
Luxembourg	53	+		87.5	5.75	5.25	5	5.75
	35		+	97.5	7	6.25	6.25	6.5
	55	+		35	4.5	5.25	4	3.25
	45	+		47.5	4	4.25	3.25	4.5
	31	+		72.5	5.5	4.25	3.25	4.5
	42		+	97.5	5.5	4	6	5.5
Average	43.5	4	2	72.92	5.38	4.88	4.63	5
Total Avg. (All Sites)	41.75	10	6	75.31	6.02	4.81	5.03	5.22
								5.44

Based on our sample size (N=16), and the total population of office employees in our pilot sites (N=144), as well as the fact that 81.25% of our sample recorded acceptable SUS scores (>68), we deduce that the confidence interval of our recorded results is 20.07% (Creative Research Systems, 2018). Hence, the avg. SUS score for (N=144), is expected to vary between 60.19 and 90.43 (+/- 15.12 from the recorded score of 75.31 for our sample), with a 95% confidence level.

6. CONCLUSION

Following a UCD process, we have designed a gamified app to be employed towards motivating employees to conserve energy at the workplace. After examining the requirements of our prospective end-users, as well as observing the contextual characteristics of their workplace environments, and inherent opportunities for energy saving therein, we proceeded to select the energy related behaviours that should be targeted towards effectively reducing energy consumption. Furthermore, we derived a game design to fit our samples' characteristics and context. A challenge-oriented, primarily team-based game scenario with fixed-timeframe bundled actions, was adopted according to the insight gained from our participants. Usability and user experience related results were well within the acceptable ranges, indicating that the UCD approach we followed led to a potentially successful app.

Our research would be better grounded through practical experimentation in a workplace environment, to verify the effectiveness of the process followed, as well as the resulting outcome application. We aim to proceed towards that direction in the future, by conducting practical experimentation in three different workplaces situated in three different EU countries, featuring the gamified app we designed.

ACKNOWLEDGEMENT

This research study is partially funded by the project ChArGED (CleAnweb Gamified Energy Disaggregation), that receives funding from the EU Horizon 2020 research and innovation programme, under grant agreement No 696170.

REFERENCES

- Ashridge (2014) *Connecting Gamification and Motivation Theory, Virtual Ashridge @ The HR director*. Available at: <https://www.thehrdirector.com/features/gamification/connecting-gamification-and-motivation-theory/> (Accessed: 25 January 2016).
- Bangor, A., Kortum, P. and Miller, J. (2009) 'Determining what individual SUS scores mean: Adding an adjective rating scale', *Journal of usability studies*, 4(3), pp. 114–123. doi: 66.39.39.113.
- Bansal, P. and Gao, J. (2006) 'Building the Future by Looking to the Past: Examining Research Published on Organizations and Environment', *Organization & Environment*, 19(4), pp. 458–478. doi: 10.1177/1086026606294957.
- Blohm, I. and Leimeister, J. M. (2013) 'Gamification: Design of IT-based enhancing services for motivational support and behavioral change', *Business and Information Systems Engineering*, 5(4), pp. 275–278.
- Bourazeri, A. and Pitt, J. (2013) 'Serious Game Design for Inclusivity and Empowerment in SmartGrids', in *First International Workshop on Intelligent Digital Games for Empowerment and Inclusion*.
- Brewer, R., Xu, Y. and Lee, G. (2013) 'Energy feedback for smart grid consumers: Lessons learned from the Kukui Cup', ... on Smart Grids, Green ..., (c), pp. 120–126. Available at: http://www.thinkmind.org/index.php?view=article&articleid=energy_2013_5_40_40157.
- Burke, B. (2014) 'How Gamification Motivates the Masses', *Forbes*. Available at: <http://www.forbes.com/sites/gartnergroup/2014/04/10/how-gamification-motivates-the-masses/>.
- Conti, J., Holtberg, P., Diefenderfer, J., LaRose, A., Turnure, J. T. and Westfall, L. (2016) *International Energy Outlook 2016, With Projections to 2040*. May 2016. Washington, DC, U.S.A.: U.S. Energy Information Administration (EIA). doi: DOE/EIA-0484(2014).
- Cool Choices (2017) *Employee Engagement Sustainability Game*. Available at: <https://coolchoices.com/> (Accessed: 10 April 2017).
- Creative Research Systems (2018) *Sample Size Calculator*. Available at: <https://www.surveysystem.com/sscalc.htm> (Accessed: 25 February 2018).
- Darby, S. (2006) 'The Effectiveness of Feedback on Energy Consumption: a Review for Defra of the Literature on Metering , Billing and Direct Displays', *Environmental Change Institute, University of Oxford*, 22(April), pp. 1–21.
- Delmas, M. A., Fischlein, M. and Asensio, O. I. (2013) 'Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012', *Energy Policy*, 61, pp. 729–739.
- Deterding, S., Sicart, M., Nacke, L., O'Hara, K. and Dixon, D. (2011) 'Gamification. using game-design elements in non-gaming contexts', *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11*, p. 2425.
- Erenli, K. (2013) 'The Impact of Gamification. Recomending Education Scenarios', *International Journal of Emerging Technologies in Learning*, 8(Special Issue 1: 'ICL2012'), pp. 15–21.
- European Commission (2017) *Buildings*, <https://ec.europa.eu>. Available at: <https://ec.europa.eu/energy/en/topics/energy-efficiency/buildings>.
- Garrett, J. J. (2011) *The Elements of User Experience: User-Centered Design for the Web and Beyond*. Second. Berkeley, CA. Available at: <http://books.google.com/books?hl=en&lr=&id=9QC6r5OzCpUC&pgis=1>.
- Geelen, D., Keyson, D., Stella, B. and Brezet, H. (2012) 'Exploring the use of a game to stimulate energy saving in households - Journal of Design Research - Volume 10, Number 1–2/2012 - Inderscience Publishers', *Journal of Design Research*, 10, pp. 102–120.
- Grossberg, F., Wolfson, M., Mazur-Stommen, S., Farley, K. and Nadel, S. (2015) *Gamified Energy Efficiency Programs*. Available at: <http://www.climateaccess.org/sites/default/files/aceee.pdf>.
- Hamari, J., Koivisto, J. and Sarsa, H. (2014) 'Does gamification work? - A literature review of empirical studies on gamification', *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 3025–3034.
- Huang, W. H. and Soman, D. (2013) *A Practitioner's Guide To Gamification Of Education*, University of Toronto - Rotman School of Management. Toronto. doi: 10.1111/j.1467-8535.2011.01259.x.
- IEA (2017) *Key World Energy Statistics*.
- IEA Digitalization & Energy Working Group (2017) *Digitalization & Energy*. Available at: <http://www.iea.org/digital/>.
- IEA EEFd (2017) *Energy Efficiency 2017*. doi: 10.1787/9789264284234-en.
- IEA Energy Data Centre (2017) *Energy Technology RD&D Budgets: Overview*. Available at: <https://www.iea.org/publications/freepublications/publication/EnergyTechnologyRDD2017Overview.pdf>.
- Knol, E. and De Vries, P. (2011) 'EnerCities-A Serious Game to Stimulate Sustainability and Energy Conservation: Preliminary Results', *eLearning Papers*, (45), pp. 1–10.
- Kotsopoulos, D., Bardaki, C., Lounis, S., Papaioannou, T. and Pramatari, K. (2017) 'Designing an IoT-enabled Gamification Application for Energy Conservation at the Workplace: Exploring Personal and Contextual Characteristics', in *30th Bled eConference*. Bled, Slovenia, pp. 369–384.

- Kotsopoulos, D., Bardaki, C. and Pramatari, K. (2016) ‘Gamification , Geolocation and Sensors for Employee Motivation Towards Energy Conservation at the Workplace’, in *Tenth Mediterranean Conference on Information Systems (MCIS) 2016*. Paphos, Cyprus, pp. 1–11. Available at: <http://aisel.aisnet.org/mcis2016/39>.
- Laugwitz, B., Held, T. and Schrepp, M. (2008) ‘Construction and Evaluation of a User Experience Questionnaire’, *HCI and Usability for Education and Work*, pp. 63–76.
- Lo, S. H., Peters, G. J. Y. and Kok, G. (2012) ‘Energy-Related Behaviors in Office Buildings: A Qualitative Study on Individual and Organisational Determinants’, *Applied Psychology*, 61(2), pp. 227–249.
- Lounis, S., Kotsopoulos, D., Bardaki, C., Papaioannou, T. G. and Pramatari, K. (2017) ‘Waste No More : Gamification for Energy Efficient Behaviour at the Workplace.’, in *Proceedings of the 1st International GamiFIN Conference, Pori, Finland, May 9-10, 2017*. Pori, Finland, pp. 129–134. Available at: http://ceur-ws.org/Vol-1857/gamifin17_p17.pdf.
- Lülfes, R. and Hahn, R. (2013) ‘Corporate greening beyond formal programs, initiatives, and systems: A conceptual model for voluntary pro-environmental behavior of employees’, *European Management Review*, 10(2), pp. 83–98.
- Lutzenhiser, L. (1993) ‘Social and Behavioral Aspects of Energy Use’, *Annual Review of Energy Environment*, 18, pp. 247–289.
- Mao, B. J., Vredenburg, K., Smith, P. W. and Carey, T. (2005) ‘User-centered design practice’, *Communications of the ACM*, 48(3), pp. 105–109.
- Matthies, E., Kastner, I., Klesse, A. and Wagner, H.-J. (2011) ‘High reduction potentials for energy user behavior in public buildings: how much can psychology-based interventions achieve?’, *Journal of Environmental Studies and Sciences*, 1(3), pp. 241–255.
- Nguyen, T. A. and Aiello, M. (2013) ‘Energy intelligent buildings based on user activity: A survey’, *Energy and Buildings*. Elsevier B.V., 56, pp. 244–257.
- Nicholson, S. (2012) ‘A User-Centered Theoretical Framework for Meaningful Gamification’, in *Games+ Learning+ Society 8.0*. Madison, WI, pp. 1–7.
- ODYSSEE (2015) *Free Energy Indicators, ODYSSEE Database*. Available at: <http://www.indicators.odyssee-mure.eu/online-indicators.html> (Accessed: 21 November 2017).
- Orland, B., Ram, N., Lang, D., Houser, K., Kling, N. and Coccia, M. (2014) ‘Saving energy in an office environment: A serious game intervention’, *Energy & Buildings*, 74, pp. 43–52.
- Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L. and Fuller, T. (2002) *User-centered Design in Games*. Boca Raton, FL: CRC Press.
- Papaioannou, T., Dimitriou, N., Vasilakis, K., Schoofs, A., Nikiforakis, M., Pursche, F., Deliyski, N., Taha, A., Kotsopoulos, D., Bardaki, C., Kotsilidis, S. and Garbi, A. (2018) ‘An IoT-Based Gamified Approach for Reducing Occupants’ Energy Wastage in Public Buildings’, *Sensors*, 18(2), p. 537.
- Pickard, T. (2015) *5 Statistics That Prove Gamification is the Future of the Workplace*, business.com. Available at: <http://www.business.com/management/5-statistics-that-prove-gamification-is-the-future-of-the-workplace/> (Accessed: 27 February 2016).
- Reeves, B., Cummings, J. J., Scarborough, J. K., Flora, J. and Anderson, D. (2012) ‘Leveraging the engagement of games to change energy behavior’, *Proceedings of the 2012 International Conference on Collaboration Technologies and Systems, CTS 2012*, pp. 354–358.
- Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I. and Pitt, L. (2015) ‘Is it all a game? Understanding the principles of gamification’, *Business Horizons*. Kelley School of Business, Indiana University.
- Robson, K., Plangger, K., Kietzmann, J., McCarthy, I. and Pitt, L. (2014) ‘Understanding Gamification of Consumer Experiences’, *Advances in Consumer Research*, 42, pp. 352–357.
- Scherbaum, C., Popovich, P. and Finlinson, S. (2008) ‘Exploring individual-level factors related to employee energy conservation behaviours at work’, *Journal of Applied Social Psychology*, 38(3), pp. 818–835.
- Seaborn, K. and Fels, D. I. (2015) ‘Gamification in theory and action: A survey’, *International Journal of Human Computer Studies*, 74, pp. 14–31.
- Stephenson, J., Barton, B., Carrington, G., Gnoth, D., Lawson, R. and Thorsnes, P. (2010) ‘Energy cultures: A framework for understanding energy behaviours’, *Energy Policy*. Elsevier, 38(10), pp. 6120–6129.
- UN News Centre (2016) “ Today is an historic day ,” says Ban , as 175 countries sign Paris climate accord’, [www.un.org](http://www.un.org/apps/news/story.asp?NewsID=53756#.WD6x9lzsFf5). Available at: <http://www.un.org/apps/news/story.asp?NewsID=53756#.WD6x9lzsFf5>.
- Uskov, A. and Sekar, B. (2015) ‘Smart Gamification and Smart Serious Games’, in *Fusion of Smart, Multimedia and Computer Gaming Technology: Research, Systems and Perspectives*. Springer International Publishing, pp. 7–36.
- Webb, E. N. (2013) ‘Gamification : When It Works , When It Doesn ’ t’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8013 LNCS(PART 2), pp. 608–614.
- Werbach, K. and Hunter, D. (2012) *For The Win: How Game Thinking can revolutionize your business*. Philadelphia, PA: Wharton Digital Press, The Wharton School, University of Pennsylvania.
- WeSpire (2017) *WeSpire - Employee Engagement Platform Powered by Behavioral Science*. Available at: <http://www.wespire.com/> (Accessed: 10 April 2017).

ORGANIZATIONAL INTEROPERABILITY: AN ASSESSMENT OF THE IMPACT OF THE SEMANTIC LAYER

Asmae Rhazale and Bouchaib Bounabat

*ALQUALSADI Team, ADMIR Laboratory, ENSIAS, Rabat IT Center
Mohamed V University in Rabat, Morocco*

ABSTRACT

In the era of social networks, citizens are now expecting efficient and user-centric public services. Public agencies and governments in general are now pressured to deliver fast-responding and 24/7 available public services.

Thus the interoperability among public agencies is becoming a major concern. The aim is to provide complex and combined public services from “a single window”. This goal can be achieved by establishing seamless exchanges between all the stakeholders. In other words, establishing interoperability at all levels, is important to guarantee efficiency of the final delivery.

Many research efforts have been conducted concerning the different layers of interoperability, especially the technical and semantic ones.

This work aims to define the relationship between the semantic and organizational layers and to measure the impact of the semantic layer on the organizational one.

KEYWORDS

E-governement, Organizational Interoperability, Semantic Interoperability, Enterprise Architecture, Social Network Analysis, Graph theory

1. INTRODUCTION

Interoperability is a multidimensional concept that can occur on multiple levels. Technical interoperability is concerned by establishing connection between machines. Semantic interoperability (SIOP) is about exchanging the meaning of shared data and establishing common ontologies. Organizational interoperability (OIOP) is on a higher level and is concerned by making business processes work together. The aim of this layer is to define how to act on the exchanged data.

Many definitions have been given to OIOP. In this work, OIOP is defined as the interaction between business processes. Kubicek et al. (Kubicek, et al., 2011) have proposed a reconceptualization of the definition of OIOP to make it more “standardized”. The authors redefined OIOP as the “business process interaction” layer and positioned it above the semantic interoperability layer.

This work attempts to formalize the relationship between semantic and organizational interoperability. The aim is to establish a mathematical indicator allowing the estimation of the impact of the semantic layer on the organizational one.

Section 2 presents in brief the multiple definitions given to OIOP and the kinds of barriers that prevent it. The semantic matter can be seen among those barriers. In fact, any exchange of information can be hugely impacted by any misunderstanding of the meaning of the data exchanged.

Section 3 is devoted to the formalization proposed to highlight the impact of the SIOP on the OIOP. It is necessary to determine, as a first step, the link between those two levels. This link could be used subsequently to evaluate the potential of OIOP between entities considering their SIOP.

Section 4 details how the coverage ratio can be used to categorize the business processes ready to interoperate and those that need a “semantic correction”.

Section 5 describes the experimental setup.

Finally, this paper is concluded with some remarks and perspectives for further works.

2. DEFINITIONS OF ORGANIZATIONAL INTEROPERABILITY

The term "interoperability" (IOP) was first introduced by the US Department of Defense in 1977 (DODD 2010.6, 1977) as "*The ability of systems, units or forces to provide services to and accept services from other systems, units, or forces and to use the services so exchanged to enable them to operate effectively together*".

Interoperability can happen at many levels:

- Technical level (TIOP): Connections between machines are set to allow the data flow sharing.
- Semantic Level: The meaning of the data involved is shared between participants.
- Organizational level: The "How to" act on the data exchanged is defined. Common goals are defined and agreements are set. Thus, the tree levels are tightly linked as described below (Figure 1) :



Figure 1. The link IOP levels

In front of the increasing complexity of information systems, improving enterprise interoperability has become a crucial element for better management. To address this issue, several research projects have been launched during the last decade and have resulted in a set of frameworks which help organizing and performing enterprise interoperability projects efficiently (Taoudi, et al., 2013); (Taoudi, et al., 2013)

This last decade, the organizational aspect of interoperability has become also a major concern to improve IOP. Therefore, it's increasingly perceived as a main domain for the research communities as well. Many different definitions have been given to OIOP. Some of them introduce the concept of exchanging and using services between the collaborative parties (Legner & Lebreton, 2007). Kubicek et al. in (Kubicek, et al., 2011) refer to OIOP in alternative terms such as collaboration and cooperation. The authors indicates that "*these terms refer to rules and regulations regarding workflows, duties and liability, cost sharing, and other issues that form the context, in which technical and semantic interoperability is established*".

In addition, the efforts of standardization have been deployed for the other levels of interoperability. International standardization committees have set many standards and protocols for the technical level. Ontologies and other means of semantic interoperability have been developed by the governments' initiatives and industrials.

The organizational level though presents a lack of formalization and standardization. Kubicek et al. in (Kubicek, et al., 2011) proposed a re-conceptualization by introducing a three-dimensional view of the OIOP. Based on the idea that OIOP is concerned by the alignment of BP as mentioned in EIF 2.0 (EUROPEAN COMMISSION, 2010), the standards for BP alignment are similar to those used in the semantic and technical layer. The authors explain that the business process alignment "*can be categorized as an additional layer above semantic interoperability*" (Kubicek, et al., 2011). They propose to "*abandon the term "organizational interoperability" and rather call the technical standards for business process (BP) alignment "business process interoperability" while summarizing the other aspects under the concept of "governance" and "implementation"*".

In the remainder of this paper, we choose the reconceptualization of the organizational interoperability introduced by Kubicek et al. We attempt to formalize the link between OIOP and SIOP to measure how the semantic layer influences the interoperability on the organizational level.

3. FORMAL LINKING BETWEEN THE SEMANTIC AND THE ORGANIZATIONAL LAYERS OF INTEROPERABILITY

As mentioned above, the OIOP is considered as the layer of business processes (BPs). It represents the interconnection between several BPs.

A business process (BP) is considered as set of activities executed in a certain order. The execution of the BP brings into play a group of terms.

In the remainder of the article, the following hypothesis (**a**) is considered: Each BP is supported by at least one application.

Given this hypothesis, the following assertion is true: Each term involved in the execution of BP has an equivalent of semantic data in the semantic layer.

Therefore, there is a correspondence between semantic data and terms. In the following paragraph we attempt to formalize this correspondence.

3.1 Defining the Correspondence Function

The organizational layer is composed of a set of BPs.

Let $n \in \mathbb{N}$ be the number of data in the semantic layer and $S_1, S_2, S_3, \dots, S_n$ the semantic data.

Each term used in the BP has a representation in the semantic layer. In other words, the meaning of the term corresponds to a combination of portions of semantic data.

This correspondence is formalized as below.

Given the hypothesis (**a**) established above, the correspondence function **Cor** is defined as:

$$\text{Cor}: T \longrightarrow [0,1]^n$$

$$t \longrightarrow (\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n)^T / \alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n \in [0,1] \text{ and } \sum_1^n \alpha_k = 1$$

$\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$ are called the coefficients of correspondence. T is the set of terms.

The term t corresponds to a combination of the portion α_1 of S_1 , α_2 of S_2 , ..., and α_n of S_n .

$\sum_1^n \alpha_k = 1$ because the term t could not be "over-represented" with the semantic data.

3.2 Defining the Correspondence Graph

The function **Cor** defined above could be represented with a directed weighted graph.

- The nodes are the terms and the semantic data and the edges are weighted with the correspondence coefficients.
- The relationship between the terms and the semantic data represents a graph $G = (V, E)$ where $V = \{t_1, t_2, \dots, t_p, S_1, S_2, \dots, S_n\}$, $p, n \in \mathbb{N}$. The order of the graph is $|G| = p + n$.

Given the definition of the function **Cor** we have :

$$\text{Cor}(t_1) = (\alpha_{11}, \alpha_{12}, \alpha_{13}, \dots, \alpha_{1n})^T, \text{Cor}(t_2) = (\alpha_{21}, \alpha_{22}, \alpha_{23}, \dots, \alpha_{2n})^T, \dots, \text{Cor}(t_p) = (\alpha_{p1}, \alpha_{p2}, \alpha_{p3}, \dots, \alpha_{pn})^T.$$

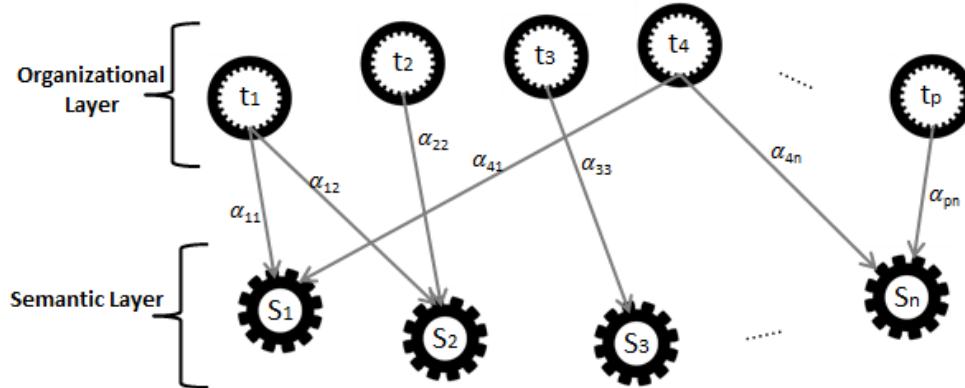


Figure 2. Correspondence graph between terms and semantic data

The graph above (Figure 2) represents the correspondence between the terms and the semantic data. The symbols used to describe terms and semantic data are different. This difference in symbols is chosen in order to highlight the difference between data nature in the two layers. The terms are a complex data that is a combination of atomic semantic data.

The graph G has the following properties:

- Between every nodes t_i and t_j , $i, j \in [1, p]$, there is no edge.
- Between every nodes S_i and S_j , $i, j \in [1, n]$, there is no edge.
- For each $i \in [1, p]$, $j \in [1, n]$ the edge, if it exists, is always directed from the node t_i to the node S_j and weighted and is unique.

The degree of the node t_i is $d^o(t_i) = d^{o+}(t_i) + d^{o-}(t_i)$. Given the properties above, $d^{o+}(t_i) = 0$.

Thus $d^o(t_i) = d^{o-}(t_i)$. Given the hypothesis (a), the assertion (b) $0 < d^o(t_i) \leq n$ is true.

3.3 Defining the Semantical Coverage Ratio

3.3.1 The Semantical Coverage Ratio of a Term

The semantical coverage ratio of a term “SCRT” is a measure determining the representation of the terms by the semantic data.

The higher the ratio is, the bigger is the representation of the term in the semantic level. The lower the ratio is, the less represented is the term. The ratio is defined as followed.

Let $m \in \mathbb{N}$ be the number of BPs present in the organizational layer.

For each business process $BP_i / i \in [1, m]$, $p_i \in \mathbb{N}^*$ is the number of terms in the business process BP_i and $t_{ij} / j \in [1, p_i]$ are the terms used in the execution of BP_i .

For each $i \in [1, m]$, $G_{cori} = (V_i, E)$ is the directed graph of correspondence. $V_i = \{t_{i1}, t_{i2}, t_{i3}, \dots, t_{ip_i}, S_1, S_2, \dots, S_n\}$ the set of nodes and E the set of edges weighted with the coefficient of correspondence.

The order of the graph is $|G_{cori}| = p_i + n$.

We have for each $j \in [1, p_i]$, $Cor(t_{ij}) = (\alpha_{ij1}, \alpha_{ij2}, \alpha_{ij3}, \dots, \alpha_{ijn})^T$.

The coverage “SCRT” of the term t_{ij} is defined as:

$$SCRT(t_{ij}) = \frac{\sum_{k=1}^n \alpha_{ijk}}{d^o(t_{ij})}$$

The SCRT of the term t_{ij} is in $[0, 1]$. The proof is as followed: given the definition of the coefficients of correspondence we have $0 \leq \alpha_{ijk} \leq 1$, thus $0 \leq \sum_{k=1}^n \alpha_{ijk} \leq n$. In addition, given the property (b) we have $0 < d^o(t_{ij}) \leq 1$. Therefore $0 < SCRT(t_{ij}) \leq 1$.

3.3.2 The Semantical Coverage Ratio of a BP

Given the definition above of the coverage ratio of a term, we can define the semantical coverage ratio “SCRBP” of a business process BP_i composed of p_i terms as:

$$SCRBP(BP_i) = \frac{\sum_{j=1}^{p_i} SCRT(t_{ij})}{p_i}$$

The SCRBP of BP_i is in $]0,1]$. The proof is as followed: we have proven in the section above that $0 < SCRT(t_{ij}) \leq 1$. Thus $0 < \sum_{j=1}^{p_i} rcov(t_{ij}) \leq p_i$. Therefore $\mathbf{0} < SCRBP(BP_i) \leq \mathbf{1}$.

The SCRBP is an illustration of the level of representation of the BP at the semantic layer. It allows to measure in order to quantify the semantic gap between the BPs and semantic data.

In the following section we discuss how this metric can be used to plan process correction in a context of OIOP.

4. ORGANIZATIONAL INTEROPERABILITY READINESS AND BUSINESS PROCESS CORRECTION PLANNING

The section above defines an indicator that allows measuring the extent of representation of a BP in the semantic layer.

Many research efforts had been conducted in the area of SIOP. Computers are now able to exchange data with unambiguous shared meaning. Recognized concepts, methods and standards such as ontologies have been set to address the SIOP.

The idea of setting the SCRBP is to measure the readiness of the BP to interoperate according to its level of SIOP. In other words, a BP semantically “good” is a BP that does not need a lot of semantical correction and the interoperability on the semantic level will be a lever for the interoperability at the level of this BP.

4.1 Organizational Interoperability Readiness evaluation scale

As mentioned before, quantifying the semantical gap in a BP can give a good idea on the readiness of the BP to interoperate.

In order to guarantee the OIOP readiness of a BP, the semantic gap has to be at its minimum. The BPs with a low SCRBP has an important semantic gap, thus it needs a high correction level and subsequently it is not ready to interoperate. On the other hand, those with a higher SCRBP do not need a lot of correction effort and thus their OIOP readiness is high and the OIOP on the level of the BP can be enhanced by an existing SIOP.

The BPs can be classified considering their SCRBPs, the correction effort level and the level of their OIOP readiness. The table below describes the proposed classification.

Table 1. Classification of BPs according to their SCRBP

SCRBP value	Level of semantic gap	Correction effort level	Level of OIOP readiness
$]0,0.5]$	Important semantic gap	Important	Not ready
$]0.5,0.8]$	Medium semantic gap	Medium	Partially ready
$]0.8,0.9]$	Minimum semantic gap	Minimum	Nearly ready
$]0.9,1]$	Insignificant semantic gap	Insignificant	Ready

4.2 Business Process Correction Planning

The indicator introduced above represents a mathematical measure of the impact of the semantic layer on the organizational layer.

The indicator can be used to identify the BPs ready to interoperate and those that need a correction.

Determining the correction priority of business process depends on the importance and criticality of the BP in the interoperating network that provides the final public service.

In an e-government context, public agencies and organization have to create value by delivering quality public services. Thus, the BPs are not all on the same level of criticality. Many factors determine the correction priority of BP relatively to another.

To judge if the ratio is good or not, not only the value of the ratio must be quantified but also the importance of the BP in the organization must be considered.

The final delivered service is a product of a public chain value based on the concept introduced by Porter's Chain Value (Porter, 1985). Many research efforts have been done in this area. Heintzman and Marson (Heintzman & Marson, 2003) introduce in figure 3 a Public Sector Service Value Chain Model that is an adaptation of the Private Sector Service Profit Chain Value.

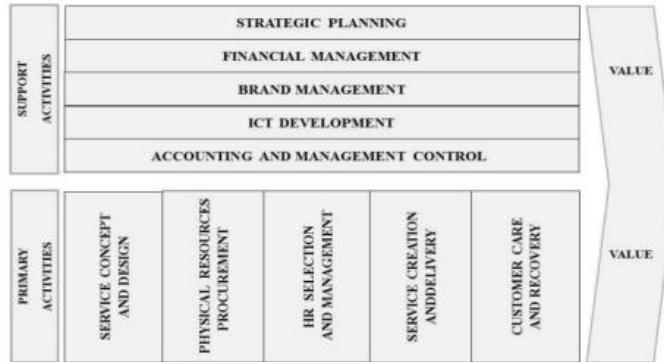


Figure 3. Public Value Chain (Heintzman & Marson, 2003)

Therefore, the priority of correction of BPs with a low SCRBP in the primary activities group and involved in a cross-organizational interaction is higher than those in the support activities one. Thus the correction planning can be based on the priority of a business process.

The table below (Table 2) illustrates how this classification can be addressed:

Table 2. Classification BP correction priorities

SCRBP value	Primary Activities BP	Support Activities BP
[0,0,5]	High correction priority	Low correction priority
[0,5,0,8]	Medium correction priority	Medium correction priority
[0,8,0,9]	Minimum adjustment effort	Minimum correction priority
[0,9,1]	No correction to plan	No correction to plan

In our future work, we propose initially to develop an application that calculates the different SCRBP in the organizational layer. This application will allow highlighting the BPs with a semantic gap. It will also identify the BPs ready to interoperate.

We will attempt also to identify the influencing information within an organization. Determining the influential data is crucial to understand how the organization works and how it can interoperate with another organization.

5. EXPERIMENTAL SETUP

The approach exposed above defines the SCRBP indicator that allows identifying the BPs with a semantic gap.

In the context of e-government cross organizational interactions, the indicator allows to plan the correction of this semantic gap in order to prepare the BP interoperability.

In addition, the influencing semantic data can also be identified. The graph depicted in figure 1 can be used to define the centrality metrics of the semantic nodes.

We have developed the platform in figure 4, which calculates the SCRBP of parametrized BPs. This application establishes the foundations to the influencing data analysis planned.

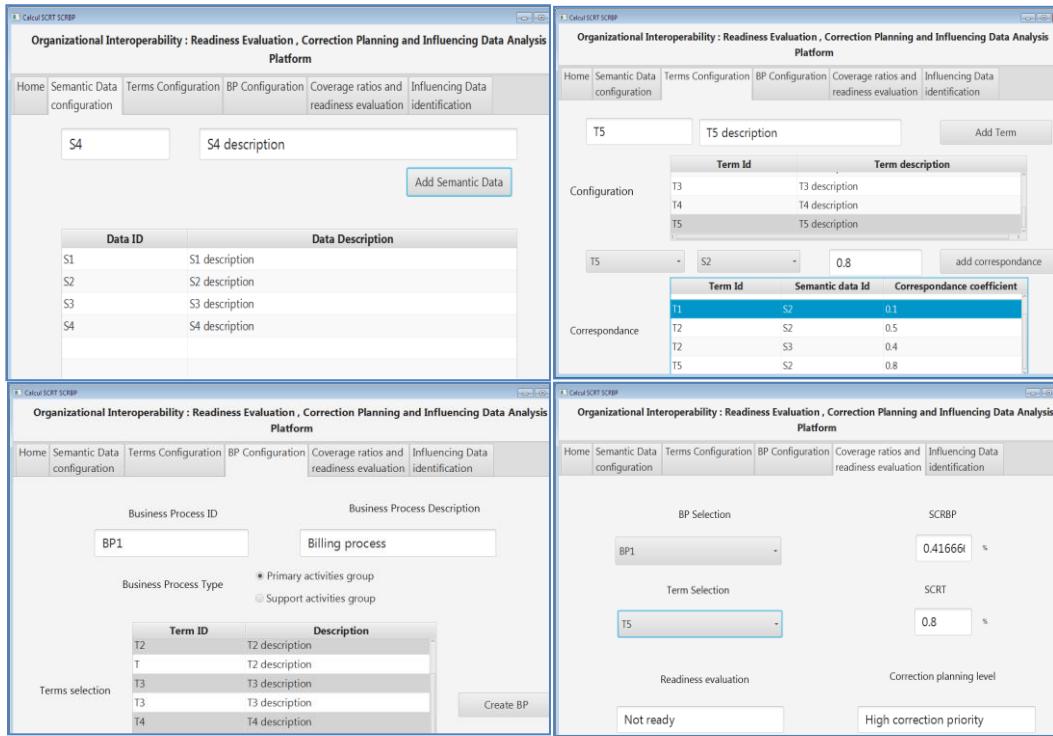


Figure 4. OIOP Evaluation Platform: Readiness Evaluation, Correction Planning and influencing Data Analysis Platform

Highlighting the impact of semantic data in a social network has been addressed in several contexts. Hashtags for example are used to identify trends in social network discussions or to measure the importance of a public subject within a community (Lycarião & Santos, 2017). In Addition, in a business context, including semantics in social network analysis can enhance achieving integration and solving enterprise knowledge management issues (C. Chelmis, 2014).

In our future work, we will attempt to define centrality measures to exhibit the influencing data within a collaborative network. Identifying the influencing data in a context of cross-organizational interaction will allow understanding cross-organizational interaction challenges and can be a negotiating lever. Another goal is to deal with the last aspect (negotiating) by taking advantage of Action Network Theory (ANT) based collaborative mode (Benqatla, et al., 2016).

6. CONCLUSION AND PERSPECTIVES

In e-government context, many public organizations and agencies cooperate to achieve common goals. This cooperation leads necessarily to interconnection between business processes to establish organization interoperability.

After reviewing some definitions of organization interoperability in section 2, we proposed a formalization of the link between the semantic and organizational layers. Before addressing OIOP between organizations, internal factors can already give an idea of the readiness of the BPs to interoperate.

In any cooperative network, technical and semantic interconnections are established. The idea of setting the SCRB indicator is to determine if a BP is ready to interoperate by evaluating its semantic gap.

The formalization of this metric is based on an oriented weighted graph. It allows setting a mathematical formula and gives the foundations of our future work on influencing data analysis. The section 3 describes the formalization proposed.

In section 4 we have discussed how this indicator can be used to evaluate OIOP readiness and in section 5 we describe an experimental setup to automate the calculation.

The main purpose of setting this indicator is to analyze influencing data within an organization and in a collaborative network in general. Assessment of interoperability in general has been a wide concern in the research community. It has been addressed mainly by using control theory (Elmir & Bounabat, 2010).

Future work should detail how this indicator allows identifying influencing data based on Social Network Analysis concepts and graph theory (Barnes & Harary, 1983) (F.M. & Martins F.M.L, 2016). How to identify strategic data? How to measure the “fingerprint” or the “volume” of semantic data in the organizational layer? As there are many centrality metrics, which measure will be the most accurate in identifying the strategic nodes? How identifying strategic data will give negotiating clues and support strategic interactions decisions?

REFERENCES

- Barnes, J. A. & Harary, F., 1983. Graph Theory in Network Analysis. Dans: Social Network. North Holland: pp. 235-244.
- Benqatla, M., D., C. & B., B., 2016. Actor network theory as a collaborative mode: The contribution of game theory in the interressement phase. Marrakech, World Congress on Information and Communication Technologies, WICT 2015, pp. 67-725.
- Chelmis, C., H. W. V. S. V. K. P., 2014. Semantic Social Network Analysis for Entreprise. Computing & Informatics, 33(3), pp. 479-502.
- DODD 2010.6, 1977. Standardization and interoperability of weapon systems and equipment within the North Atlantic Treaty Organization. NATO.
- Elmir, B. & Bounabat, B., 2010. Integrated public e-services interoperability assessment. International Journal of Information Science and Management Issue, Issue SPL.ISSUE2, pp. 1-12.
- EUROPEAN COMMISSION, 2010. Towards interoperability for European. [En ligne] Available at: <http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A52010DC0744> [Accessed le 06 08 2017].
- F.M., C. & Martins F.M.L, M. R., 2016. Social Network Analysis: Concepts and Definitions. Social Network Analysis Applied to Team Sports Analysis, Volume XII, pp. 7-17.
- Heintzman, R. & Marson, B., 2003. People, Service and Trust – is there a Public Sector Service Value Chain?. Oeiras,
- Kubicek, H., Cimander, R. & Scholl, H., 2011. Organizational Interoperability in E-Government - Lessons from 77 European Good Practice Cases.
- Legner, C. & Lebreton, B., 2007. Preface to the Focus Theme Section: "Business Interoperability" Business Interoperability Research: Present Achievements and Upcoming Challenges. Dans: Electronic Markets. pp. 176-186.
- Lycarião, D. & Santos, M. A. d., 2017. Bridging semantic and social network analyses: the case of the hashtag #precisamosfalar sobre aborto (we need to talk about abortion) on Twitter. Information, Communication & Society, 20(3), pp. 368-385.
- Porter, M., 1985. Competitive Advantage: Creating and Sustaining Superior Performance. (Republished with a new introduction, 1998.) éd. NY: Free Press.
- Taoudi, A., Bounabat, B. & Elmir, B., 2013. Quality based project control using interoperability as a quality factor.. Marrakesh; Morocco.
- Taoudi, A., Bounabat, B. & Elmir, B., 2013. Toward a methodology to control interoperability projects execution. Enschede; Netherlands, LNBIP, pp. 186-197.

THE TRANSFORMATIONAL POTENTIAL OF CIVIL SOCIETY

József Veress

Corvinus University, 1093 Budapest Fővámtér 8, Hungary

ABSTRACT

The civil society due to its transformational dynamism is a major driver of the emergence of cooperative, sharing, and genuinely sustainable knowledge society. Historically the civil society is simultaneously the domain and the outcome of the civil activism. This activism interplays with the dynamism of civil society, capitalizes on and catalyses its transformational character. This transformational capacity provides the civil society organizations' potential to facilitate changes in the broader environment, creates their capability to affect and shape the resultant pattern of interplay between digitalization and the approximation among the societal macro-sectors by enhancing their association-prone character.

KEYWORDS

Transformational Dynamism, Civil Activism, Digitalization, Macro-Sectors' Convergence, Knowledge Society

1. INTRODUCTION

Historically the civil society emerged together with the industrial society. Its members' activism played crucial role in practical implementation of the principles arising from the 'glorious triad' of liberty, equality, and solidarity (once coined as fraternity). This activism generated institutional changes which (i) contributed to enhanced social productivity providing the potential to liberate 'increasing volume' of time from wage work and (ii) enforced changes enabling to mobilize growing part of this free time for voluntary activities.

The paper analyses causes and mechanisms of the civil society organizations' dynamism and discusses their transformational outcomes in context of the emergence of knowledge society. It explores how its interplay with the civil activism can affect the enactment (Orlikowski 2000) of the new, primarily digital technologies what in turn has the potential to promote more cooperative dynamism across social fields by shaping patterns of approximation among the macro-sectors.

2. TRANSFORMATIONAL DYNAMISM OF CIVIL SOCIETY ORGANIZATIONS

The study of civil society organizations, their emerging networks and the related changes helps to explore sources and outcomes of their transformational dynamism generating the potential of empowering social agency by 'going after the small picture' (Giddens 1984). The exploration of community clusters representing broad array of civil society organizations provide ample empirical data for combined deployment of the methodological pluralism (Van de Ven and Poole 2005), an extended version of the scientific realism (Bhaskar 1978, Tsoukas 1989), and the ideal-type constructs that Weber (1949) proposes. This setup enables to identify sources and mechanisms of the civil society organizations' dynamism, its transformational impacts on their broader environment by identifying emerging long-term transformational tendencies - currently in phase of their nascence.

2.1 Transformational Dynamism: Sources and Mechanisms

The exploration of the transformational dynamism unfolded through recursive triangulation among the (i) empirical data extracted from five clusters of communities, (ii) the relevant research literature, and (iii) the

emerging constructs which the iterative analytic efforts identified deploying methodological pluralism. It indicated that the interactions among members of civil society organizations taking place in the empirical domain (Bhaskar, 1978; Tsoukas, 1989) create feed back changes unfolding in actual domain (Bhaskar, 1978; Tsoukas, 1989) and impacting multiple dimensions simultaneously. These alterations are mutually catalytic what facilitates also their patterned aggregation into re-emergence of the volunteers' organizations which possess transformational dynamism. This transformative impact affects simultaneously the interacting members, their activities and relationships, as well as their organization and its wider environment (Table 1).

Table 1. Components of the civil society organizations' transformational dynamism

Personal context:	Empowerment
	Individuation
Relationships:	Power relations
	Institutional changes
Activities:	Work
	Competition
	Value creation
	Resourcing
	Social agency
Community alterations	Social capital and trust
self-transformation:	Networking self-upgrading
	New dialectics of cooperation

The volunteers' intertwined intra- and inter-personal dialogues carry out meaning and decision making (Stacey 2010) and simultaneously aggregate into self-communication which "...multiplies and diversifies the entry points in the communication process. This gives rise to unprecedented autonomy for communicative subjects to communicate at large" (Castells, 2009:135) (Figure 1). Such autonomy facilitates and capitalizes on the robust institutional shift to dual primacy of non-zero-sum approach and interdependence therefore the volunteers' self-communication enacts and re-creates (primarily) association-prone institutional settings. This autonomy enables to bring cooperation into competitive environments (Benkler 2011) and to maintain collaboration despite the presence of robust institutional isomorphic pressures (DiMaggio and Powell 1983) that the broader environment 'promotes' generating domination-seeking and colliding instead of cooperative relational dynamism.

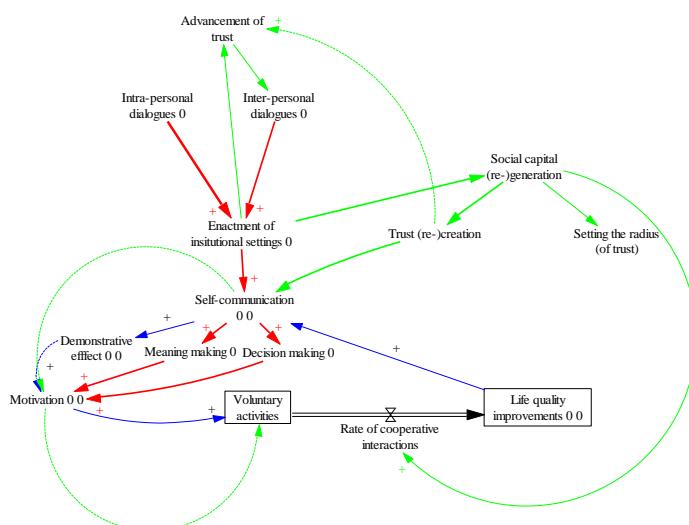


Figure 1. Self-communication re-generating motivation and trust

The volunteers due to such enhanced autonomy can carry out collective pursuits through communicative interactions what enables to socialize or ‘participate for the sake of participation’. The volunteers’ perceive their contributions to common pursuits as passionate and sharing co-creation, as non-wage work which improves their shared life quality - and re-generates their motivation to participate in and contribute to collective efforts. It amends primary motivation to volunteer that the often tacit wish to socialize creates.

The cooperative institutional-relational dynamism in the civil society organizations feeds back with robust transformations in characteristics of power facilitating to replace hierarchies with broader horizontalization tendencies. The volunteers don’t accept attempts of domination and control, their power (relations) possess horizontal, non-zero-sum, non-hierarchical, shared and sharing character. This integrative power or ‘power with’ (Kreisberg 1992) is non-zero sum therefore it could be strengthened or increased through sharing by simultaneously enabling mutual empowerment.

Due to these trends the civil society organizations become commons which serve as shelters against alienation and estrangement tendencies and facilitate their members’ empowering individuation. Their empowerment “[unfolds as a] multi-dimensional social process that helps people gain control over their own lives. It ... fosters power in people, for use in their own lives, their communities, and in their society, by acting on issues that they define as important... To create change we must change individually to enable us to become partners in solving the complex issues facing us. In collaborations based on mutual respect, diverse perspectives, and a developing vision, people work toward creative and realistic solutions. This synthesis of individual and collective change...is our understanding of an empowerment process” - explain Page and Czuba (1999). “...There is an important distinction between... - what could be called selfish individualism - and what is sometimes referred to as individuation...Beck and Giddens...argue. Individuation is the freeing up of people from their traditional roles and deference to hierarchical authority, and their growing capacity to draw on wider pools of information and expertise and actively chose what sort of life they lead. Individuation is...as Beck points out... about the politicization of day-to-day life; the hard choices people face...in crafting personal identities and choosing how to relate to issues such as race, gender, the environment, local culture, and diversity” - describes the individuation Grenier (2006:124-125).

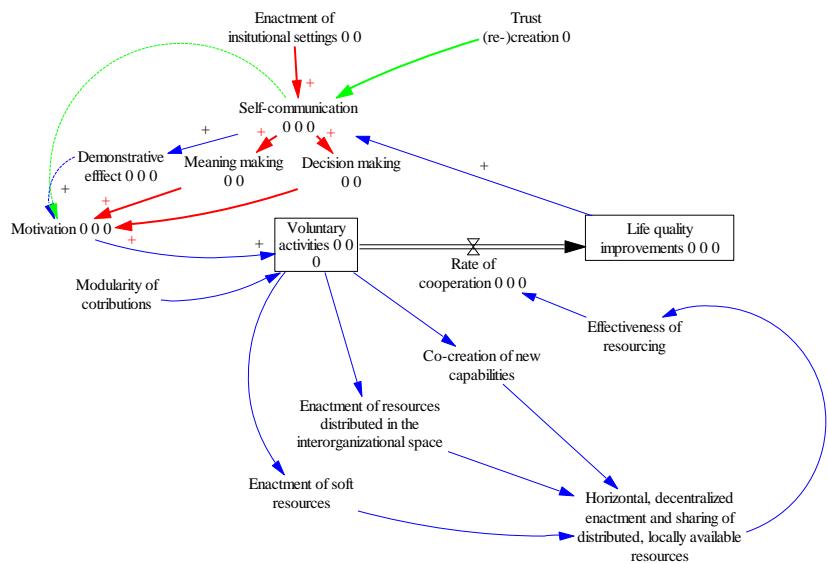


Figure 2. The volunteers’ communicative interactions interplaying with feedback loops

These tendencies are mutually catalytic and their interplay amplifies cooperative atmosphere. The collaboration facilitates to combine the members’ individual capabilities and through symbiotic and synergistic tendencies improves the effectiveness of collective resourcing (Csányi 1989) (Figure 2). The volunteers tend to limit the complexity, size, and resource intensity of particular tasks. Such “modularity of contributions” (Benkler 2011) enables during their parallel, distributed, and mutually adaptive interactions simultaneously mobilize also resources therefore the volunteers’ resourcing follows horizontal, decentralized, distributed and sharing pattern. The autonomy which their self-communication provides enables to follow

diametrically opposite trends in resourcing than the market sector which tends to maximize the resource intensity and complexity of the tasks what allows decreasing the number of the wage workers. By contrast the civil society players aim to increase the number of participants. It facilitates to raise also the overall volume of the contributions since the volunteers' communicative interactions simultaneously carry out the distributed and locally available resources through horizontal and decentralized enactment and sharing. The collaboration facilitates to capitalize on networked patterns of resourcing, which has "...the core assumption...that giving oneself to the larger networked community optimizes the value [also the shared power and resources] of the group as well as its individual members...[similarly to the] Internet"(Rifkin, 2011:268). The volunteers share information, knowledge, and other soft resources similar to creativity and various cognitive, relational, emotional, and psychological energies. Since the soft resources are non-depletable and non-rivalrous (Bollier, 2007:28) these could be shared and also multiplied what helps to extend and upgrade the collective resource base.

The civil society entities are the most important sources of the co-creation and amplification of social capital and trust (Fukuyama 1999) which are indispensable resources also for market and public sector organizations. The social capital is "...an informal norm that promotes cooperation between two or more individuals... [These norms] must be instantiated in an actual human relationship [and generate] trust...[which is] epiphenomenal, arising as a result of social capital but not constituting social capital itself"(Fukuyama 1999:1). The abundance of social capital strengthens cooperative relationships and enables to re-generate and increase trust and extend its radius and the growingly trustful atmosphere amplifies the motivation to cooperate (Figure 1 - above).

Since the abundance of social capital facilitates to extend the radius of trust it can cross over and reach beyond the boundaries of particular organizations. Such networking among members of various groups and entities amplifies their relationships cooperative dynamism and facilitates the emergence of quasi-field(s) where organizational boundaries cease to divide volunteers into various groups. This enables to overcome and prevent the re-emergence of exclusive and fragmented cooperation unfolding only 'within (organizational) boundaries' and which often is oriented against other groups or individuals in the name of group solidarity. The qualitative shift to inclusive and un-fragmented cooperation prevents its self-alienation - a paradox when intra-organizational cooperation generates inter-organizational competition. Consequently, the changes in nature of cooperation interplay with simultaneous self-transformation, the networking self-upgrading of civil society organizations.

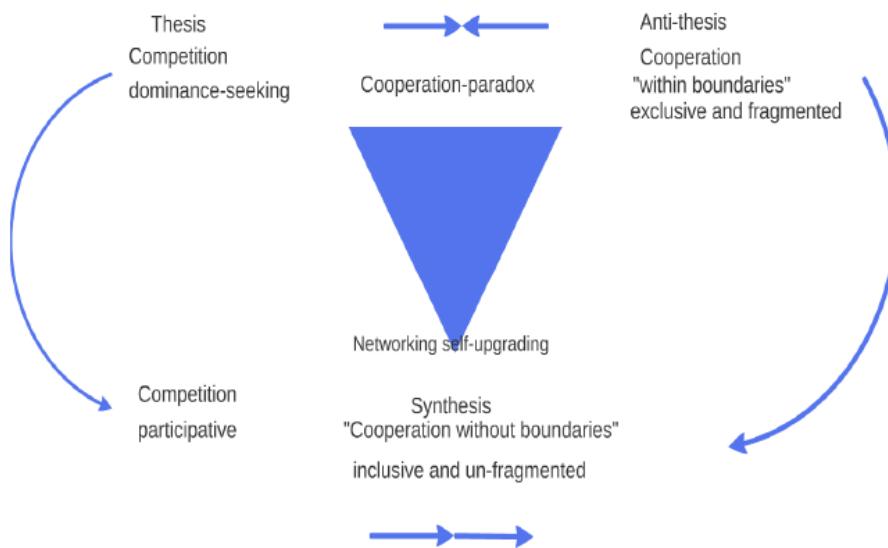


Figure 3. New dialectics of cooperation and competition

When collaboration among individuals belonging to diverse entities turns into inclusive and un-fragmented the competition also becomes participative. Indeed the volunteers compete through improved participation and attempts to provide more effective contributions to common efforts instead of seeking

dominance over each other. Their inclusive and un-fragmented collaboration “without boundaries” appears together with the emergence of altered, participative competition and their interplay follows altered dialectics (Figure 3 – above). Due to the new dialectics between cooperation ‘without boundaries’ and participative competition the quasi-fields of civil society entities catalyse association-prone dynamism across social fields - it provides the capacity of social agency.

2.1.1 Self-organizing Structuration Generating Transformational Dynamism

The volunteers’ intertwined sense and decision making intra- and inter-personal dialogues aggregate into self-communication which enacts association-prone cultural schemas (Sewell 1992), taken for granted perceptions (Perez 1992) and structures (Giddens 1984). “Structures are not the patterned social practices that make up social systems, but the principles that pattern these practices... Structures do not exist concretely in time and space except as “memory traces, the organic basis of knowledgeability” (i.e., only as ideas or schemas lodged in human brains) and as they are “instantiated in action” (i.e., put into practice)”(Sewell 1992:6). Since the volunteers suppose and expect (high probability of) reciprocity of cooperative behaviour, i.e. mutually advance trust they can start (self-) communication which enacts association-prone institutional settings by re-generating trust and facilitating cooperative communicative interactions (Habermas 1995). The association-prone institutional settings that the self-communication enacts regenerate trust and enable the volunteers’ communicative interactions - operate simultaneously as social capital and organizing platforms. Indeed the in depth analysis of the processes taking place in real domain (Bhaskar, 1978; Tsoukas, 1989) indicates that the civil society organizations’ transformational dynamism is dynamic resultant of the interplay among association-prone reconfiguration of structuration and continuously unfolding self-organizing.

The volunteers’ self-communication enacts association-prone institutional settings which catalyse the cooperative character of their communicative interactions and (the pattern of their) aggregation. The cooperative interactions generate change processes in multiple dimensions - including institutional, relational, power, communication and resourcing - that aggregate into continuous patterned (re-) emergence of commons. The association-prone institutional settings operate as soft organizing platforms, facilitate continuous self-organizing that carries out “organizing without organization” (Shirky 2008).

The self-communication simultaneously generates the volunteers’ awareness of improvements in shared life quality which their communicative interactions create. Such demonstrative effect generates motivation, facilitates to repeat the patterns of communicative interactions which generate perceived improvements in effectiveness of resourcing (Figure 2 - above). Such perception serves as selective factor which facilitates to repeat particular (communicative) interactions and to aggregate them into continuous self-organizing. The patterned interplay of self-reinforcing feedback loops can unleash “cooperation trap”(Csányi 1989), generate spiralling up cooperation tendencies. The patterned (re-)emergence of the commons helps to increase the complexity of cooperative pursuits. It serves as important mechanism of growing creativity, capacity to innovate, generate change - creates and amplifies the capacity of (social) agency. Since self-organizing enables to “organize without organization” the growing functional complexity is independent from structural complexity, don’t generate bureaucratic tendencies.

In civil society organizations the continuous self-organizing is intertwined with simultaneous changes in structuration (processes) (Giddens 1984; Sewell 1992; Orlikowski 2000), their association-prone reconfiguration. Such re-shuffling simultaneously capitalizes on and (re-)creates cooperative changes in power, communication and sanctions, i.e. in all three systems of interactions (Figure 4 and 5). The collaboration becomes the “primary structure” in civil society entities. It takes place through the volunteers’ mutual co-inspiration and resource enactment and replaces domination that unfolds through authorization and allocation. This shift is intertwined with switch to reciprocity and sharing operating as facilities of power - by replacing authority and property. These alterations are intertwined with simultaneous transformation of the very power. The volunteers’ cooperation can be mutually empowering and they are unwilling and unready to tolerate attempts to exercise hierarchical control and domination. In civil society organizations the integrative ‘power with’ (Kreisberg 1992) is non-domination and non-hierarchical, associational and lateral which can be shared and sharing.

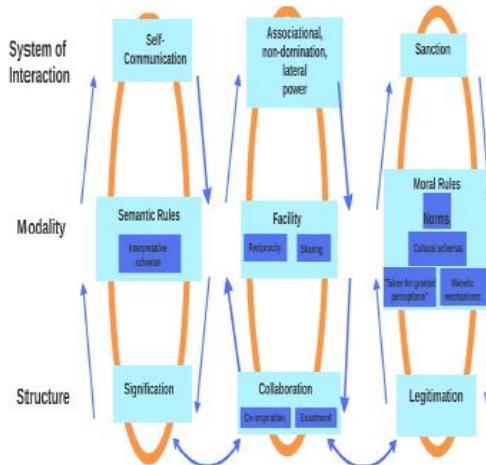


Figure 4. Modified Dimensions of the Modalities of Structuration – based on Stillman (2006: 150)

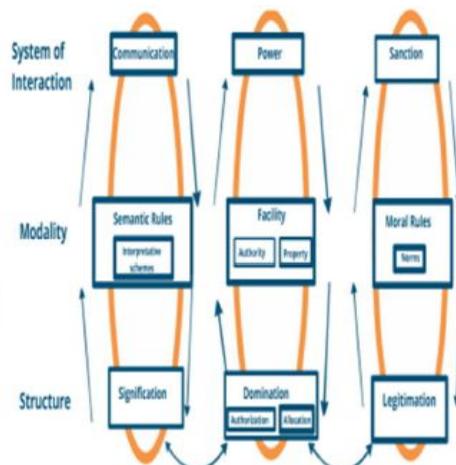


Figure 5. Dimensions of the Modalities of Structuration – Stillman (2006: 150)

The volunteers due to their willingness to socialize and inclination to collaborate generate ‘positive sanctions’ for cooperative behaviour. They enact moral rules, including norms (Giddens 1984), taken for granted perceptions (Perez 2002), and mimetic mechanisms (Scott 1995) possessing robust and growing association-prone character. Their feedbacks play important role in driving and shaping recursive daily interactions that carry out routines and praxis constitutive of everyday life (Perez, 2002). Through deliberate un-learning (Scharmer 2007) the volunteers can accelerate association-prone changes in taken for granted perceptions and in other patterns of moral rules (re-)shaping the recursive daily activities (Perez, 2002).

The volunteers’ self-communication enacts interpretative schemas and semantic rules, follows horizontal, lateral patterns, and is compatible with cooperative, association-prone dynamics of institutional settings. This constellation allows establish and maintain autonomy - the self-communication replaces and prevents the re-emergence of hierarchical, top-down (patterns of) communication. Since the broader environment is characterized by primacy of hierarchies and dominance-seeking competition the endorsement of the lateral approach has eminent importance. Consequently, the association-prone character of signification and legitimization can establish, maintain and amplify cooperative dynamics in local cultures by facilitating to compensate and offset effects - similar to institutional isomorphic pressures (DiMaggio and Powell, 1983) - generated by dominance-seeking competition in the broader environment. The association-prone re-configuration of structuration feeds back with the civil society organizations’ transformational dynamism providing their capability and capacity of self-empowerment and social agency.

2.2 Self-empowerment of the Civil Society: Civil Economy – Patterns of Approximation among Societal Macro-Sectors

The civil society is simultaneously the domain and the ‘product’ of its members’ political activism which drives both its emergence and institutionalization as a societal macro-sector. The civil society is a relatively new historical phenomenon appearing 2-300 years ago and rather tightly connected with (the development of) the industrial society. The (formal) right to carry out self-organizing voluntary activities is connected with and created by the individuals’ legal freedom, which is characteristic (and necessary) for the industrial era. However, the legal equality and liberty brings about transformation into wage-worker in individual context. Although the increasing social productivity in the industrial era potentially enables to liberate more and more time from wage work, only the self-organizing political activism of civil society enacts this potential through enforcing the institutionalization of standards of decreased work time and changing patterns of redistribution.

The civil society organizations due to their transformational dynamism possess a tendency to networking self-upgrading into quasi-fields characterized by new dialectics of inclusive and un-fragmented cooperation. These fields interplay with emergence of large-scale patterns of cooperation similar to Wikipedia and communities of free and open source software (F(L)OSS). These - probably rather rudimentary - large-scale patterns of cooperation are generative and constitutive of transitions to a knowledge driven society taking place through mutual approximation among the three macro-sectors. These are “...a core vector through

which the transition to a networked society and economy ...happening" (Benkler 2011). These can operate as precursor and catalyst of systemic change(s) similarly to merchant capital acting in early days of industrial era. The market and public sectors and the civil society interplay in multiple ways, frequently exhibit diverging or even controversial tendencies. However, these diverging trends, somewhat paradoxically, can also aggregate into convergence bringing about association-prone patterns of mutual approximation among the macro-sectors Figure 6.

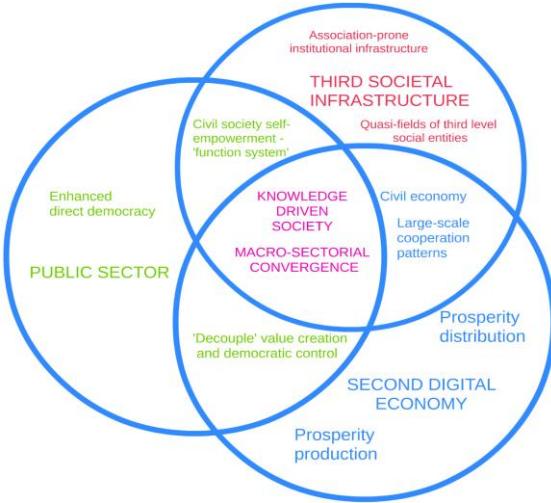


Figure 6. Components of macro-sectorial convergence (...)

The cooperative tendencies are intertwined with increasingly association-prone character of the institutional infrastructure, the "...wider social and cultural context...[the] environments [which] create the infrastructures - regulative, normative, and cognitive - that constrain and support the operation of individual organizations" (Scott, 1995:151). The institutional infrastructure's growingly cooperative character interplays with primacy of non-zero-sum approach and interdependence which facilitates to establish cooperative relational trends also in competitive environments (Benkler 2011).

The strengthening cooperative character of institutional infrastructure and the enhanced self-empowerment of the civil society through emergence of quasi-fields of self-upgrading third level (Vitányi 2007) social entities are mutually catalytic. Their interplay operates as a third societal infrastructure which is catalyst of mutual approximation among the macro-sectors by following cooperative, association-prone tendencies affecting also social fields. (The roads, rail- and, water-roads, pipelines, electric grids, lines of transmission and telegraph lines aggregate into globally inter-linked and expanding networks of transport and communication capacities constituting the first societal infrastructure. This includes 'traditional' forms of telecommunication, facilitates global mobility of goods, service delivery and "human capital". The second societal infrastructure consists of the global network(s) of mostly digital information-communication technologies most often connected through the Internet. It creates and amplifies the individuals' quasi-instant mobile connectivity with (growingly) global reach. The current emergence of the "Internet of Things" (IOT) may physically re-link the first and second societal infrastructure.) The third societal infrastructure facilitates association-prone changes in market and public sectors and amplifies them in the civil society.

The association-prone dynamism of third societal infrastructure is connected with emergence of a civil economy including large-scale patterns of cooperation. This interplay feeds back with the (character of) accelerating digitalization tendencies shaping resultant pattern(s) of emerging second digital economy (Arthur 2011). The stronger is this interplay the bigger is the probability of successful (re-)connection of the (re-)distribution of prosperity with its production. It is the political activism which can strengthen the feedbacks between the association-prone character of third societal infrastructure and the digitalization trends by generating and amplifying the cooperative and sharing character of the convergence among the macro sectors and of their aggregation into a cooperative, sharing and sustainable knowledge society.

3. CONCLUSIONS

The transformational potential of the civil society organizations is the resultant of the interplay between continuous self-organizing and the reconfiguration of structuration providing their robust association-prone dynamism. It feeds back with the volunteering members' political activism that can generate and enact enhanced patterns of direct democracy facilitating the civil society's self-empowerment. These tendencies are also constitutive of its capacity to operate as function system of the society (Reichel 2012) which can ensure "...the provision of stability for joint collective action for something greater than just individual benefits...for the common good and social coherence ...to solve...[also wicked] problems that are not solved by any other part of society"(Reichel, 2012:58-60). Indeed its transformational capacity provides the civil society organizations' potential to facilitate changes in its broader environment: amplify the association-prone character of interplay between digitalization and the societal macro-sectors' convergence by acting as major driver of the emergence of a cooperative, sharing, and genuinely sustainable knowledge society.

REFERENCES

- Arthur, W.B. 2011. 'The second economy'. McKinsey Quarterly, October 2011. https://www.mckinseyquarterly.com/Strategy/Growth/The_second_economy_2853
- Benkler, Y. 2011. The Penguin and the Leviathan The Triumph of Cooperation over Self-Interest. Crown business.
- Bhaskar, R. 1978. A realist theory of science. Hassocks, England: Harvester Press
- Bollier, D. 2007. 'The growth of the Commons paradigm' In: Hess C. and Ostrom E. (eds.) 2007. Understanding knowledge as Commons From theory to practice. MIT Press. pp.27-41.
- Castells, M. 2009. Communication power. Oxford University Press.
- Csányi, V. 1989. Evolutionary systems and Society - A General Theory. Duke University Press.
- DiMaggio, P.J. and Powell W.W. 1983. 'The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields' American Sociological Review 48:147-60.
- Giddens, A. 1984. The Constitution of Society Outline of the Theory of Structuration. Cambridge: Polity Press.
- Grenier, P. 2006. 'Social Entrepreneurship: Agency in a Globalizing World' In: Nicholls, A. (ed.) 2006. Social Entrepreneurship: Agency in a Globalizing World. Oxford University Press.
- Habermas, J. 1995. Moral Consciousness and Communicative Action. Polity Press.
- Kreisberg, S. 1992. Transforming power: Domination, empowerment, and education. Albany NY State University.
- Orlikowski, W. J. 2000. Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. Organization Science. Vol. 11, No. 4 (Jul. - Aug., 2000), pp. 404-428
- Page, N. and Czuba, C. E. 1999. 'Empowerment: What is it?' Journal of Extension. [Online] October 1999 Volume 37 Number 5 <http://www.joe.org/joe/1999october/comm1.php>
- Perez, C. 2002. Technological Revolutions and Financial Capital Dynamics of Bubbles and Golden Ages. Edward Elgar.
- Reichel, A. 2012. 'Civil Society as a System' In: Renn, O., Reichel, A. & Bauer, J. (eds.) Civil Society for Sustainability A Guidebook for Connecting Science and Society. Bremen: Europäischer Hochschulverlag GmbH
- Rifkin, J. 2011. The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World. Palgrave Macmillan.
- Scharmer, O. 2007. Theory U: Leading from the Future as it Emerges – The Social Technology of Presencing. Cambridge MA: SOL Press.
- Scott, W. R. 1995. Institutions and Organizations Foundations for Organizational Science. Sage Publications.
- Sewell, Jr., W. H. 1992. 'A theory of structure: duality, agency, and transformation'. The American Journal of Sociology. 98(1): 1-29.
- Shirky, C. 2008. Here comes Everybody. Penguin Press.
- Stacey, R.D. 2010. Complexity and organizational reality Uncertainty and the need to rethink management after the collapse of investment capitalism. Routledge.
- Tsoukas, H. 1989. The validity of idiographic research explanations. Academy of Management Review 14(4):551-561
- Vitányi, I. 2007. New society – new approach Attempts to critically analyse new characteristics of the human world. (in Hungarian) Budapest: Napvilág kiadó.

MODELLING THE ENROLMENT ESERVICE OF A UNIVERSITY USING MACHINE LEARNING TECHNIQUES

Ainhoa Yera¹, Iñigo Perona¹, Olatz Arbelaitz¹ and Javier Muguerza¹

¹*University of the Basque Country, Faculty of Informatics, Donostia, Spain*

ABSTRACT

This work analyses the navigation in the enrolment eService of the University of the Basque Country. A complete data mining process shows that successful and failure navigation behaviours can be automatically modelled using data mining techniques. On the one hand, unsupervised machine learning techniques have shown that both aspects, where the users navigate and how they do it, affect to the success or failure of a navigation session. In both cases, the sessions in clusters labelled as success or failure, have more than 80% probability of being of that type. Besides, using supervised learning we are able to automatically distinguish the two navigation types with an accuracy rate of 98 % and to identify their main characteristics. Thus, we think that this research is a suitable basis to improve the eService analysed in a near future.

KEYWORDS

eGovernment, eSociety, eServices, Web Usage Mining, Navigation Models

1. INTRODUCTION

eGovernment and eSociety have become a primary trend in the information revolution and almost all the governments in the world have been part of it (Taylor et al., 2007). eGovernment is an important area of the information-technology (IT) innovation. For this reason, governments and institutions overall, are including eGovernment in the government processes (Alshehri and Drew, 2010). The development of the eGovernment is an important task, not only because it is quickly changing the way the institutions provide information and services to the citizen, but also because it is becoming an integral part of their strategies (Zhang et al., 2014).

According to the EU eGovernment Action Plan 2016-2020 (European Commission, 2016) that is currently in force, for 2020, public administrations and public institutions in the European Union should be open, efficient and inclusive, providing borderless, personalized, user-friendly, end-to-end digital public services to all citizens and businesses in the EU. With regard to the inclusiveness and accessibility principle proposed in this Action Plan, public administrations should design digital public services that are inclusive by default and cater for different needs such as those of the elderly and people with disabilities.

To this regard in 2016 the European Parliament and the Council of the EU established a Directive to make the websites and mobile apps of public sector bodies more accessible (European Parliament and Council of the European Union, 2016) based on the European standard EN 301 549 (ETSI, 2014). This standard recommends following the Web Content Accessibility Guidelines 2.0 (WCAG) 2.0 (W3C, 2012) to make the European ICT Products and services accessible for all. However, the standards do not ensure an effective accessibility and more efforts should be done to adapt the web content to the user needs and the usage context.

On this matter, data mining and modelling techniques can be helpful to get user navigation patterns and to improve the development of the eGovernment. In fact, on the one hand, they enable to predict the interactive behaviour and on the other hand, they are useful to evaluate the suitability of the content of a link, the information architecture of a website and the design of a web page (Vigo and Harper, 2013). Web usage mining techniques (Fujimoto et al., 2011) can be used to extract knowledge from observed actions, for example, to get user profiles (Schiaffino and Amandi, 2009).

In addition, introducing the above mentioned methods in the eGovernment context makes two main contributions (Gugliotta et al., 2005): first, they allow to make eGovernment services adaptable to a diversity of users and the second, they improve the integration of these services.

As a step to these contributions this paper presents a research result of the collaboration with the University of the Basque Country (UPV/EHU). Since February 2016 the university has been providing us with access to the navigation data of its whole website. In order to provide clues for future service improvements, our main goal has been to model the university enrolment area (www.ehu.eus/web/sarrera-acceso) as an eService and to extract as much knowledge as possible from it through data mining processes.

Initially, we analysed the structure and content of the whole website of the university, in order to identify the parts related to the enrolment area. Then, we studied the usage of this particular area extracting the navigation sessions of the users from the log files stored in the servers, and labelling them as success or fail based on the end of the navigation. Finally, we used supervised and unsupervised algorithms to answer three meaningful questions: whether the area where a user navigates and how the user navigates (the two types of information used to represent user sessions) affected the success or failure of her/his navigation, if both sources, the navigation area and how the navigation was done, were closely related and if it is possible to foresee if new sessions will be successful or not just analysing the beginning of the navigation.

The paper proceeds with the background of the research in Section 2. The next sections, Section 3 and Section 4, describe the research and the results achieved respectively. Finally in Section 5 we present the main conclusions of the study.

2. BACKGROUND

The UPV/EHU is the public University of the Basque Country with campuses over the three provinces of this region: Biscay, Gipuzkoa and Álava. This institution was established in 1980 and it has around 45,000 students and a staff of around 3,500 workers.

In this research we analysed the usage of the web page of the UPV/EHU, www.ehu.eus and more specifically we were interested in the enrolment e-Service. This university has an online enrolment procedure that can be completed using an IT application called GAUR. However, this process requires to be logged and the institutions have difficulties to provide such data due to privacy issues. Thus, we focused in the navigations of the enrolment area whose main domain is www.ehu.eus/web/sarrera-acceso.

The enrolment area can be accessed using the top menu (*University access* option) displayed in all URLs of the site. This area provides information about the university (staff, contact and location), the access to the university (types of access, academic calendar, admission and enrolment procedure, degree offer...) and scholarships. The main web page of the enrolment area is shown in Figure 1 below.

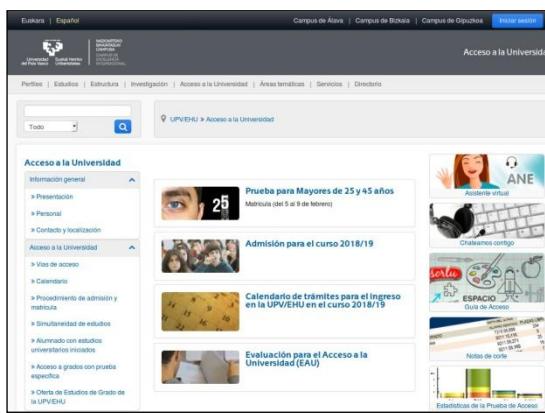


Figure 1. The main web page of the enrolment area of the UPV/EHU www.ehu.eus/web/sarrera-acceso

3. SYSTEM DEVELOPMENT

3.1 Data Acquisition and Preprocessing

This research work analysed the navigation in the area that supplies enrolment information in the website of the UPV/EHU, from the perspective of an eService. To that end, we concretely analysed the navigation data of 49 days, from 23/02/2016 to 12/04/2016.

The first task was to pre-process the log file and select the interesting part. To do so, we only considered the user sessions with at least one URL of the enrolment area (www.ehu.eus/web/sarrera-acceso) and with a minimum length of three clicks, that is, with a meaningful navigation. As a result, a total of 25,467 sessions were selected, around 6 % of the 416,354 sessions available.

In the path to model the enrolment area as an eService the second step was to define a user session classification. In real eServices the users have particular goals or tasks (enrol in a course, apply for a document...) and the user session classification (whether they have finished the task successfully or not...) is evident. However, in our context the users did not have a specific task and thus, we used another kind of criteria to enable the classification of the sessions. Our classification is based on the last web page visited by the user navigation; specifically on the type of the last URL. Hence, we matched the last URLs visited by the users with the achievements of their navigations.

In particular, the type of URLs visited in the sessions were analysed considering two criteria: the content of the URL (whether the text or the links were dominant) and the area of the URL (whether it corresponded to the enrolment area or not). Regarding the content of the URL, those URLs with text format (.pdf / .doc / .docx) were classified as of content type. In the remaining cases, we used the *LCIndex* (Link Content index) (Arbelaitz et al., 2016) described in expression (1) to classify the URLs as of content type or of scatter type (when links were dominant):

$$LCIndex = \frac{Nwords - NwordsLinks}{Nlinks}; URL\ type = \begin{cases} scatter, & LCIndex \leq 10 \\ content, & LCIndex > 10 \end{cases} \quad (1)$$

Where, *Nwords* is the number of words appearing in the web page, *NwordsLinks* is the number of words used in the links of the page and *Nlinks* is the number of links in the web page.

Hypothetically, users visiting the area under focus will be interested in enrolling. Consequently we considered the user sessions finished in a URL with plenty of enrolment information (content type) to be of success type. Conversely, the user sessions ended in web pages with little information (scatter type) were considered to be of failure type. In other words, we classified as success type sessions the ones ended in a content type URL from the enrolment area. In contrast, we considered as failure type sessions those finished in a URL of scatter type and from any area. In this first approach, the sessions ended in a content type URL outside the enrolment area were not considered. In fact, that type of sessions could be linked to success, outside the enrolment context. For example, if the user aimed to get information from another area. Table 1 shows the defined session classification and the number of examples obtained in the database for each type of session. In total, 10,734 success type sessions (42.1 %) and 14,733 failure type sessions (57.9 %) were obtained.

Table 1. User session classification of the enrolment area of the UPV/EHU

Type of session	No. Sessions	Last URL of the session	
		Area	Type
Success	10,734 (42.1 %)	Enrolment	Content
Failure	14,733 (57.9 %)	Enrolment Not Enrolment	Scatter Scatter

To carry out the pre-processing of the log files including the URL filtering, the sessioning and the users' sessions classification, a java project was implemented in the Eclipse platform V3.8.1¹ and finally executed as a jar file in the terminal of Ubuntu 14.04 LTS.

¹ <http://archive.eclipse.org/eclipse/downloads/drops/R-3.8.1-201209141540/>

3.2 Data Preparation

In the pre-processing carried out two files were obtained: one with the sequence of URLs visited in each user navigation session and another one with several navigation attributes for each URL visited by the users extracted from the log files and the sessioning process (Request order, IP, user ID, session ID, URL ID, elapsed time, method, request size, reference etc.). Then, using a script implemented in bash², a Unix shell and command language, an extra field was added to each user session defined in the first file and each URL defined in the second file, in order to specify the type of sessions and URLs visited (success / failure).

Using both files we created two databases (DBs) with the selected sessions: a DB revealed where the users navigated whereas the other one showed how they did that navigation. In the first one, the user sessions were represented with the sequences of URLs accessed by the users. In the second one, based on the information of the log files, a vector of mechanical attributes was created to represent the user sessions, using several scripts implemented in bash. The attributes that represented the user sessions in the second DB were computed according to the time, the URL classification (content / area) or the number of clicks. Table 2 shows the list of attributes mentioned and their description.

Table 2. Attributes used to represent the user sessions

Attribute	Description
No. click	Number of clicks (length of the session).
No. scat. %	Number of scatter type URLs / length of the session
No. cont. %	Number of content type URLs / length of the session
No. enr. %	Number URLs of the enrolment area / length of the session
No. not-enr. %	Number URLs from outside the enrolment area / length of the session
No. ind. %	Number of times the start page of the enrolment area (index) is visited / length of the session
No. ref-sear. %	Number of URLs that have web search engine as reference
T-ses	Duration of the session (s)
T-click_avg	Average duration of a click (s)
T-scat._avg	Average duration of a click on a scatter type URL (s)
T-cont._avg	Average duration of a click on a content type URL (s)
T-enr._avg	Average duration of a click on a URL of the enrolment area (s)
T-not-enr._avg	Average duration of a click on a URL outside the enrolment area (s)
T-ind._avg	Average duration of a click on the start page of the enrolment area (s)
No. cont.-scat.	Number of transitions content type URL- scatter type URL
No. scat.-cont.	Number of transitions scatter type URL- content type URL
No. enr.-not-enr.	Number of transitions enrolment area URL- outside enrolment area URL
No. not-enr.-enr.	Number of transitions outside enrolment area URL- enrolment area URL

To better understand the databases, next we show a particular user session from the database with the sequences of URLs visited and the values (not normalized) of the attributes computed in the second DB for the same session according to the information of the log files:

- Session 1 - URLs visited:
 - 1) <http://www.ehu.eus/es/web/medikuntza-odontologia/medikuntza-14-1>
 - 2) <http://www.ehu.eus/es/web/medikuntza-odontologia/gasteiz>
 - 3) http://www.ehu.eus/documents/1546271/2600354/horario6_vitoria_castellano_2014-2015.pdf
 - 4) <http://www.ehu.eus/es/web/medikuntza-odontologia/medikuntza-plana>
 - 5) <http://www.ehu.eus/es/web/medikuntza-odontologia/tramiteak>
 - 6) <http://www.ehu.eus/es/web/medikuntza-odontologia>
 - 7) <http://www.ehu.eus/eu/web/sarrera-acceso/gutxieneko-notak>
- Session 1 - attributes computed:

No. click = 7, No. scat. = 43 %, No. cont. = 57 %, No. enr. = 14 %, No. not-enr. = 86 %, No. ind. = 0, No. ref-sear. = 29 %, T-ses. = 98 s, T-click_avg = 14 s, T-scat._avg = 14 s, T-cont._avg = 14 s, T-enr._avg = 21 s, T-not-enr._avg = 13 s, T-ind._avg = 0, No. cont.-scat. = 2, No. scat.-cont. = 2, No. enr.-not-enr. = 0, No. not-enr.-enr. = 1.

² <http://www.gnu.org/software/bash/>

4. MODELLING OF THE ENROLMENT ESERVICE

The goal of our research was not only to automatically detect each type of session, that is, to build a two-class classifier that distinguished between success and the failure type sessions. Our main challenge was to improve this eService and thus, we analysed four main issues that could provide us useful information for that task: the usefulness of the navigation style and area to discriminate between the sessions labelled as success and failure; the closeness between the two perspectives analysed and the possibilities to foresee the type of new session just analysing the beginning of the navigation. We next explain in detail the answers found to the questions drawn up.

4.1 Session Discrimination based on Navigation Area

In order to analyse the effectiveness of the navigation area to discriminate the type of sessions defined, we used PAM (K-medoids) algorithm (Kaufman and Rousseeuw, 1990) in the DB built with sequences of URLs. In fact, PAM is a clustering algorithm that allows to group sequences into clusters. Broadly, in a clustering procedure the number of clusters selected is ideally small, as it contributes to create clusters with as many cases as possible of the same type. In our case, as the different K tested (50, 75, 100 ...) hardly influenced the structures of the clusters, we selected the smallest value tested (K = 50) for the algorithm used.

To evaluate the discernment power of the approach for the two types of navigation sessions, we focused on the clusters where the superiority of success or failure cases was over 74 %. This percentage was selected on the one hand because the total number of sessions grouped in those types of clusters was significant (the 42 % of the whole DB), and on the other hand because it provided a suitable representation of each type of session (12 % of success and 29 % of failure). Considering the results, these are the most significant findings:

- We obtained eight clusters where the success class is dominant, being the proportion of success type sessions higher than 74 %.
- Using the 12 % of the sessions of the DB (the eight clusters linked with success) we were able to group the 24 % of the success type sessions.
- We obtained 17 clusters where the failure is dominant, with a proportion of failure type sessions higher than 74%.
- Using the 29 % of the whole sessions (the 17 clusters linked to the failure) we were able to detect the 45 % of the failure type sessions.

Table 3 summarizes the types of examples grouped in the eight success type clusters and the 17 clusters of failure type.

Table 3. Results of PAM (K = 50) in the DB built with sequences of URLs

Attribute	Clusters with a no. Success-sessions \geq 74 %	Clusters with a no. Failure-sessions \geq 74 %
No. clusters	8	17
No. success-sessions	2,551 (81.9 %)	842 (11.2 %)
No. failure-sessions	564 (18.1 %)	6,669 (88.8 %)
No. sessions-clusters	3,115	7,511
No. sessions-DB (%)	12.2 %	29.5 %

From Table 3 it can be drawn that half of the clusters (25) have a proportion of one class or the other one that exceeds a 74 %. In the whole DB, we have a total of 14,733 user sessions of failure type and in the clusters named like that, there are 6,669 sessions. Thus, although the number of failure type sessions is 58 %, the ones that are grouped in the failure named clusters have more probabilities to be of failure class, 89 %. Similarly, being the success type user sessions 42 % of the whole DB, in the selected clusters this proportion increases up to 82 %. In sum, it seems that there is a connection between the areas visited during the navigation (URLs) and the success / failure classification, structure that the clustering is able to get automatically. Hence, we can gather that there is a chance to automatically classify the navigation of new users of the UPV/EHU enrolment eService, based on the areas they use (URLs).

4.2 Session Discrimination based on Navigation Style

By contrast, in order to analyse the effectiveness of the second approach to discriminate the type of sessions defined, we ran the K-means algorithm (Lloyd, 1982) using $K = 50$ in the database with the mechanical attributes of the sessions (Table 2) previously normalized (normal distribution). Then, we selected the clusters with a superiority of success or failure cases over 74 %, which grouped 43 % of the total number of sessions of the DB (14 % of the success and 29 % of failure). Regarding the results these are the relevant issues:

- We obtained six clusters with a proportion of success type sessions higher than a 74 %, gathering the 14 % of the total number of sessions defined in the DB.
- Using the 14 % of the sessions we were able to group the 30 % of the success type sessions.
- 21 clusters were obtained where the failure type sessions were over 74 %.
- Using the 29 % of the existing sessions we were able to detect a 46 % of the failure type sessions.

These results are summarized in Table 4 below.

Table 4. Results of K-means ($K = 50$) in the mechanical attributes DB

Attribute	Clusters with a no. Success-sessions $\geq 74\%$	Clusters with a no. Failure-sessions $\geq 74\%$
No. clusters	6	21
No. success-sessions	3,212 (87.7 %)	544 (7.4 %)
No. failure-sessions	451 (12.3 %)	6,842 (92.6 %)
No. sessions-clusters	3,663	7,386
No. sessions-DB (%)	14.4 %	29 %

As it was the case for the previous DB, analysing Table 4 we see that in more than half of the clusters, the proportion of one of the types of sessions defined or the other is higher than 74 %. In this case in the failure named clusters there are 6,842 sessions of failure type. As happened before, although failure type sessions represent 58 % of the DB, the sessions within the mentioned clusters have higher probabilities to be of failure type, more exactly 93 %. Likewise, being the success type sessions 42 % of the DB, this number greatly increases in the clusters named as success, up to 88 % precisely. Thus, the results show that the navigation style is discriminant for the two types of navigations defined, success and failure.

4.3 Comparison of the Navigation Area and Navigation Style

Analysing the results, both points of view, the navigation area and the navigation style, seem to have a similar ability to discern between success and failure type sessions. Accordingly, in order to know if the two perspectives were closely related or not, we then compared the partitions of the two PAM clustering procedures using the Jaccard index (Jaccard, 1908).

The Jaccard index provided a value of 0.04 in this comparison, a very low value showing that both results are quite different. In light of that, we can claim that in the navigation of the UPV/EHU enrolment area, the area (URL) and the navigation style described by the mechanical attributes are independent, what suggests that the design of each concrete URL does not affect much to how the user navigates. Hence, regarding the task of classifying the navigation of the new users of this eService, the two features mentioned (area / mechanical attributes) would be useful and it remains to be analysed whether both systems are complementary or not.

4.4 Automatic Classifier System based on the Navigation Style

Finally, considering the future goal of building a system able to classify new user sessions of the enrolment eService analysing just the beginning of the navigation, we carried out two approaches to automatically classify the sessions using 10 C4.5 (Quinlan, 2014) and 10 CTC (Consolidated Tree Construction) (Ibarguren et al., 2015) trees using the DB with the mechanical attributes. In fact, these two supervised learning approaches will provide us not only the specific discrimination capacity of the system to classify new sessions as success or failure, but also a concrete description of the mechanical attributes to be used in the process.

For this task, the sessions of the dataset were chronologically ordered and the first 25,000 were selected for experiments. Then, as shown in Figure 2 below the new DB was divided in ten parts of 2,500 sessions ($Fold_i \mid 1 \leq i \leq 10$) and each part was again divided into 10 segments of 250 sessions ($F_{ij} \mid i, j \in \mathbb{N}, i, j \leq 10$). Every split respects the chronological order as it would happen in exploitation in a real system. As in an ordinary 10 fold-cv procedure, the first nine parts of this DB ($Fold_i \mid 1 \leq i \leq 9$) were used for training whereas the last one with newest sessions ($Fold_{10}$) was kept for test. To build each of the ten trees a particular segment from the ten parts of the DB available was used (first = F_{i1} , second = F_{i2} , third = F_{i3} etc.), but always using these concrete segments from the nine first parts ($F_{ij} \mid 1 \leq i \leq 9, 1 \leq j \leq 10$) as training (2,250 sessions) and the newest segment ($F_{10j} \mid 1 \leq j \leq 10$) for test (250 sessions). This way we ensured that the data were equally time distributed among the trees so that they had similar learning processes and that the newest sessions ($Fold_{10}$) were used for test.

The algorithms were implemented in Visual C++ although they are also available as official WEKA packages (C4.5 and J48Consolidated³).

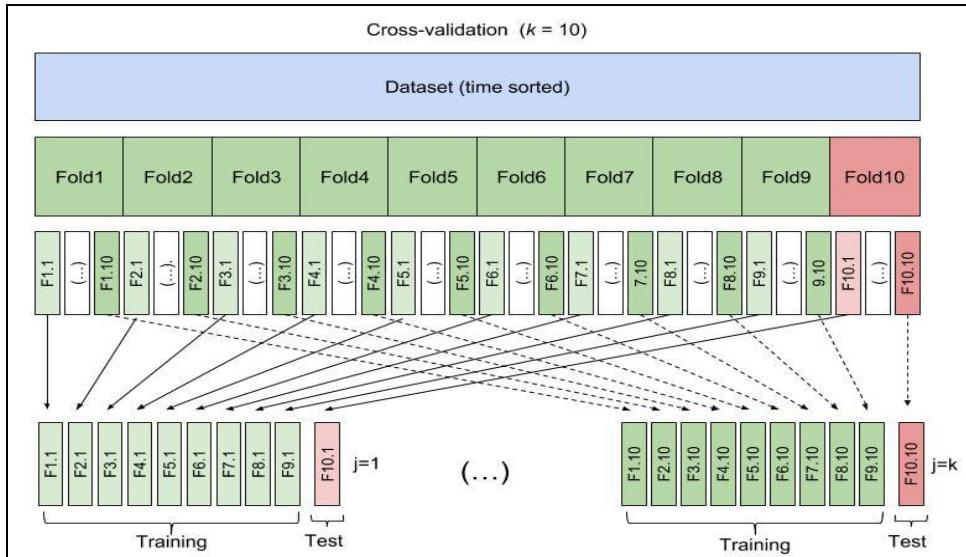


Figure 2. Procedure to build 10 C4.5 and 10 CTC trees using the DB with the mechanical attributes

Ideally, the best classifier will be the one with better classification performance, that is, the one with a low classification error and a simple and stable explanation. Regarding the average error the C.45 decision trees achieved a lower value (0.0500) than the CTC ones (0.0626). Conversely, the CTC trees achieved a higher average value (0.9828) for the Area Under the ROC (AUC) than the one achieved by the C.45 (0.9665) ones. These low classification errors and high AUC values show that both types of trees are able to discriminate between the success and failure navigations defined, however, in order to better compare them a paired t-test was carried out using both metrics. The t-test revealed on the one hand that the AUC of the CTC approach was significantly better than the AUC of the C4.5 approach (p-value > 0.05) and on the other hand, that there were not statistically significant differences between the classification errors of the two options (p-value < 0.05). Concerning the explaining capacity of the trees, we noticed that the structures of the CTC trees were simpler, with values on average for the number leaves and for the number of internal nodes, 14.4 and 13.4 respectively, lower than the ones achieved by the C.45 trees, 44.3 and 43.3. Accordingly, the explanation provided by the CTC approach was proof to be more stable, achieving an average value of common nodes among the different trees (21.81 %) higher than the one obtained in the C.45 approach (6.02 %).

Therefore, the CTC approach was found to be the best option as it achieved a significantly better AUC value and it provided more stable and simple explanations. According to the structure of the CTC trees these are the main mechanical attributes to differentiate each type of session (success / failure): the average duration of a click on a URL of content or scatter type (T-cont._avg and T-scat._avg), the proportion of

³ <http://www.sc.ehu.es/aldapa/weka-ctc/>

content type URLs in the session (No. cont. %) and the number of content type URL– scatter type URL transitions (No. cont.-scat). More specifically, the CTC highlighted the following rules to discriminate each type of navigation:

Main rules to detect failure type sessions:

- ($T\text{-cont._avg} \leq 13.95$ s) AND ($T\text{-scat._avg} > 14.06$ s)
- ($T\text{-cont._avg} \leq 13.95$ s) AND ($T\text{-scat._avg} \leq 14.06$ s) AND ($\text{No. cont. \%} \leq 15\%$)

Main rules to detect success type sessions:

- ($T\text{-cont._avg} > 13.95$ s)
- ($T\text{-cont._avg} \leq 13.95$ s) AND ($T\text{-scat._avg} \leq 14.06$ s) AND ($\text{No. cont. \%} > 15\%$) AND ($\text{No. cont.-scat.} \leq 0.01$)

Therefore, failure type sessions are closely linked with short times in content type URLs ($T\text{-cont._avg} \leq 13.95$ s) and long times in scatter type URLs ($T\text{-scat._avg} > 14.06$ s). This could represent the navigation of those users who are not able to find certain information. In addition, the sessions with small values for the two attributes mentioned $T\text{-cont._avg}$ / $T\text{-scat._avg}$, and low proportion of content type URLs ($\text{No. cont. \%} \leq 15\%$) have more probabilities to be of failure type. Conversely, success type sessions have a close relation with longer times on average in content type URLs ($T\text{-cont._avg} > 13.95$ s). To a lesser extent, short times in content and link type URLs, high proportions of content type URLs and an absence of content-scatter type URL transitions ($\text{No. cont.-scat.} \leq 0.01$) lead more easily to success type sessions.

Finally, it should be remarked that the rules mentioned above, are very similar to the ones provided by the C.45 trees. This fact reinforces the validity of the mechanical attributes noted to discriminate the types of sessions defined. Thus, we think that the conclusions achieved are will be very effective to classify the new user sessions and improve the enrolment eService of the UPV/EHU in the future.

5. CONCLUSIONS

In this research we modelled the enrolment eService of the UPV/EHU using web mining techniques. First of all, based on the end of the navigations of this eService (the area and the kind of content of the last click) we defined two types of sessions: success and failure. In order to obtain as much useful information as possible, we represented the navigation sessions in two different ways, based on the navigation area (sequences of URLs) and based on the navigation style (mechanical attributes).

The application of PAM and K-means clustering algorithms in the databases built with sequences of URLs and with the mechanical attributes respectively came out with partitions where half of the clusters had a high proportion ($> 74\%$) of one of the navigation types defined, showing that the two perspectives used give rise to automatically detect the two navigation types defined.

The two partitions obtained before were compared using the Jaccard index to compare the two partitions. The low value provided by this index (0.04) suggested that the navigation area and the navigation style were not connected.

Finally, two supervised algorithms, C.45 and CTC, where used to build decision trees with the mechanical attributes DB. Although both types of trees performed well, according to the paired t-test carried out the CTC approach provided a significantly better AUC value ($0.9828 > 0.9665$), as well as more simple and stable explanation. Analysing the structure of the CTC trees, we discovered which attributes have more influence when trying to discriminate the navigation behaviours. Specifically, failure type sessions were more related with small times in content type URLs and high times in scatter type URLs. In contrast, success type URLs were closely linked to small times in scatter type URLs.

Thanks to this work, we have taken the first step to model the enrolment eService of the UPV/EHU and considering the results, we can state that web mining techniques are useful in that process.

As future work, firstly we should analyse if the two approaches used to represent the DB are complementary or not. If so, we should improve the accuracy of the systems. We also want to do an in-depth analysis of the models created in order to anticipate to the future users, and to identify and improve those elements having a negative influence on the usability.

To conclude, we would like to use the conclusions of this eService analysis as a kind of guidelines to improve other kinds of eServices and in general, to disseminate all the knowledge that can be useful to improve eGovernment.

ACKNOWLEDGEMENT

This work has been funded by the following units: Firstly, by the University of the Basque Country UPV/EHU (PIF15/143 grant). Secondly, by the research group ADIAN that is supported by the Department of Education, Universities and Research of the Basque Government, (grant IT980-16). Finally, by the Ministry of Economy and Competitiveness of the Spanish Government, co-founded by the ERDF (eGovernAbility, TIN2014-52665-C2-1-R).

REFERENCES

- Alshehri, M. and Drew, S., 2010. Challenges of eGovernment Services Adoption in Saudi Arabia from an e-Ready Citizen Perspective. *Education*, 29 (5.1).
- Arbelaitz, O., Lojo, A., Muguerza, J. and Perona, I., 2016. Web mining for navigation problem detection and diagnosis in Discapnet: A website aimed at disabled people. *Journal of the Association for Information Science and Technology*, 67(8), pp. 1916-1927.
- ETSI, 2014. *EN 301 549 V1.1.1 (2014-02). Accessibility requirements suitable for public procurement of ICT products and services in Europe*.
- European Commission, 2016. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: EU eGovernment Action Plan 2016-2020 Accelerating the digital transformation of government (COM (2016) 179 final)*.
- European Parliament and Council of the European Union, 2016. Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector. *Official Journal of the European Union*, Brussels, 1-103.
- Frank, E., Hall, M. A., & Witten, I. H. (2016). The WEKA workbench. Data mining: Practical machine learning tools and techniques. Burlington: Morgan Kaufmann.
- Fujimoto, H., Etoh, M., Kinno, A. and Akinaga., Y., 2011. Web user profiling on proxy logs and its evaluation in personalization, *In Procs. of the 13th Asia-Pacific web conference on Web technologies and applications (APWeb'11)*, Springer-Verlag, Berlin, pp. 107-118.
- Gugliotta, A., Cabral, L., Domingue, J., Roberto, V., Rowlatt, M., Council, E. C. and Davies, R.A., 2005. Semantic web service-based architecture for the interoperability of e-Government services. *WISM 2005*, 21.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. H., 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1), pp. 10-18.
- Ibarguren, I., Pérez, J.M., Muguerza, J., Gurrutxaga, I. and Arbelaitz, O., 2015. Coverage-based resampling: Building robust consolidated decision trees. *Knowledge-Based Systems*, 79, pp. 51-67.
- Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise de Sciences Naturelles*, 44, pp. 223-370.
- Kaufman, L. and Rousseeuw, P.J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley-Interscience.
- Lloyd, S., 1982. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28 (2), pp. 129-137.
- Quinlan, J.R., 2014. *C4. 5: programs for machine learning*. Elsevier.
- Schiaffino, S. and Amaldi, A., 2009. Intelligent user profiling. In M. Brammer (Ed), LNAI 5640 chapter Intelligent user profiling, *Artificial Intelligence: An International Perspective*, Springer-Verlag, pp. 193-216.
- Taylor, J., Miriam, L. and Organ, J., 2007. Information-intensive government and the layering and sorting of citizenship. *Public Money and Management*, 27(2), pp. 161-164.
- Vigo, M. and Harper, S., 2013. Challenging information foraging theory: screen reader users are not always driven by information scent. *In Procs. of the 24th ACM Conference on Hypertext and Social Media*, ACM, NY, USA, pp. 60-68.
- W3C, 2008. Web Content Accessibility Guidelines (WCAG) 2.0, <http://www.w3.org/WAI/intro/wcag.php>.
- Zhang., H, Xiaolin, X. and Jianying, X., 2014. Diffusion of eGovernment: A literature review and directions for future directions. *Government Information Quarterly*, 31(4), pp. 631-636.

MEDIA INNOVATION AND BUSINESS MODELS: THE CASE OF END-TO-END IMMERSIVE AUDIOVISUAL SERVICES

Heritiana Ranaivoson and Simon Delaere
Imec-SMIT, Vrije Universiteit Brussel, Brussels, Belgium

ABSTRACT

This paper analyses how innovation and business models are related in the case of immersive audiovisual services, with a focus on three end-to-end services: Jaunt, NextVR and Immersive Media. To do so, it first briefly discusses the issues at stake in terms of media innovation, in particular the fact that content innovation is often neglected or misunderstood. It then applies the Business Model Matrix (Ballon, 2009) to these three cases. In particular, the cases are compared in terms of vertical integration, interoperability, revenue model, positioning, user involvement, and the role of content innovation. The paper concludes with a comparison of the cases. It expands the rapidly growing literature on media innovation by using a business model methodology, providing the first instance of an in-depth, structured comparison of immersive audiovisual services from a business point of view. It aims to fill a knowledge gap by dealing with a sector almost ignored by business or economic academic literature.

KEYWORDS

Immersive Experiences, Media, 360 Video, Omnidirectional, Virtual Reality, Interactivity, Business Models, Media Innovation

1. INTRODUCTION

Immersive audiovisual services – characterised by features such as omnidirectional video, interactivity and multi-screen output – have been around for quite some time. However, recent technological advances in various parts of the production chain (from capture, stitching and coding to low-latency transmission and, of course, the rise of Head Mounted Displays for output) have been at the basis of a market surge which, beyond the hype cycle, could prove to be a disruptive factor in the way audiovisual content is produced, distributed and consumed. As with many of these evolutions, new players try to enter the market while established majors choose between focusing on the traditional business model and product line, fully engaging with the new technologies or choosing a hybrid model by entering into strategic partnerships.

Virtual Reality (VR) now has the potential to become one of the next big computing platforms, as we saw with the PC and smartphone (Goldman Sachs, 2016). This is visible in investments made (Digi-Capital, 2017). It is expected that the market will grow rapidly in the coming years (Goldman Sachs, 2016; Digi-Capital, 2016), in particular for devices. While gaming is considered the leading application of VR (Gopal, 2016) and the early driver of uptake of high-end VR headsets (Clarke, 2016), video will also drive adoption of this technology (CC Insight, n.d.). Immersive AV content notably has a potentially positive impact on viewers' engagement. (Chambel, Chhaganlal, & Neng, 2011) However, such impact may also be due to the character of novelty, and therefore may not last when VR has become mainstream.

Since immersive audiovisual experiences – at least in the way they are delivered nowadays, are a relatively new phenomenon, stakeholders involved are still defining the right business model for these services. They innovate in how they produce and deliver content, and how they use it to capture value. Therefore, the main research questions in this paper are: 1) how do companies innovate in VR? 2) what are the design choices they made in articulating a business model for their innovative services? 3) can patterns be discerned in these design choices, thereby characterising a first generation of end-to-end immersive services?

The paper is structured as follows: first, the research questions above are grounded in existing insights and theory concerning media innovation on the one hand, and business models on the other hand. This results in a practical framework for analyzing and comparing different immersive services. Subsequently, the new immersive audiovisual market is delineated and a generic value network is outlined, providing insight in what “end-to-end” service provisioning signifies in this context. Then, the three cases studies are presented. The paper concludes with a discussion of the case studies and some avenues for future work.

2. MEDIA INNOVATION AND CHALLENGES FOR IMMERSIVE AV

Media innovation shares most of the main features of innovation, while also having a few specificities (Lindmark, Ranaivoson, Donders, & Ballon, 2013). Innovation concerns the introduction of something new with an element of valorisation (or utilisation) within it (OECD & Eurostat, 2005; Schumpeter, 1942). In other words, innovation involves putting an invention into practical use. For media innovation as for innovation in general, the discussion on the level of novelty (Freeman & Soete, 1997) and the opposition between product and process innovations (Cave & Frinking, 2007) are recurring issues.

Specifically with regard to media however, Lindmark et al. (2013) argue that the articulation between product and process innovation should be discussed in more detail, notably because content innovation is at the crossroads between the two. Content innovation is usually not well understood by academics, policy-makers and professionals (Bleyen, Lindmark, Ranaivoson, & Ballon, 2014). In contrast, this paper argues, following Ranaivoson et al. (2017a) that content innovation is crucial to ensure the adoption of 360 video.¹ The importance of content innovation has started being taken into account by other market observers, e.g. Goldman Sachs (2016). Immersive AV content requires new storytelling with different writing and producing techniques (Clarke, 2016), e.g. with more long takes, since cutting may make give nausea to users. One can draw a comparison with 3D technology. For VR even more than for 3D, it is not possible to simple port a film over. VR also faces some of the same challenges that hindered 3-D, namely viewers need to buy special equipment to view the broadcast, making the cost and convenience of such equipment an important additional factor (Ranaivoson et al., 2017).

Another challenge, of economic nature, relates to the lack of a sustainable model for the creation of content and a clear pathway to monetization (Clarke, 2016). Content distributors and device manufacturers are aware of such challenge, and develop different strategies, including producing content themselves and attracting content producers. For example, as of Sept 2015, 200,000 developers had registered to create games on Oculus’ VR platform. Oculus expected 100 games will be available in 2016 (including 20 developed internally) (Goldman Sachs, 2016). CCS Insight believes that although augmented reality and virtual reality are two very different technologies, they each have the potential to deliver transformative experiences. (CC Insight, n.d.).

3. BUSINESS MODELLING

In order to empirically analyse these strategies, we propose a business model framework allowing to analyse what design choices 360 video companies have made in order to create and capture value, with a particular focus on the role of content innovation in their overall strategy.

¹ Ranaivoson et al. (2017) puts the availability of means of display, in particular diverse head-mounted displays, at the same level of importance. They are out of scope for this paper due to its limited size.

Table 1. Business Model Matrix

CONTROL PARAMETERS		VALUE PARAMETERS	
Value Network Parameters	Functional Architecture Parameters	Financial Model Parameters	Value Proposition Parameters
Combination of Assets	Modularity	Cost (Sharing) Model	Positioning
Vertical Integration	Distribution of Intelligence	Revenue Model	User Involvement
Customer Ownership	Interoperability	Revenue Sharing Model	Intended Value

The origins of the term business model are found in the writings of Peter Drucker (2006). Drucker asked about the customer, what is valued and how to earn money. Although present for over half a century, the concept has only gained prominence in the last decades, especially during the dot-com revolution of the 2000s (Sawy & Pereira, 2012). As Ballon (2009) explains, with the progression of digitization, the concept gradually shifted from mainly dealing with the ‘logic of creating and capturing value’ (Keen & Qureshi, 2006; Magretta, 2002) towards ‘the development of an unambiguous ontology that can serve as the basis for business process modelling and business case simulations’ (Haaker, Faber, & Bouwman, 2006). As a result, business models developed towards encompassing a complex set of design choices concerning a specific value network, functional architecture, financial models, and eventual value propositions made to the user. Since a lot of choices are involved, the notion of control became as important as the concept of value in analysing business models (Ballon, 2009).

The specific analytical framework used in this paper, The Business Model Matrix, is based on exactly such conceptualisation. This approach emphasizes roles and relationships between different actors within one ecosystem, rather than confining to processes within single businesses. The matrix combines four parameters of business models: (1) the value network showing how actors and roles, resources and capabilities are distributed in the value network, (2) the functional architecture showing how elements thereof contribute to the value creation process, (3) the financial construction containing revenue streams and revenue sharing models, and (4) the value proposition referring to the product or service offered to the end user. Not all elements of the matrix are of equal relevance for this paper and information on some parameters proved impossible to find. We therefore focus on a subset of parameters, which are further defined.

Level of vertical integration: Vertical integration relates to the scope of tasks one firm takes over in the value creation process. The higher the level of ownership and control over successive stages of the value chain, the higher the vertical integration. *Interoperability:* Interoperability refers to the ability of technological systems to directly exchange information and services with other systems, and to the interworking of services and products originating from different sources. *Revenue model:* The main specification of this aspect is to examine the business model in question of income stream (direct/indirect). *Positioning:* This aspect deals with the question whether to position a product or service as a complement to a particular set of existing products and services, or rather as a substitute to them. It is not least a question of marketing issues including branding, identifying market segments, establishing consumer trust, detecting competing products or services, and identifying the most relevant attributes of the product or service in question. *Customer involvement:* Customer involvement refers to the role of users in the business model and the scope of their integration possibilities. ‘Prosumers’ are enabled to produce and consume content and services in parallel in the network.

The BM Matrix is applied to the analysis of three cases: Jaunt, NextVR and Immersive Media. Information was gathered using mainly online sources. Results benefitted from the feedback of experts within the EU H2020 project ImmersiaTV.

4. CASE STUDY SELECTION: END-TO-END IMMERSIVE PLATFORMS

In general, two main Value Chains can be distinguished within the value network for immersive audiovisual services: Hardware & Software on the one hand, and Content on the other hand. To generate immersive content, such content needs to be captured, then stitched together and then produced; the resulting product is distributed before being displayed. Every step of this content chain relies on specific hardware and software: 1) Cameras for capture; 2) Software to ensure the processing (i.e. stitching) of the captured content; 3) Software allowing to edit the processed content (selecting scenes, adding graphics and subtitles, sound engineering etc.); 4) Content delivery networks, specific codecs, players etc. to ensure distribution, and 5) Various devices for displaying and controlling the experience. Figure 1 provides an overview of the immersive audiovisual service value network.

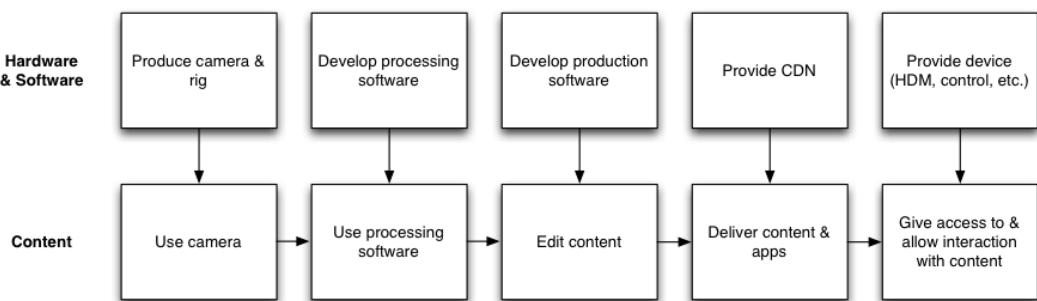


Figure 1. Immersive audiovisual service value network

Such overall value network for immersive audiovisual services consists of many small players, contributing specific assets to VR products (e.g. producers of cameras, production companies, HMD providers etc.). For our case study, however, we have chosen three end-to-end platforms, i.e. providers of immersive experiences and/or of solutions covering production, distribution and sometimes even display aspects of immersive audiovisual content. This includes platforms giving access to live events since this implies an integration of activities from capture to production or distribution. At the moment, there are already several providers of such experiences², three of which will be evaluated in this paper. They are chosen as they provide interesting cases when all steps in the value chain can be touched upon rather than focusing on one aspect of production and delivery of immersive audiovisual content.

5. EMPIRICAL ANALYSIS OF THREE END-TO-END PLATFORMS

5.1 Jaunt

Founded in 2013 and based in Palo Alto (US), Jaunt is a VR company focused on cinematic virtual reality. It is one of the most prolific 360 degree, cinematic VR content providers over the last couple of years, with a dozen or so experiences already (James, 2015) (in June 2015). In February 2016, Jaunt was included on American business magazine *Fast Company's* list of *The World's 50 Most Innovative Companies 2016*. Jaunt was recognized "for becoming the first VR media company". In June 2017, its VR app was officially launched in China (Jaunt VR, 2017).

5.1.1 Value Network

Jaunt has created an end-to-end solution to record, edit, produce, and deliver stereoscopic virtual reality experiences for various content – including narrative storytelling, music, travel/adventure, and sports (Clark, 2015). They have developed an integrated pipeline of hardware and software tools, including their

² Other examples include Condition One, LiveLike, LiveSphere, Perfant, Reality Lab, Scratch VR, Specular Theory, Teradek, Total Cinema 360, VantageTV, Vahana, Voke VR, VRcade and Vrtify.

own 360-degree cameras Jaunt ONE (J1-24G), released in 2015;³ proprietary computational photography algorithms to transform the recorded video data; (Jaunt VR, n.d.-b) a full suite of post-production tools including editing, colour correction and compositing; Jaunt Player (Jaunt's VR video player); and finally the Jaunt app that allows worldwide distribution for Cinematic VR experiences across all platforms and headsets (Jaunt VR, n.d.-a). For example, CBS' behind-the-scenes report about an unmanned mission to Mars, has been added to the Jaunt app (Clarke, 2016). Thus, Jaunt also serves as a one-stop shop that can handle the distribution side, making available its projects through Oculus, iOS, and Android apps (Bishop, 2015).

Since August 2016, Jaunt gives access to 300 videos created by third parties through Jaunt Publishing, a programme that allows the growing community of professional VR creators to publish their high-quality VR content directly to the Jaunt VR app. Through an online portal on the company's website, Jaunt Publishing enables creators to submit their work for consideration and publishing across the Jaunt platform. Once content is approved by an internal review board, creators will have access to Jaunt Cloud, and the publishing tools within including transcoding, "deep-links," support for premium spatial audio formats like Dolby Atmos, and processing and preparation for distribution on all VR platforms. (Jaunt VR, 2016)

In 2015, Jaunt launched Jaunt Studios, a division to create original VR content. (Clark, 2015; Clarke, 2016; Zeitchik, 2015) As the company's CEO stated in an interview: "We spent the past two years focusing entirely on building out the technology necessary to do content creation in VR (...) and the final missing piece of the puzzle is the content piece." (Bishop, 2015). Jaunt aimed to create as many as 1,000 pieces of content in the next 18 months, including scripted entertainment as well as sports, concerts and other live-event programming. (Zeitchik, 2015) Significant collaborations to produce content have included ABC News and Jaunt's release of 6 short videos in 2015 (Palmeri, 2016), as well as a partnership with Condé Nast Entertainment (CNE) to create several VR series that "explore Condé Nast's strong portfolio of travel, lifestyle, fashion, sports, and technology content." (Zeitchik, 2015)

In sum, Jaunt is strongly vertically integrated, as it provides an end-to-end solution to produce and distribute 360 video content (except for devices allowing to access and watch content). Furthermore, it is getting more and more integrated since it now produces content itself. In parallel, Jaunt has opened its distribution platform to third party creators.

5.1.2 Functional Architecture

As the company claims and several sources confirm, interoperability is a key objective for Jaunt. They aim at delivering content to the wide array of mobile devices and VR hardware in the industry (Pitchbook, 2015). Also Jaunt Cloud not only works with their own cameras but also with the Nokia Ozo (Nelson, 2016). The main reason, Jaunt claims (notably when promoting Jaunt Publishing), is that multichannel distribution can be seen as a sometimes paralyzing challenge for creators in the nascent VR industry (Jaunt VR, 2016). Creations need to be available as broadly as possible.

5.1.3 Financial Model

Jaunt's revenue model is mainly indirect. Therefore, partnerships are crucial – as exemplified previously. They are also the early sources of funding for content. (Clarke, 2016) End-users are important because they are targets for the commissioning entity (for example for advertising), but they never provide direct revenue.

It is worth noting though that Jaunt is still experimenting with ways to earn money and makes no secret of it. Thus, while Jaunt was initially not targeting to generate sales through device sales or rentals (Sjöblom, 2015), the Jaunt ONE camera is currently available for rent or purchase via a small number of resellers. (Jaunt VR, n.d.-a) Furthermore, Jaunt plans to test whether people would like to pay to download an episode or series. (Clarke, 2016).

5.1.4 Value Proposition

Jaunt is distinctively emerging into the cinematic industry (Sjöblom, 2015), with its unique offer. It could however be that it will replace existing, traditional experience. Its positioning is therefore as a substitute. Jaunt proposes a high involvement of professional users. This is the case for both partners (who finance the content produced by Jaunt) or creators. In particular Jaunt Publishing is a form of user-generated content platform aimed at professional creators of 360 video content. Involvement of end-users is in contrast quite low.

³ The first professional-grade camera system specifically designed for capturing high-quality 360°, stereoscopic, cinematic VR experiences, which features 24 camera modules (Jaunt VR, n.d.-a)

5.2 NextVR

Founded in 2009 and based in Laguna Beach, CA, [44] NextVR aims to enable the transmission of live, long-form virtual reality content in broadcast quality, for sporting events, concerts, cinematic productions, etc. Its focus is on live streaming of VR videos of live events.

5.2.1 Value Network

NextVR has an end-to-end solution, including cameras and encoders, platforms for distribution, as well as playback (Sjöblom, 2015). NextVR has its own stereoscopic 360-degree VR Cam. In terms of production, besides its pieces of software, NextVR owns a virtual reality production truck, to "allow for rapid deployment at an arena or stadium for the increasing demand for live virtual reality content." It is designed to "plug and play", meaning it can go to any given venue and deliver a multi-camera, live stereoscopic VR experience complete with fully mixed 3D VR audio. (Hamedy, 2016) Regarding distribution NextVR says they can stream the content over an average home internet connection using their proprietary streaming solution. (Lang, 2014) It also proposes the NextVR app, inside which there is a range of regularly scheduled content, including dedicated partners like FOX Sports and Live Nation. (NextVR, n.d.) Finally, NextVR is producing content, in particular live content (sport, concerts). In May 2016, the company was live-streaming about one sporting event a week. What they want to do next is to "build a gravitational pull" of content that people return to time and time again to see what's happening next — like they would with any other channel. (Hines, 2016)

In sum, NextVR's Value Network is vertically integrated, with the exception of display. There is no indication however towards further vertical integration.

5.2.2 Functional Architecture

NextVR's service does not show much interoperability. The NextVR App can be downloaded on the Oculus Store. (Oculus, n.d.-a) It requires a Samsung Gear VR headset, which is powered by Oculus, and a Samsung mobile phone. (NextVR, n.d.) However NextVR mentions interoperability as an aim to expand their offerings to other hardware (NextVR, n.d.; Ranaivoson et al., 2017) such as the Google Daydream View headset.

5.2.3 Financial Model

NextVR's revenue model is fully indirect. NextVR only uses its technology to profit by streaming live events. Thus, the company is not selling the camera system – they are rather using it to produce content. (Lang, 2014) The broadcasting of 360 videos of live events is sold as a service, but not to the audience. Until now, NextVR has not charged their end-users for their app, experiences, and live streams. (NextVR, n.d.) For these reasons, partnerships are crucial. NextVR's current and former partners include FOX Sports (notably with a five-year partnership started in 2015), Live Nation, NBC Sports, HBO/Golden Boy, Turner Sports, and CNN. (NextVR, n.d.)

5.2.4 Value Proposition

NextVR positions itself as the leader in broadcasting live events in virtual reality. (Oculus, n.d.-b) This makes the company complementary to the event and to the traditional broadcasting of the event. This also explains why it is always closely working with event organisers and broadcasters. This close cooperation corresponds to a high user involvement to ensure the correct functioning of the live streaming of the event. On the other hand, consumers are not much involved.

5.3 Immersive Media

Based in Kelowna (Canada) and founded in 1994, Immersive Media is a historical player in the Virtual Reality sector with the production in 1995 of the world's first full motion, fully immersive movie. Roughly a decade later, the organization launched its first 360°camera, the Dodeca. The first big undertaking to put it to use was the Geoimmersive City Collection project. Since 2010, Immersive Media's focus has shifted to media and entertainment providing radical technological approaches to traditional media. This includes live events streaming, in 360°to computers and mobile devices. (McGovern & Poss, 2016)

In 2014 Immersive Media entered into a joint venture with Digital Domain, one of the main visual effects studios in the world. This led to the creation of IM360, a provider of VR content solutions. The aim was to leverage the firms' complementary technology portfolios for the fast-rising VR market. In 2015, Digital Domain Holdings Ltd purchased Immersive Media. (Johnson, 2017; McGovern & Poss, 2016).

5.3.1 Value Network⁴

Immersive Media provides an end-to-end solution, including cameras and encoders, platforms for distribution, as well as playback, with their own software to process, store and deliver the content (Sjöblom, 2015). This notably applies to live experiences, from and to anywhere in the world, relying on Immersive Media's imLive System to simultaneously stitch, encode and stream live 360 video capture, while also recording direct-to-disc.

Immersive Media has proprietary HD 360 cameras and capturing systems. For production, the IM360 VR Toolkit provides directors the capabilities for shooting, capturing and directing virtual reality content. The VR Toolkit uses Immersive Media's imLive System to integrate the Director into a virtual video village for real-time monitoring/directing of VR shoots. IM360's proprietary server platform encompasses a network of web services and client applications providing interactive media management solutions for immersive content. Through their branch IM360, Immersive Media has an app development division solely dedicated to the creation, design, development and integration of VR-based apps for all popular devices & OSs. Around 2014, Immersive Media increased the focus on content production by creating IM360 in a joint-venture with Digital Domains. Thus, IM360 claims to have produced the first-ever live boxing match in 360° in January 2016 (Burns, 2016).

In sum, Immersive Media is vertically integrated (excluding display). Integration is increasing since Immersive Media is getting more involved in VR content production.

5.3.2 Functional Architecture

Immersive Media aims at interoperability. Their objective is universal playback of 360 video experiences across all major platforms such as desktop, mobile and VR devices. The platform also makes it easy for viewers to share their favourite immersive experiences directly on Facebook, Twitter, Pinterest, Google+ or other social networks. (Immersive Media, n.d.-b)

5.3.3 Financial Model

Immersive Media combines direct and indirect revenue models. The indirect revenue model relies on the partnerships the company has developed over the years. An interesting example is the development of a 360 Video App for Taylor Swift's *Blank Space* in 2014. The app allows to watch and interact with the video. In terms of interaction for example the user can wander in different rooms rather than follow the singer, they can also find collectibles in the various rooms. The app relies on a partnership between Immersive Media, Taylor Swift's label and the bank American Express. Actually, the app is entitled *American Express Unstaged: Taylor Swift Experience*. While one does not need to have an American Express account to experience the interactive music video, it was however required for access to exclusive ticket and event offers. (AMEX, n.d.) From a creative point of view, Immersive Media worked with Radical Media and Grammy Award winning director Joseph Kahn. (Spock, 2015) The project was very well received and even won an Emmy (the first for Immersive) (McGovern & Poss, 2016).

In addition, though, Immersive Media has direct revenues through the offering of its products. In particular it sells its cameras, the targets being professional users (Sjöblom, 2015). It is also possible to purchase the im360 (iOS) SDK to create fully customizable 360° experience. (Immersive Media, n.d.-b).

5.3.4 Value Proposition

Immersive Media insists on their technological uniqueness, e.g. "For more than *twenty years*, we've been the experts at designing & building 360° technology and experiences." (Immersive Media, n.d.-a) They claim to own the most patents and intellectual properties in the industry covering 360 video and virtual reality. (IM360, n.d.) Value Proposition can be considered mainly as complementary as Immersive Media is working

⁴ Some of the information in this section was retrieved from the homepage, a services pages, a technology page and an experiences page of IM360 (www.im360.info). However, following the company's merger, these pages have been deleted.

in partnerships with other stakeholders, such as brands or event organizers by creating interactive experiences for them. Examples include the New York Time's VR platform and SyFy's Expanse app. Immersive Media includes the ability to render videos, but also apps and content management. (Dawson, 2015) User involvement is high with professional users, e.g. sports event organisers or in the case of the American Express unstaged: the Taylor Swift Experience case. However, and in contrast to the two other cases, it is lower for activities where Immersive Media is selling pieces of software or hardware.

6. DISCUSSION

The following table summarises the main business model features of our three cases. The cases show significant similarities in the design choices made. First, they are all vertically integrated; this is to be related to the fact that their revenue model is, at least mainly, indirect and that they have a high user involvement. Actually, all three features point to the importance of partnerships for the three companies. These partnerships are more than a mere commissioning of content. They take place with different types of stakeholders, including broadcasters, organisers of live events, brands, etc. They rely on various configurations: with one or several partners, for content that is available for free or only for subscribers. Reasons for such partnerships from these companies' point of view include that they are a source of income, but also that they reinforce their reputation.

The cases differ as far as interoperability and positioning are concerned. Interoperability seems less of a must for NextVR, maybe because it represents an additional cost. Jaunt's positioning is rather substitutional because the experience they offer could replace what traditional media is offering. On the other side, NextVR and Immersive Media propose services or products complementary to the way media (and events) are currently functioning. This is consistent with the analysis currently made that VR will succeed thanks to its complementarity, with fans still mainly watching the event on their television sets, but using VR for enhancements: game recaps, highlights of a particular play, features, brief forays into watching a certain portion of the action live, etc. (Clark, 2015; Zeitchik, 2015).

Table 2. Overview of case study findings

	Jaunt	NextVR	Immersive Media
Vertical integration	Yes & move towards content production	Yes	Yes & move towards content production
Interoperability as an objective	Yes	No(t yet)	Yes
Revenue model	Mainly indirect	Indirect	Mainly indirect
Positioning	Substitute	Complement	Complement
User involvement	High	High	High, with exceptions

These business model features clearly point towards the now acknowledged importance of content innovation. This is particularly the case for vertical integration, which includes content production in different manners. This integration is a recent evolution for Jaunt and Immersive Media. This is comparable to the strategy followed by distribution platforms such as HTC Viveport, Facebook's Oculus Home and Google's YouTube VR (Ranaivoson et al., 2017). The first two are involved in first party content production while the latter, via Google Jump, provides technologies to help VR content creators.

7. CONCLUSION

This paper has highlighted the main business model features, and has derived the analysis of the articulation of media innovation in the case of the emerging VR market. To do so, it has first provided a value network of the production and distribution of immersive audiovisual content. It has then provided an in-depth analysis of three companies providing end-to-end immersive audiovisual services.

Taking immersive audiovisual services as a case, this paper has demonstrated how content innovation is crucial for media, and how it becomes apparent in the business model features of end-to-end platforms. This does not exhaust the topic of how to ensure such innovation is sustainable from the content producers and distributors' point of view. In particular, it raises questions in terms of how such innovation could be promoted by innovation and media policies.

ACKNOWLEDGMENTS

This paper has been written in the context of ImmersiaTV (<http://www.immersiatv.eu>), a project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688619.

REFERENCES

- AMEX. (n.d.). Amex UNSTAGED: Taylor Swift Experience. Retrieved November 29, 2017, from <https://www.americanexpress.com/us/content/unstaged-app/faqs.html>
- Ballon, P. (2009). *Control and Value in Mobile Communications: A Political Economy of the Reconfiguration of Business Models in the European Mobile Industry* (SSRN Scholarly Paper No. ID 1331439). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=1331439>
- Bishop, B. (2015, April 28). VR company Jaunt is starting a virtual reality movie studio - The Verge.pdf. *The Verge*. Retrieved from <http://www.theverge.com/2015/4/28/8507091/jaunt-studios-interview-virtual-reality-movies-360-degree-camera>
- Bleyen, V.-A., Lindmark, S., Ranaivoson, H., & Ballon, P. (2014). A typology of media innovations: Insights from an exploratory study. *The Journal of Media Innovations*, 1(1), 28–51.
- Burns, C. (2016, February 11). PGA Tour Golf VR made real with IM360. Retrieved November 29, 2017, from <https://www.slashgear.com/pga-tour-golf-vr-made-real-with-im360-11426550/>
- Cave, J., & Frinking, E. (2007). Public Procurement for R&D, 24.
- CC Insight. (n.d.). Augmented and Virtual Reality Devices to Become a \$4 Billion-Plus Business in Three Years. Retrieved November 29, 2017, from <http://www.ccsinsight.com/press/company-news/2251-augmented-and-virtual-reality-devices-to-become-a-4-billion-plus-business-in-three-years>
- Chambel, T., Chhaganlal, M. N., & Neng, L. A. R. (2011). Towards immersive interactive video through 360° hypervideo (p. 1). ACM Press. <https://doi.org/10.1145/2071423.2071518>
- Clark, W. (2015, September 28). Jaunt's CEO on its \$65M funding and the humanity-advancing mainstream future of virtual reality. Retrieved November 29, 2017, from <https://venturebeat.com/2015/09/28/jaunts-ceo-on-its-65m-funding-and-the-humanity-advancing-mainstream-future-of-virtual-reality/>
- Clarke, S. (2016, March 1). Will VR rock in 2016 - TBI Vision TBI Vision.pdf. *TBI Vision*. Retrieved from <http://www.tbivision.com/highlight/2016/03/will-vr-rock-2016/555472/>
- Dawson, T. (2015, November 5). IM360 VR to Enable Virtual Reality Content Creation | Androidheadlines.com. Retrieved November 29, 2017, from <https://www.androidheadlines.com/2015/11/im360-vr-to-enable-virtual-reality-content-creation.html>
- Digi-Capital. (2016). Virtual, augmented and mixed reality are the 4th wave | NEWS >> Digi-Capital. Retrieved November 29, 2017, from <http://www.digi-capital.com/news/2016/07/virtual-augmented-and-mixed-reality-are-the-4th-wave/#.WMEfUy0rLDc>
- Digi-Capital. (2017). Record \$2.3 billion VR/AR investment in 2016 | NEWS >> Digi-Capital. Retrieved November 29, 2017, from <http://www.digi-capital.com/news/2017/02/record-2-3-billion-vrar-investment-in-2016/#.WMEawi0rLDc>
- Drucker, P. F. (2006). *The Practice of Management* (Reissue edition). New York, NY: HarperBusiness.

- Freeman, C., & Soete, L. (1997). *The Economics of Industrial Innovation, 3rd Edition*. The MIT Press. Retrieved from <http://ideas.repec.org/b/mtp/titles/0262061953.html>
- Goldman Sachs. (2016). *Virtual & Augmented Reality. Understanding the race for the next computing platform* (Equity Research).
- Gopal, R. (2016). VR Movie Revolution Goes Dutch. Retrieved November 29, 2017, from <http://www.ccsinsight.com/blog/vr-movie-revolution-goes-dutch>
- Haaker, T., Faber, E., & Bouwman, H. (2006). Balancing customer and network value in business models for mobile services. *International Journal of Mobile Communications*, 4(6), 645–661.
- Hamedy, S. (2016). World's first virtual reality production truck has arrived. Retrieved November 29, 2017, from <http://mashable.com/2016/04/20/virtual-reality-production-truck-nab-show/>
- Hines, N. (2016). NextVR Founder: The Hardware Is Here; We're Just Waiting on the "Flood of Content." Retrieved November 29, 2017, from <https://www.inverse.com/article/15410-nextvr-founder-the-hardware-is-here-we-re-working-to-bring-the-flood-of-content>
- IM360. (n.d.). TECHNOLOGY. Retrieved November 29, 2017, from <http://www.im360.info/technology.html>
- Immersive Media. (n.d.-a). Hardware – Immersive Media. Retrieved November 29, 2017, from <http://immersivemedia.com/hardware/>
- Immersive Media. (n.d.-b). Software – Immersive Media. Retrieved November 29, 2017, from <http://immersivemedia.com/software/>
- James, P. (2015, June 23). Jaunt and Google Announce "High-End" Cinematic Content Collaboration. Retrieved November 29, 2017, from <https://www.roadtovr.com/jaunt-and-google-announce-high-end-cinematic-content-collaboration/>
- Jaunt VR. (2015, November 19). Inside the Jaunt ONE: Why a VR Content Company Had to Build the World's Best VR Camera. Retrieved November 29, 2017, from <https://www.pdnonline.com/gear/cameras/video/inside-the-jaunt-one-why-a-vr-content-company-had-to-build-the-worlds-best-vr-camera/>
- Jaunt VR. (2016, August 4). Jaunt Further Expands Its Cinematic VR Ecosystem With The Launch Of Jaunt Publishing. Retrieved November 29, 2017, from </news/jaunt-publishing>
- Jaunt VR. (2017, June 13). Jaunt VR App Officially Launches in China on Xiaomi VR Headset. Retrieved November 29, 2017, from </news/jaunt-vr-app-officially-launches-china-xiaomi-vr-headset>
- Jaunt VR. (n.d.-a). Jaunt ONE. Retrieved November 29, 2017, from <https://www.jauntvr.com/jaunt-one/>
- Jaunt VR. (n.d.-b). Technology - Jaunt. Retrieved November 29, 2017, from <https://www.jauntvr.com/technology/>
- Johnson, L. (2017). This Vp of VR and AR Productions Was Once Google Street View's Driver No. 6 [Adweek.com]. Retrieved from <http://www.adweek.com/digital/how-google-street-view-driver-no-6-spent-his-summer/>
- Keen, P., & Qureshi, S. (2006). Organizational transformation through business models: a framework for business model design. In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on* (Vol. 8, p. 206b–206b). IEEE.
- Lang, B. (2014, September 12). NextVR's Stereoscopic 360 Degree VR Cam Uses \$180,000 Worth of RED 6K Cameras. Retrieved November 29, 2017, from <https://www.roadtovr.com/nextvr-stereoscopic-360-degree-vr-cam-uses-180000-worth-of-red-6k-cameras/>
- Lindmark, S., Ranaivoson, H., Donders, K., & Ballon, P. (2013). Innovation in small regions' media sectors. Assessing the impact of policy in the Flemishcase. In *Storsul T., Krumsvik A. H., (eds), Media Innovations. A Multidisciplinary Study of Change* (pp. 127–144). Gothenburg: Nordicom. Retrieved from www.nordicom.gu.se/sites/default/files/publikationer-hela-pdf/media_innovations_a_multidisciplinary_study_of_change.pdf
- Magretta, J. (2002). Why business models matter. *Harvard Business Review*, 80(5), 86–92.
- McGovern, T., & Poss, J. (2016, January). From Streetview to the battlefield and beyond. Immersive Media's 360° Viewing Revolution. White Paper, Immersive Media.
- Nelson, N. J. (2016, February 9). Virtual Reality's Distribution Frontier Is Heating Up. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/turnstyle/virtual-realitys-distribu_b_11837214.html
- NextVR. (n.d.). NextVR. Retrieved November 29, 2017, from <http://nextvr.com/>
- Oculus. (n.d.-a). NextVR. Retrieved November 29, 2017, from <https://www.oculus.com/experiences/gear-vr/858258597574484/>
- Oculus. (n.d.-b). Oculus Gear VR Experiences. Retrieved November 29, 2017, from <https://www.oculus.com/experiences/gear-vr/>
- OECD, & Eurostat. (2005). *Oslo Manual*. OECD Publishing. <https://doi.org/10.1787/9789264013100-en>

- Palmeri, C. (2016, October 13). For a Clue to Disney's View of Future, Look at Startups It Aids. *Bloomberg.com*. Retrieved from <https://www.bloomberg.com/news/articles/2016-10-13/for-a-clue-to-disney-s-view-of-future-look-at-startups-it-aids>
- Pitchbook. (2015). *Virtual Reality. 2015 Analyst Report*.
- Ranaivoson, H., Delaere, S., & Xia, Q. (2017). *D5.1 Market analysis report*. ImmersiaTV.
- Sawy, O. A. E., & Pereira, F. (2012). *Business Modelling in the Dynamic Digital Space: An Ecosystem Approach* (2013 edition). Heidelberg: Springer.
- Schumpeter, J. A. (1942). *Capitalism, socialism, and democracy*. New York, London,: Harper & Brothers.
- Sjöblom, S. (2015). Competitive Intelligence—Conducting an Analysis of a Business Environment. Retrieved from <http://dspace.cc.tut.fi/dpub/handle/123456789/23507>
- Spock, J. (2015, February 12). Taylor Swift 360 Degree Music Video By Immersive Media. Retrieved November 29, 2017, from <https://virtualrealityreporter.com/360-degree-music-video-taylor-swift-blank-space/>
- Zeitchik, S. (2015, April 28). VR Watch: Jaunt forms studio, hiring Lucasfilm veterans to run it. *LA Times*. Retrieved from <http://venturebeat.com/2015/09/28/jaunts-ceo-on-its-65m-funding-and-the-humanity-advancing-mainstream-future-of-virtual-reality/>

SPORT TOURISM: IMPROVING VISITORS' EXPERIENCE THROUGH WEB 2.0 TOOLS IN SOCCER CLUBS' MUSEUMS AND STADIUMS

Deborah Moraes Zouain, Paola Bastos Lohmann, Marcio Ferreira Bezerra,
Gabriela De Laurentis Cardoso, Kaarina Barbosa Virkki and Pedro Henrique Domingues Curvello
Universidade do Grande Rio - UNIGRANRIO
Rua da Lapa, 86, 9th Floor – Lapa – Rio de Janeiro - Brazil

ABSTRACT

Information and communication technology is increasingly integrated in economic sectors with the purpose of improving access and attracting more consumers. In the tourism market, travel processes are facilitated through online platforms, whether for booking hotel rooms, restaurants, buying tickets or for getting information and maps for better directions. The periods in which visitors experience most search interaction with their destination are the ones prior to and during the trip, as they are involved in planning and making decisions about the places they will visit. In this process, websites and applications are among the main sources of information. By using these tools, destinations have come to include the “smart” concept, according to which they become increasingly competitive and seek to provide visitors with an experience of excellence, as well as to improve their own management. Therefore, the goal of this article is to investigate how brasiliian and foreigners’ soccer clubs are using technological tools to enhance the visitor experience and their interaction with those same visitors. Therefore, a sample of 21 soccer clubs’ websites were analysed using 10 criteria based on the concepts of Web 2.0, as proposed by López et al. (2010). Results showed that social networks are very important tools for interacting with website users, as well as an extensive use of multimedia to divulge services, including videos, pictures and sound. Although there is still a misusage of the virtual interaction between clubs and visitors. Additionally, it was found that soccer clubs and museums should invest in web 2.0 tools as a way to improve sport tourism experience for visitors.

KEYWORDS

Sports Tourism, Smart Tourism, Information Technology, Web 2.0

1. INTRODUCTION

“Tourism is one of the world’s most important economic sectors, accounting for 9.8% of the GDP and 5.4% of exports worldwide, which corresponds to 7.6 trillion dollars annually” (Turner and Freiermuth, 2016, p.1). According to Wright (2007), one of the fastest growing segments in this worldwide industry is sports, in which people from all over the world travel in search of new sports experiences, both as practitioners and spectators. These activities include visits to stadiums, museums, giftshops and events.

The technological revolution centered around information and communication technologies (ICT) imposes a new pace on society, leading to new ways of communicating and managing, according to Hassan (2011). Thus, new technologies are changing people's routines, behaviors and ways of relating to each other. From this revolution, people can produce and disseminate information in real time with the help of the internet, and consequently share experiences and have persuasive power over potential consumers. In the tourism context, besides serving as a source of interaction and information, the use of technology becomes a competitive advantage as it can positively influence the visit experience.

For Hassan (2011), the ICTs seem to be one of the most critical areas for both tourism success and the way of promoting destinations in the future, since smart systems seize technology-generated resources to provide smart applicabilities. Muñoz and Sánchez (2013) point out that the challenge for the tourism industry is to integrate the evolution of ICTs into tourism destinations, i.e., to turn them into smart tourism destinations.

The consumer shifted from the industrial era's isolation to present-day's connectivity, from insecurity to information and from passivity to activity. In this way, the impact of connected, informed and active consumers begins to manifest in different ways, allowing greater interaction not only between demand and supply, but also between demand and demand and supply and supply itself (Prahalad & Ramaswamy, 2004).

According to López et al. (2010), the internet has become a place for participation, collaboration and social exchange through Web 2.0 tools, thus becoming the largest contemporary source of information on tourism destinations, hotel options, transport, food and local attractions, mainly via websites which provide real-time information for better decision making about the trip.

Thus, in order to understand the use of technological tools in sports attractions, this article will comparatively analyze the official websites of 21 Brazilian and foreign soccer clubs that offer stadium and museum visits, according to criteria based on Web 2.0 concepts that promote a greater interaction with users and visitors. To that purpose, a qualitative exploratory study was conducted using secondary data collection from the following clubs: Real Madrid, Barcelona, Benfica, Porto, Sporting, Boca Juniors, River Plate, Chelsea, Manchester United, Arsenal, Bayern Munich, Juventus, Grêmio, Internacional, São Paulo, Corinthians, Palmeiras, Atlético Paranaense, Coritiba and Botafogo.

2. LITERATURE REVIEW

2.1 Information Technology and Web 2.0

According to Zhang et al. (2012), ICTs aim to meet tourists' individual needs, to provide enhanced trip quality as well as satisfaction with tourism products and services, to promote innovation in tourism management, and to optimize resources in order to improve their competitiveness.

Silva and Mendes Filho (2016) advocate technology as a source of innovation, strategic differentiation and competitive advantage to co-create experiences in the future, when multiple levels of engagement and interaction networks allow a richer, more customized experience, as well as distinct value for the tourist, with connectivity being essential for more customized information anytime and anywhere. "This technological and social revolution was made possible by the creation of the Internet in the early 1980's and has taken on gigantic proportions in recent years with the creation of mobile technologies" (Paffrath & Cassol, 2014, p. 429).

The ICTs revolution is having a deep, fundamental impact on the way trips are advertised, distributed, sold and delivered, since the business that covers it is practically formed by information. Buhalis and Amaranggana (2015) said that making a destination smart requires incorporating technology into this web of connections. Among the actors are tourism companies (hotels, restaurants, airlines, tour operators, etc.), other industries and support organizations (arts, entertainment, sports, etc.) and institutions that manage destinations, as well as local residents and other tourists, according to Crouch (2011). Therefore, destinations are complex coproduction networks.

According to Beritelli et al. (2007), the success of the individual actors involved and the destination as a whole depends on the effective coordination and integration of resources, products and services. "This phenomenon of active participation, engagement, customization, collaboration, and content sharing is known as Web 2.0" (O'Reilly, 2005, p.1). According to MacArthur (2007), examples of these tools can be: blogs, wikis, forums, podcasts, Facebook, YouTube and iGoogle.

According to Lopez et al. (2010), using new tools on websites and taking advantage of their potential means providing a unique way of promoting new cultural and learning opportunities and making them available. To the clubs, this potential exists particularly through the involvement of users in active participation, thus improving their museum and stadium visiting experience.

2.2 The Smart Phenomenon

The current use of Web 2.0 tools, applying these new technologies into daily life appear to offer sharing and learning further about the place you have interest, or you are about to visit, for example. The user experience in a website is no longer only a passive source of information but a dynamic and interactive one. The new consumer according to Euromonitor 2018 tendencies report (Angus, 2018), points out that shifting consumer

attitudes and behaviours will continue to cause disruption for business in 2018, with mobile technology and internet accessibility playing a key role in shaping these changes.

The soccer clubs are heading a maturity as "smart marketing" organizations as they develop a sense of connecting the club with their fans using technology to build a stronger link, such as using their social network to tell more about the backstage of their training, matches and next moves, for example. Facilitating processes like buying a ticket online to visit the soccer stadium is also another "smart" move to engage more followers and rise the visitation number. The soccer itself, is part of a larger network in the soccer business, and offers to their fans a greater connection, active experiences, such as visiting the stadium, knowing more about the soccer players and the team, among other.

According to Boes et al (2015), the term "smart" is a marketing word for anything that is embodied or enhanced by technology. This increasing dependence of tourism market on ICTs arises from the need to adapt to consumers' new behaviors, as they now use communication media for greater convenience when planning a trip, from the time of researching what destination to visit, to buying tickets, booking accommodation, and closing the cycle with feedback, which consists of the sharing of experiences on social media. "Social networks allow digital marketing organizations and tourists an environment to interact, comment and share their individual experiences and build a community environment" (Hays et al, 2012 apud Neuhofer et al., 2012).

The traveler essentially goes through three phases during which he constantly absorbs information: before, during and after the trip. In the first phase, he researches transportation, attractions, documents, accommodation, food, among others, this being a period of extensive searching on websites, social networks and applications. In addition, the Internet is an ICT that makes content widely available quickly and easily, thus allowing the emergence of websites that specialize in marketing tourism online, according to Vicentini and Hoppen (2002).

During the trip, which is the peak of his experience, the tourist comes in contact with several actors which make up the tourism system, and the integration between them can occur with the help of technologies for producing and sharing information in real time. Examples include buying air tickets, getting information about tourist attractions and tickets, city maps featuring public transport routes, recommended restaurants, among others.

As for the post-trip period, tourists share experiences, reinforcing one of the main uses of the internet, which is to allow anyone to be a creator of content, with travelers reporting their opinions about accommodation, restaurants and touristic attractions. User-generated-content (UGC) "has the potential to influence consumers' travel planning and decision-making process. For instance, online consumers are likely to employ usergenerated reviews for their travel planning so long as they perceive the source to be credible—regardless of whether the review reflects the actual truth or not". (Ayeh et al, 2013, p. 447).

2.3 Sports Tourism

The segment's rapid growth, as Ribeiro (2014) notes, is related to a number of factors such as the democratization of sports, investments in sports infrastructures, and the growth in tourism and mobility options. These activities, according to Ross (2001) and Carvalhedo (2003), can be summarized in three types of behavior: active sports, event and visiting tourism (this last being related to museums and stadiums, for example). With regard to sports tourism, Ross (2001) says it tends to be nostalgic, as it is related to a search for the past, like in museums, halls of fame, among others.

"Today's tourists are not satisfied simply with tourist attractions, they also seek greater, deeper knowledge of the places they visit" (Pezzi e Vianna, 2015, p. 170). Entertainment-consuming citizens tend to connect to experiences that are relevant to and in line with their lifestyle, and define certain experiences as representing something more than the apparently ordinary ones, according to Hiller (2011).

In the case of soccer, museums and stadiums give sports lovers the opportunity to be in places that are inaccessible on match days, where their idols were when they achieved their victories and shaped history. This allows the visitor to explore his imagination and relive remarkable moments.

Because this is an attraction that involves passion, clubs understand the importance of fans' opinions and invest in increasingly efficient channels to understand what their demands are, how they have rated their visits, and the improvements that can be made. The continuous development of web tools belonging to a new generation known as Web 2.0 has opened new possibilities for clubs to communicate and interact with their audiences, according to Antinucci (2007) and Dawson (2008).

3. METHODOLOGICAL PROCEDURES

This study is a qualitative and exploratory research, a model that, according to Zikmund (2000), fits all who seek to discover ideas and intuitions, in an effort to acquire greater familiarity with the phenomenon studied, to explore alternatives or discover new ideas.

According to Gil (2007), the qualitative research allows a deeper investigation of questions, in pursuit of their meaning, based on the perception of the phenomenon within its context. In addition, we conducted bibliographical research, mainly through books, articles and periodicals.

As for the data collection technique, from September 04 to 15, 2017, we investigated the official websites of 21 soccer clubs – 8 of which Brazilian and 13 foreign – which offer museum and stadium visits. With regard to the selection process, we chose Brazilian clubs participating in the 2017 Brazilian Championship A Series, as well as foreign global brands considered some of the main tourist attractions in their cities. The Brazilian clubs are: Grêmio, Internacional, São Paulo, Corinthians, Palmeiras, Atlético Paranaense, Coritiba, Botafogo. The foreign clubs are: Boca Juniors and River Plate (Argentina), Chelsea, Manchester United and Arsenal (England), Bayern Munich (Germany), Juventus (Italy), Real Madrid and Barcelona (Spain), Benfica, Porto and Sporting (Portugal) and Paris Saint-Germain (France).

The goal is to comparatively investigate the use of clubs' website tools, using 10 criteria that serve as facilitators and provide greater visitor autonomy, to check the degree of integration with the "smart" concept, as well as the improvements that can be made. We chose these criteria because they deal with the possibility of greater website-visitor interaction through purchase, information and knowledge about the place, as well as information sharing on social networks, thus encompassing important concepts of the experience, which begins before being at the actual venue.

To that purpose, we designed a checklist with a structure indicating the presence or absence of a given feature or tool, based on the Web 2.0 concepts, and inspired by the studies of López et al. (2010). In their study, they investigated the presence of Web 2.0 tools in the websites of museums in 5 countries, and highlighted, in their results, the tools' real importance for the impact on the quality of the services proposed. The authors used 24 criteria to analyze 4 different types of museums: arts, natural sciences, humanities and specialized (e.g., museums of shoes, chocolate).

Thus, we selected criteria that could support a better evaluation of the sample of clubs, such as detailed information on visits, online purchasing, virtual tour, forums, spaces for comments, social network sharing and video content channels, totaling 10 items subdivided in two groups:

Static/non-interaction tools (providing information or purchasing features, rather than promoting interaction with visitors):

- ✓ Basic information;
- ✓ Availability in other languages;
- ✓ Online ticket sale;
- ✓ Multimedia (audio/videos/podcasts/animations) for showing information or further details on a specific subject;
- ✓ FAQ (Frequently Asked Questions)/Queries.

Dynamic/Interactive tools (offering the possibility of interaction with visitors):

- ✓ Online assistance: helping users with their doubts;
- ✓ Forum/comments: possibility to share opinions and interact with others;
- ✓ Video content channels;
- ✓ Buttons to link site resources to users' personal files on sharing networks (Yahooweb, Facebook, Twitter);
- ✓ Virtual tour.

4. RESULTS

In line with the ideas proposed by the concepts of smart tourism, we examined the implementation of technological tools at the websites of the attractions studied. Thus, we designed two analysis tables, each divided in five criteria, Table 2 being for static and 3 for dynamic tools.

Table 1. Presence of static/non-interaction tools in the clubs' websites

Clubs	Online sales	Basic information	Other languages	Internal Multimedia	FAQ
Grêmio	YES	YES	YES	NO	YES
Internacional	NO	YES	NO	NO	NO
São Paulo	NO	YES	NO	YES	NO
Corinthians	YES	YES	NO	YES	YES
Palmeiras	YES	YES	NO	NO	NO
Atlético Paranaense	NO	YES	NO	NO	NO
Coritiba	NO	YES	YES	YES	NO
Botafogo	NO	YES	YES	YES	NO
Real Madrid	YES	YES	YES	YES	YES
Barcelona	YES	YES	YES	YES	YES
Benfica	YES	YES	YES	YES	NO
Porto	NO	YES	YES	NO	NO
Sporting	NO	YES	YES	NO	NO
Boca Juniors	NO	YES	NO	YES	NO
River Plate	YES	YES	YES	YES	YES
Chelsea	YES	YES	YES	YES	YES
Manchester United	YES	YES	YES	YES	YES
Arsenal	YES	YES	NO	YES	YES
Bayern Munich	YES	YES	YES	YES	YES
Juventus	YES	YES	YES	YES	NO
Paris Saint-Germain	NO	NO	YES	YES	NO

Source: the authors' own elaboration.

According to the analysis of the five criteria, our findings are as follows:

a) Online Sales: A feature widely used by websites, online ticket sales have had an increase in popularity in the tourism industry. For consumers, this service promises time savings (Sigala, 2013). We found that this service is only available in 12 sites (57%), and, among the Brazilian clubs, only Palmeiras and Corinthians offer this feature. It is worth mentioning that Boca Juniors and Paris Saint-Germain, places largely sought by tourists, do not offer the service, which can be considered a weakness.

b) Basic information: Analysis of information about opening hours, ticket prices, on-site box office, tour itinerary, program, directions. Except for Paris Saint-Germain, all websites provided, in a non-standardized way, practical information so that visitors could get to the stadiums and museums. Because travel information is among the most popular and visited contents in the internet (Law and Leung, 2000a; Buhalis and Law, 2008), it is important to have them correctly presented to tourists in order to allow them to prepare effectively, relying on the accuracy of the information displayed.

c) Other languages: 14 of the 21 websites had the option of translating contents into other languages. Not having this tool can be considered a serious weakness to all the stadiums that wish to position themselves as an international tourist attraction. With the increase in the number of travelers, it is of utmost importance that websites offer the possibility of reading contents in other languages, especially in English, when a place positions itself as a tourist attraction.

d) Internal multimedia: Through this criterion, the study sought to determine whether the websites offered photos, videos and sounds for exploring more about the visit to the place. Such tools are a captivating form of presentation that allows visitors to learn about the experience the club has to offer. 15 of the 21 websites explored such features, investing in several videos and photos of the venues.

e) FAQ - Frequently Asked Questions: The most common doubts among users can be presented in this tool, where website administrators list a number of commonly asked questions and provide the answers so that new visitors can be more easily informed. Only 8 websites offer this feature. Among Brazilian clubs, only Corinthians and Grêmio do it.

Below are considerations on dynamic tools, which allow a greater interaction with the website.

Table 2. Presence of dynamic/interactive tools on club websites

Clubs	Virtual Tour	Forum	Online Assistance	YouTube	Buttons to social networks
Grêmio	NO	NO	NO	YES	YES
Internacional	NO	NO	NO	YES	YES
São Paulo	NO	NO	NO	NO	YES
Corinthians	NO	NO	NO	YES	YES
Palmeiras	NO	NO	NO	YES	YES
Atlético Paranaense	NO	NO	NO	NO	YES
Coritiba	NO	NO	NO	YES	YES
Botafogo	NO	NO	NO	NO	YES
Real Madrid	NO	NO	NO	YES	YES
Barcelona	NO	NO	NO	YES	YES
Benfica	YES	NO	NO	NO	YES
Porto	NO	NO	NO	NO	YES
Sporting	NO	NO	NO	YES	YES
Boca Juniors	NO	NO	NO	NO	YES
River Plate	NO	NO	NO	YES	YES
Chelsea	NO	NO	NO	YES	YES
Manchester United	NO	NO	NO	NO	NO
Arsenal	NO	NO	NO	NO	NO
Bayern Munich	NO	NO	NO	NO	YES
Juventus	YES	NO	NO	YES	YES
Paris Saint-Germain	NO	NO	NO	NO	YES

Source: the authors' own elaboration

According to our observation of the websites visited, our findings are as follows:

a) Virtual tour: It is a modern interactive tool which allows the user to browse through places by means of virtual reality, thus experiencing a preview of what he will see in the sports attraction. According to Buhalis and Law (2008), it allows the tourist to "experience" the place without actually visiting it.

Although it is a modern tool for presenting venues, only 2 websites include this feature, i.e., Benfica and Juventus. It is worth noting that the Juventus' is one of the "smartest" and most modern websites, and its identity has been restructured through a branding project carried out by Interbrand, with the aim of internationalizing the brand of the club (Interbrand, 2017).

b) Forum / Comments: A space for comments can be used to promote the venue if it receives positive reviews, thus helping potential visitors decide whether to go or not. It is worth remembering that websites like TripAdvisor are based on the idea that travelers rely on other travelers' opinions to plan their trips, or at least to be satisfactorily helped by their decision (Miguens, Baggio and Costa, 2008); yet, none of the 21 websites provided this interaction space.

c) Online Assistance / Chat: Like FAQs, online assistance is a useful tool to address queries not displayed on the website. According to Flouri and Buhalis (2004), these tools are very useful for obtaining information on tourist destinations. For example, people with reduced mobility or parents with baby strollers may want to know if the place is accessible, in order to avoid problems in their visit.

d) Video content channels: 52% of the websites (11 out of 21) had YouTube channels where they post videos, mainly of matches, goals and interviews with players and leaders; it is an option to be further developed. Many businesses already have video content channels, such as the ones on YouTube, which they use to promote their products and services, thus motivating consumption. In the case of museum and stadium tours, it helps to give visibility to new spaces open to visitors, as well as new items and/or trophies in museums, for example.

e) Social Networks: 19 of the 21 soccer club websites have accounts in various social networks. The websites of Arsenal and Manchester United were the only ones whose visit section did not offer this option, which is surprising, as they are major tourist attractions in their cities, with thousands of visitors each year. Social networks are already part of the daily lives of most internet users. In addition to individuals, many companies use this tool to publicize and promote their business, and this is no different with sports attractions.

5. CONCLUSIONS AND FINAL CONSIDERATIONS

With the technological revolution and the integration of its various tools into everyday life, the human being started to interact differently with his world, enjoying easier access to information and communication technologies, substantially changing his way of interacting with the market and other people. Besides providing access to detailed information, the use of online platforms includes auxiliary tools to facilitate the consumption of tourism services. Therefore, more independence is provided for travelers, who are seeking their own information and making their own decisions about destinations and services with less or no participation of travel intermediaries, according to Ayeh, et al. (2013).

Web 2.0 tools are a powerful way to obtain data directly from clients, mainly through questionnaires, with, in the case of football club sites, users being invited to fill out registration forms with specific demographic information about them, sports preferences, demands for products and services, feedback, among others. The benefits to football clubs are that they can track the access of the fans, the use of these facilities and the time spent. This, according to Paul (1996), helps football clubs build large user databases with rich and detailed marketing information, which helps to better understand the needs of web customers and product customization.

In the present study, among the five criteria involving static website activity, related to online purchasing and information provision about the sports attractions, most of the websites presented conclusive, clarifying options. A strength worth mentioning is the possibility to read the contents in other languages, which shows concern with foreign tourists. In contrast, a weakness to be highlighted is that 10 websites did not offer online sales, which can cause some rejection among those who wish to plan their trip.

With regard to user-website interaction, clubs should work harder to develop Web 2.0 tools in order to create a greater proximity to the public and attract their visit to sports venues. Among the Brazilian stadiums, more than half do not offer online sales or contents in other languages. It is worth remembering that in 2019 Brazil will host another mega-event, the Copa América, which brings together the continent's main national teams.

Therefore, investing in these tools can expand demand and attract national and foreign tourists, not only to watch matches. Revenue from using sports facilities on non-match days has become a powerful income source in clubs' finances. Real Madrid and Barcelona, for example, have attracted over 1 million visitors in one year to their tours and museums, making them highlights for those who visit their home cities (Chad 2013; Raupp, 2016). In the Brazilian case, after the construction of the new arenas for the 2014 World Cup, which have high quality structures, the venues can turn themselves into potential income generators through the attraction of new visitors.

It is recommended to boost virtual tours and social network sharing options, i.e. the visitors-website interaction, which are linked to the concept of smart tourism. The limitation of this study was the lack of direct information from users and visitors, thus restricting a more in-depth analysis on their actual experience with the website, and what criteria are most important to users. Our suggestion for further research is a qualitative-quantitative study involving in-depth interviews with sport managers and visitors or potential visitors, in real and in virtual world and to discuss the tools analyzed.

This information would allow a quantitative study with visitors to learn about their perception of the clubs'/stadiums' websites and to check whether or not they use the tools or find them relevant, in addition to collecting their suggestions to enhance the websites, with a view to improve the experience and increase satisfaction.

REFERENCES

- Angus, A. (2018). *Top 10 Global Consumer Trends for 2018 Emerging Forces Shaping Consumer Behaviour*. Euromonitor International.
- Antinucci, F. (2007). *Musei virtuali. Come non fare innovazione tecnologica*. [Virtual museums. How not to make technological innovations.] Roma: Laterza.
- Ayeh, J. K., Au, N. and Law, R. (2013). "Do we believe in TripAdvisor?" examining credibility perceptions and online travelers' attitude toward using user-generated content. *Journal of Travel Research*, Vol. 52, No. 4, pp. 437-452.
- Beritelli, P., Bieger, T., and Laesser, C. (2007). Destination governance: Using corporate governance theories as a foundation for effective destination management. *Journal of Travel Research*, Vol. 46, No. 1, pp. 96-107.
- Boes, K., Buhalis, D. and Inversini, A. (2015). Conceptualising Smart Tourism Destination Dimensions. Springer International Publishing Switzerland: *Information and Communication Technologies in Tourism*. pp. 391-403.
- Buhalis, D. and Amaranggana, A. (2014). *Smart Tourism Destinations*. In Xiang, Z. and Tussyadiah, I. (Eds.). Information and Communication Technologies in Tourism. Heidelberg, Germany: Springer.
- Buhalis, D. and Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet – The state of eTourism research. *Tourism Management*, Vol. 29, No. 4, pp. 609-623.
- Carvalhedo, A. (2003) "Tourism as a cultural legacy of the Olympic games". In M. De Moragas, C. Kennett and N. Puig (eds). The Legacy of the Olympic Games 1984-2000, 195-204. Lausanne: International Olympic Committee.
- Chade, J. (2013). Real enche estádio até quando não joga. O Estado de S.Paulo. Available: <http://economia.estadao.com.br/noticias/geral,real-enche-estadio-ate-quando-nao-joga-imp-,1028661> (Accessed 20 July 2017).
- Crouch, G. I. (2011). Destination Competitiveness: An Analysis of Determinant Attributes. *Journal of Travel Research*, Vol. 50, No. 1, pp. 27–45.
- Dawson, B. (2008). Facilitating innovation: Opportunity in times of change. *Museum Management and Curatorship*, Vol. 23, pp. 313-331.
- Flouri, E. and Buhalis, D. (2004). *Wireless Technologies for tourism destinations*. In: Information and communication Technologies in tourism 2004, the 11th ENTER Conference in Cairo, Egypt, pp. 27-38.
- Gil, A. C. (2007). *Como elaborar projetos de pesquisa*. 4. ed. São Paulo: Atlas.
- Guardia, S. R. R. and Mendes Filho, L. A. M. (2016). Estratégias necessárias para estruturar e desenvolver com competitividade Natal-RN como destino turístico inteligente. *Cultur*. Vol. 10, No. 2.
- Hays, S., Page, S. J. and Buhalis, D. (2012). Social media as a destination marketing tool: Its use by national tourism organisations. *Current Issues in Tourism*, pp. 1-29.
- Hassan, H. A. (2011). *Tecnologias de Informação e Turismo: e-tourism*. Dissertação (Mestrado em Lazer) – Departamento de Geografia e Turismo, Faculdade de Letras da Universidade de Coimbra.
- Hiller, M. (2011). *Uma experiência Estética e de Marketing em Museus*. Available: http://www.faculdadefundetec.com.br/img/revista_academica/pdf/artigo_hiller.pdf (Accessed 17 July 2017)

- Interbrand. (2017). Juventus Futebol Clube muda as regras do jogo e expande seu universo com a Interbrand. Available: <http://interbrand.com.br/newsroom/juventus-futebol-clube-muda-as-regras-do-jogo-e-expande-seu-universo-com-a-interbrand/> (Acessed 10 July 2017)
- Law, R. and Leung, R. (2000). A study of airlines' online reservation services on the internet. *Journal of Travel Research*, Vol. 39, No. 2.
- López, X., Margapoti, I., Maragliano, R. and Bove, G. (2010). The presence of Web 2.0 tools on museum websites: a comparative study between England, France, Spain, Italy and the USA, *Museum Management and Curatorship*, Vol. 25, No. 2, pp. 235-249.
- MacArthur, M. (2007). *Can museums allow online users to become participants?* Washington, DC: AAM.
- Miguéns, J., Baggio, R. and Costa, C. (2008). Social media and tourism destinations: TripAdvisor case study. *Advances in Tourism Research*, pp. 26-28.
- Muñoz, A. L. Á and Sánchez, S. G. (2013). Destino Turístico Inteligente. *Harvard Deusto Business Review*. 224, 58-67.
- Neuhofer, B., Buhalis, D. and Ladkin A. (2012). Conceptualising technology enhanced destination experiences. *Journal of Destination Marketing and Management*, Vol. 1, pp. 36-46.
- O'Reilly, T. (2005). *What is Web 2.0. Design patterns and business models for the next generation of software.* Available: <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- Paffrath, R. E. and Cassol, V. J. (2014). Gaming abroad: o uso e gamificação no projeto de um sistema para apoio a turistas. XIII SBGames. pp. 429-437.
- Prahala, C. K., & Ramaswamy, V. (2004). Co- creating unique value with customers. *Strategy & Leadership*, 32(3), 4–9. <https://doi.org/10.1108/10878570410699249>
- Paul, P. (1996). Marketing on the internet. *Journal of Consumer Marketing*, Vol. 13, pp. 27-39.
- Pezzi, E., and Vianna, S. L. G. (2015). A Experiência Turística e o Turismo de Experiência: um estudo sobre as dimensões da experiência memorável. *Turismo em Análise*, Vol. 26 No. 1, pp. 165-187.
- Raupp, I. (2016). Barcelona gera 3 R\$ 3 bi no ano e é o motor econômico e turístico da cidade. Available: <http://globoesporte.globo.com/futebol/futebol-internacional/futebol-espanhol/noticia/2016/12/barcelona-gera-r-3-bi-no-ano-e-e-motor-economico-e-turistico-da-cidade.html> (Acessed 26 June 2017).
- Ribeiro, N. A. I. (2014) *O Turismo Desportivo como estratégia de competitividade para destinos turísticos: o caso da região de Lisboa* (Dissertação de Mestrado, Instituto Superior de Gestão, Lisboa, Portugal) Available: <http://hdl.handle.net/10400.26/9425>
- Ross, S. D. (2001). Developing sports tourism. Illinois: University of Illinois at Urbana-Champaign.
- Sigala, M. (2013). A framework for designing and implementing effective online cupouns in tourism and hospitality. *Journal of Vacation Marketing*, Vol. 19, No. 2.
- Vicentini, I. C. and Hoppen, N. (2002). Tecnologia da informação aplicada aos negócios de Turismo no Brasil. *Turismo – Visão e Ação*, Vol. 5, No. 11, pp. 83-94.
- Wright, R. K. (2007). Planning for the Great Unknown: The Challenge of Promoting Spectator-driven Sports Event Tourism. *International Journal Of Tourism Research*, Vol. 9, No. 5, pp. 345-359.
- Turner, R. and Freiermuth, E. (2016). Travel and tourism: economic impacts 2016 world. World Travel and Tourism Council. Available: <https://www.wttc.org/-/media/files/reports/economic%20impact%20research/regions%202016/world2016.pdf> (Acessed 15 June 2017)
- Zhang, L. Y., Li, N., Liu, M. (2012). On the basic concept of smarter tourism and its theoretical system, *Tourism Tribune*, Vol. 27, No. 5, pp. 66-73.
- Zikmund, W. G. (2000) *Business research methods* (5a. ed.) Fort Worth, TX: Dryden.

HEART RATE VARIABILITY ANALYSIS AS A PREVENTIVE INSTRUMENT AGAINST STRESS FOR EMPLOYEES

Thomas Keller, Giovanni Annunzio and Elke Brucker-Kley
*Zurich University of Applied Sciences, School of Management and Law
St.-Georgen-Platz 2, 8401 Winterthur, Switzerland*

ABSTRACT

Disease-related absences due to mental illness have risen sharply in recent years. The aim of this work is to develop a key performance indicator system that can be used to describe the chronic stress burden on employees in a company. Based on the state of the art of heart rate variability (HRV) and stress medication, it was deductively concluded on possible indicator systems which were subsequently evaluated. A simple and composite system of indicators has been developed, which makes it possible to determine the tendency of adaptability (health and well-being) to external and internal influences of employees. This suggests a tendency towards chronic stress for employees. The key indicator systems are able to produce a valid final result without any user interaction. They provide information on how many employees show the corresponding tendency during the survey period. With the help of the indicator systems, companies can take appropriate measures to protect their employees from serious illnesses.

KEYWORDS

Heart Rate Variability, Stress, Indicator System, Prevention

1. INTRODUCTION

The measurement of heart rate variability (HRV) for the objective determination of the chronic (permanent) stress load of employees shows a high practical relevance, since stress is omnipresent in the workplace. The figures published by the DAK (2016) show that the number of sick leave due to mental illness has risen sharply in recent years. The study has shown that absenteeism has increased by more than 200% between 1997 and 2015. A permanent psychological stress at the workplace can lead to a mental illness, such as burnout (Berger et al., 2012: 1366).

This has negative consequences for the employee, but also for the employer. The aim of this work is to record the chronic (permanent) stress burden of employees in companies and to visualize it as a key performance indicator system. The aim is to reduce the work-related stress of the employees and thus optimize their work performance. This is made possible by providing companies with a tool to visualize the chronic stress on their employees. Using this tool, the company can influence the work-related stress of employees.

In order to be able to make a statement about the chronic stress burden on employees, it must be determined using a suitable measurement method. According to Mockenhaupt (2009: 354), chronic stress can be assessed using various methods, i.e., self-assessment, observations, tests or measurements.

For the present work, it was decided to collect the chronic stress load of employees on the basis of measurements. The measurement method can be easily implemented in a company. The employees are equipped with a measuring device and after a predefined period of measurement, the devices are returned and the data can be evaluated.

Since stress cannot be measured directly, it must be determined by so-called surrogate parameters (Haurand et al., 2015: 145). Stress influences various physiological parameters via the vegetative nervous system, such as heart rate or blood pressure (Haurand et al., 2015: 145). These changes are mediated by means of released endogenous hormones, such as cortisol (Haurand et al., 2015: 145). In order to measure the

values of the body's own hormones, saliva samples would have to be taken from the employee, which would then be evaluated. The authors question the practicability of this approach for a larger group of people. The samples would have to be taken at different intervals in order to show the temporal course of the stress values. For this reason, it was decided to collect heart rate variability as a physiological parameter in order to draw conclusions about the stress. HRV can be raised by a non-invasive procedure. For this purpose, employees must be equipped with a measuring device that can record an electrocardiogram (ECG).

Today, the literature mainly describes procedures in which HRV is evaluated for individual test persons via short-term (few minutes) or long-term measurement (24 hours). There do not seem to be any systems of identification numbers for companies to represent the chronic stress burden on entire groups of people.

HRV analysis is a very powerful tool. If the ECG is collected precisely, it can diagnose various diseases. In order to protect employees, it was decided that the key figure system should not allow conclusions to be drawn about individuals, but should only provide a statement about the stress level of the company in general. In the event of a possible implementation of the KPI model in the company, the data protection of the employees must be guaranteed.

2. METHODOLOGY

This work presents a simple and a composite key figure system model for discussion. These are based on the methods of HRV. From the analysis of the literature it could be determined that there are no uniform standards for the implementation of the entire HRV analysis process. This inevitably leads to difficulties when comparing different studies or testing the validity of the results obtained. For this reason, it was decided to close this gap by documenting the process used, the methods used and their parameters for HRV analysis. This allows independent verification of the results after completion of the work.

2.1 Research Process for HRV Analysis

A simple model was created for this purpose. The model is intended to depict the research process for HRV analysis from an abstract perspective. The research process was based on the following questions:

- Which process must be used for an HRV analysis?
- Which methods with which parameters must be used?

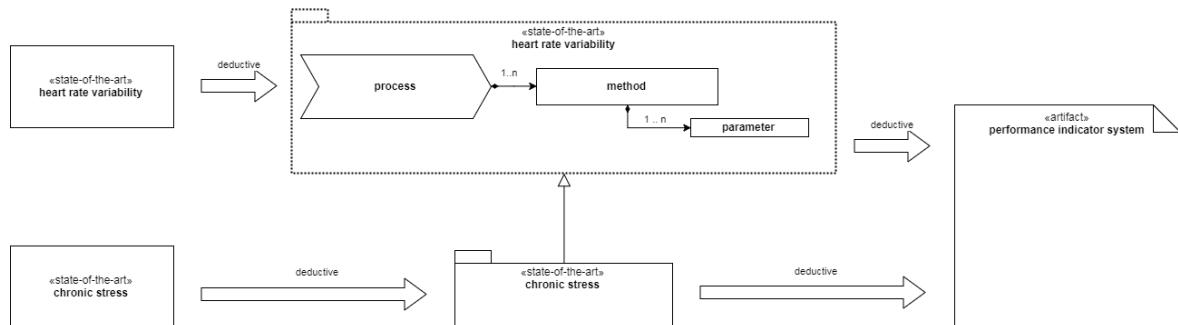


Figure 1. Research process of the investigation

This resulted in the overview shown in Figure 1. The basis is the current state of the art of HRV and chronic stress literature. Starting from the HRV literature has been deductively concluded on the standard process, the standard methods and their parameters of today's state of the art (Annunzio, 2017: 20 ff). The same was done for the state of the art of chronic stress literature (Annunzio, 2017: 41 ff). In this step, only those methods that have a reference to the HRV analysis (subset of the HRV methods) were taken into account. Based on this abundance of methods, the system of indicators was concluded (Annunzio, 2017: 59).

2.2 The Process of HRV Analysis

The research community agrees that the guidelines of the Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology (abbreviation: Task Force) represented by Malik et al. (1996) define the standard process of HRV analysis (see: Figure 2).

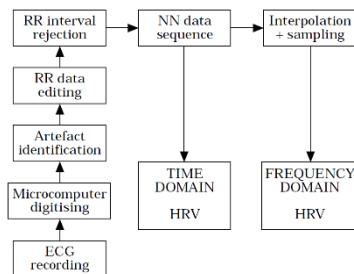


Figure 2. Steps in the process Analysis of HRV (Malik et al., 1996: 365)

The process steps from Figure 2 of Malik et al. (1996: 365) can be divided into the following steps:

1. Recording of the data: The ECG signal is recorded
2. Preparation of the data: The R-Peaks are detected, the artifacts are detected and cleaned.
Note: In order to calculate the HRV parameters from the frequency domain, the data (R-peak) must first be interpolated and sampled again (data points on the time axis are not equidistant).
3. Evaluation of the data: The calculation of the HRV parameters (time/non-linear range)
4. Interpretation of the data: Interpretation of HRV parameters (time / non-linear range)

Note: The steps described above apply only to the calculation of the HRV parameters from the time and non-linear range.

2.3 Methods and Parameters of HRV Analysis Methods

Following are the four process steps further explained.

2.3.1 Recording of Data

The ECG signal is recorded with a recording device. The central question in this section is which parameters must be considered when choosing the recording device. The authors Wittling & Wittling (2012: 146 ff) identified the following three:

1. Sampling of the ECG signal: The sampling must be selected in such a way that the sampling theorem of Nyquist-Shannon is not violated. A sampling frequency of 500-600Hz is considered a good compromise between sampling accuracy and data volume by the authors Wittling & Wittling (2012: 148), as higher sampling frequencies (1000-4000Hz) do not contribute to a better HRV analysis result.
2. Quantization of the ECG signal: The assignment of a measuring point (sampling point) to a discrete binary value is called quantization. Wittling & Wittling (2012: 147) showed that even with a low resolution of 8-bit, there is no relevant impact on the HRV analysis results.
3. Digitisation of the ECG signal: From a metrological point of view, the step is an analog-to-digital conversion which is carried out in the recording devices by the A/D converter (Wittling & Wittling, 2012: 148).

Conclusion: The authors Wittling & Wittling (2012) describe the data recording in detail using various examples and were thus able to explain their results plausibly and comprehensibly.

2.3.2 Processing of Data

The raw data recorded by the recorder must be processed. The RR intervals form the basis for calculating the HRV parameters. They represent the time elapsed between two heartbeats. In the ECG signal, this is the time difference between two consecutive highest points. These highest points are called R-peak in technical

jargon. The RR intervals must be identified from the raw data, checked for plausibility (artifact yes/no?) and corrected if necessary. The data preparation can be divided into the following steps:

1. R-peak detection: In this step, the R-peak must be reliably detected in the ECG signal. Peltola (2012: 3) lists some algorithms based on Hilbert transformation, digital filters, pattern recognition and wavelet transform, but he notes critically that there is no standard in the literature.
2. Artifact detection: Artifacts may occur during recording. There are two types of artifacts: physiological and other artifacts. The physiological artifacts are arrhythmias of the heart and can be traced back to deviations from the normal action sequence of the heart. The remaining artifacts are either caused by the measuring device or by the displacement of the measuring sensors relative to the skin or to the heart (Hoos et al., 2006: 189). The consequence of this is that the R-peak detection misses heartbeats, e. g., missing beats or disturbances on the ECG signal are detected as R-peak, e. g., additional beats (Hoos et al., 2006: 189) and this again artificially increases the variability and thus improves it. In the literature no standard regarding artifact detection could be found, often the algorithm was named by Berntson et al. (1990), but without going into the details.
3. Artifact handling: If an artifact has been found, you have to decide which strategy to choose. When applying the strategy, it is important to be aware that the choice of strategy has an influence on the HRV analysis result. Peltola (2012: 7) concluded in his analysis that there is no consensus in the research community as to which strategy should be applied and when. The artifact content of the data must also be kept in mind, as Wittling & Wittling (2012: 165) make the following recommendation: Since artefact loads of more than 10% should not be interpreted or excluded from the study (in scientific studies <2% and in practice <5%). The following strategies were found in the analysis of the literature:
 - Ignore it: The artifacts from the data are ignored and included 1:1 in the calculations. This strategy can artificially increase the variability (missing/additional beats).
 - Remove: The artifacts are removed from the data. Important information can be lost and the database becomes smaller.
 - Correct it: The artifacts from the data are replaced by one or more new values. This is often done by means of an interpolation algorithm (linear, cubic, non-linear). Kaufmann et al. (2011: 1162) believe that linear and cubic interpolations are still the standard for HRV analysis.

Conclusion: In summary, the recommendation of Peltola (2012: 8) can be passed on here. He advises that further studies in this area are necessary.

Note: Physiological artifacts can also occur in healthy people (Peltola, 2012: 3). The HRV analysis is based on the electrocardiogram (ECG), which contains the excitation image of the heart. In order to be able to assess the function of the autonomic nervous system (ANS) with the help of HRV, it is assumed that the RR intervals (= distance between two heartbeats) on which the calculation is based are derived from artifact-free QRS complexes of a sinus rhythm (=normal heartbeat) (Hoos et al., 2006: 188).

2.3.3 Analysis of Data

The evaluation of data consists of two steps, the validation of the input data and the calculation of the HRV parameters:

1. Validation: The input data must be validated in order to answer the question of whether they may be used for further calculations. To calculate the HRV parameters, which are based on statistical methods (subset of HRV parameters from the time domain), the data must be normally distributed. Kaufmann et al. (2011: 1165) used the Kolmogorov-Smirnov test (NIST/SEMATECH, 2012) to check the data for their normal distribution.
2. Calculation: For the quantification of HRV, methods from the time and frequency range and non-linear analyses can be used (Sammito et al., 2014: 12). Sammito et al. (2014: 15–18) provide an overview of the established HRV parameters. In addition to each HRV parameter, they specify which indicator is used, e. g.: it is an indicator of the overall variability, short-/long-term variability, and so on. Furthermore, they provide a recommendation for the HRV parameters for the individual measurement duration.

Conclusion: For validation, the methods can be derived from the calculation methods used, but these must be disclosed in order to enable a comparison. Many HRV parameters are available for the calculation and these seem to be uniformly documented in the literature.

2.3.4 Interpretation of Data

The HRV parameters only provide a benefit if they can be put into a context. Thus, the interpretability of HRV parameters is a central point for practicability. The challenge is to interpret the HRV parameters correctly, because the HRV parameters are the result of all external and internal influences on the person (Sammito & Böckelmann, 2015: 82). The following factors influence HRV: alcohol, breathing, fitness, gender, cardiovascular disease, heat, cold, cold, body fat/body weight, noise, age, age, psychiatric illnesses, smoking, harmful substances, shift work with night shift, metabolic diseases, stress/mental tension, circadian rhythm / time of day (Sammito et al., 2014: 19 ff).

The Task Force of Malik et al. (1996: 380) has set standard values, but they do not go into the influencing factors in more detail. Today it is known that the variability is highest in younger years and decreases with age (Sammito et al., 2014: 21). In the book by Eller-Berndl (2015: 152–153) we refer to the HRV score of Buchhorn, which takes age into account. A further complication is that each HRV parameter must be evaluated separately, e. g. the HRV parameter SDNN may only be compared with the same measurement duration and similar activities (Eller-Berndl, 2015: 29). The higher the activity level, the higher the resulting SDNN (Eller-Berndl, 2015: 29).

Conclusion: In some cases, reference values or standard values for HRV analysis can be found in the literature. However, there are still no guidelines on how strongly (quantitative) the influencing factors affect HRV and how they influence each other.

2.4 A homogeneous Template for a HRV Analysis

In order to compare future studies, the template from Figure 3 is proposed.

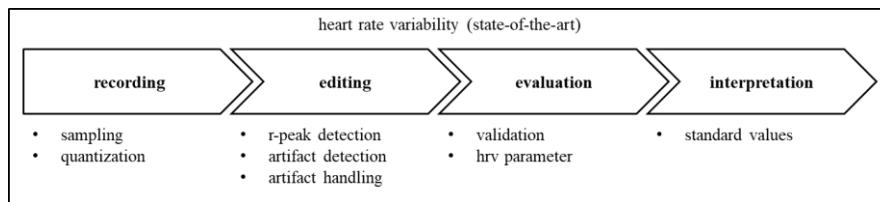


Figure 3. HRV Analysis template

The scientific papers can be compared more easily if the reader is told which methods and parameters have been used in their investigations.

Table 1. Checklist for a homogeneous HRV Analysis

Process step	Description
Recording	What sampling frequency was selected? Which quantization was chosen?
Processing	Which R-peak detection algorithm, with which parameters, was used? Why was this algorithm chosen and not another one? Which artifact detection algorithm, with which parameters, was used? Why was this algorithm chosen and not another one? Which artifact handling strategy, with which parameters, was used? Why was this strategy chosen and not another?
Analysis	Which method was used to validate the input data? Why was this method chosen and not another? Which HRV parameters have been calculated? Why were they calculated and not others?
Interpretation	Which reference/standard values were consulted for the interpretation? Why were these reference/standard values consulted and not others?

Table 1 lists the items for a standardization of HRV analysis. The authors should disclose the methods used (incl. parameters) and why they used this method. This should help to compare simple studies in the future. The research community must also agree on which methods will be chosen as the standard, so that as many analyses as possible can be carried out using the same methods. This will ensure that the studies can be compared with each other in the future.

3. PROPOSED KEY INDICATOR SYSTEMS

Every employee has received a recording device (chest belt) from the company, which continuously sends the data to the server that processes it. On weekends, the data is stored temporarily on the recorder and is not synchronized until the next company entry. The model of the simple key figure system is illustrated in Figure 4. The trend for each individual employee is calculated every day at the same time.

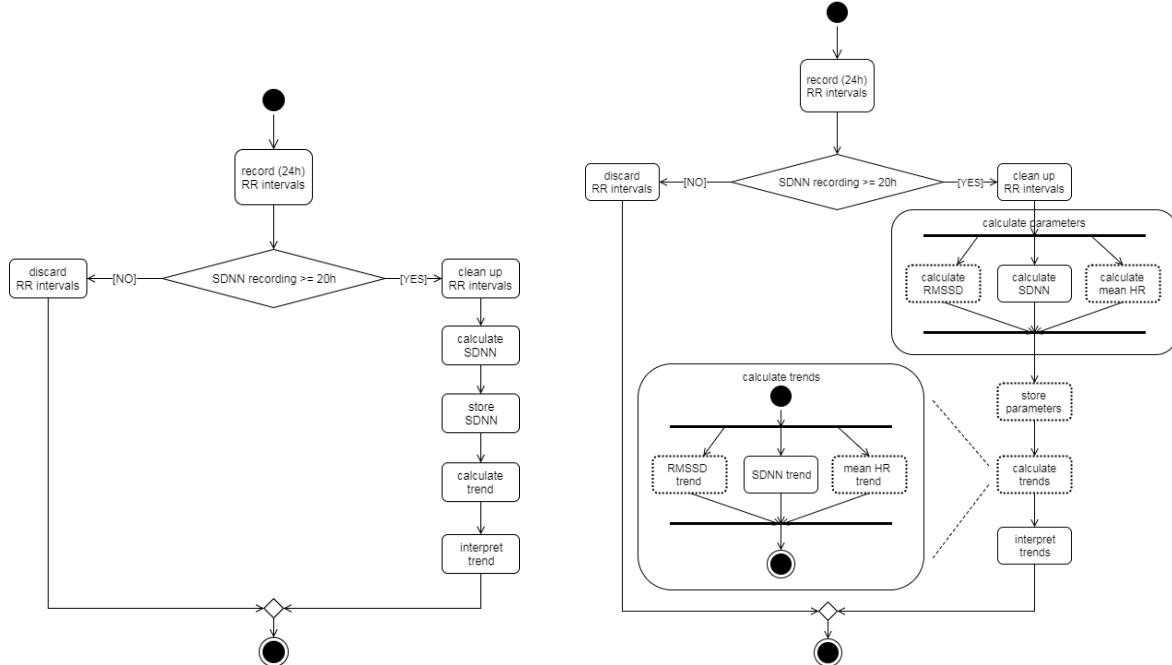


Figure 4. Simple (left) and composite (right) key figure system

The trend analysis for the simple approach consists of the following steps. In the first step, the data is analyzed for completeness after one day (24h). Completeness is given if the RR intervals have been recorded for at least 20 hours, regardless of whether interruptions have occurred in between. If the completeness is not given, the RR intervals are discarded and a data gap arises for this day. If completeness is given, the data are corrected (artifact extraction and artifact cleaning) and checked for quality (normal distribution and artifact content). Records that are not normally distributed or that have an artifact infestation of more than 5% are excluded. The SDNN is calculated from the cleaned data and stored in the next step. This results in an SDNN value for every employee over the course of time. This SDNN time series is the basis for a trend analysis. By means of linear regression (a straight line is laid through the time series), the trend (corresponds to the gradient of the straight line) can be read for each employee. The interpretation of the data takes place via the trend. The trend provides information as to whether the variability improves (positive slope of the trend line) or worsens (negative slope of the trend line).

For the more precise determination of chronic stress, two additional HRV parameters have been added to the simple KPI system: RMSSD and the mean heart rate (HR) (both highlighted with dotted boxes in Figure 4). The steps "store parameters" and "calculate trends" also had to be adapted. Two additional HRV parameters are now required for trend determination. In order to speak of a chronic stress trend, the conditions B1-B3 [decreasing SDNN==TRUE] AND [sinking RMSSD==TRUE] AND [increasing mean HR==TRUE] must be fulfilled at the same time.

4. DISCUSSION

The composite performance indicator system was tested with a limited test data set of seven employees for one company. To confirm the results, further tests with different test data are required. Among other things, it

would be necessary to check the data of test persons who have been diagnosed with chronic stress by a specialist. This could be used to determine the recognition rate of the composite performance indicator system.

The measurement duration of 15 days was too short, not least due to the poor data quality. Due to the volatility of the data, especially on weekends, it takes at least 14 days of continuous data recording. It must also be further investigated what the optimal measurement duration is in order to be able to make a valid statement about the chronic stress load.

The recording duration was chosen for practicability reasons shorter than 24 hours. In order to meet 100% of the requirements of the HRV analysis, the data must be available for the entire 24h. Manufacturers of recording devices are required to develop their equipment more stable against interruptions and with a higher wearing comfort.

The size of the groups to be examined must be investigated. For large groups, it is difficult to initiate measures that will also reach the people affected. For a medium-sized company with more than 200 employees, it would make sense to group the participants by department. In this way, measures can be initiated specifically for these groups specifically. For example, if assembly workers in day and night shifts have been working overtime for months, it makes sense that this group is shown in isolation in the performance indicator system so that a statement can be made about their chronic stress load. Otherwise, it is only possible to order company-wide measures and hope that those affected can benefit.

Further HRV parameters are required for more precise recognition of chronic stress. In this respect, a valid HRV measure for the determination of sympathetic nervous system activation is to be mentioned, which should be investigated in the future.

The participant's behaviour is influenced by measuring HRV. We must always bear in mind what incentives are created and how they are realised by employees. This should be reviewed in a wide-ranging study. There were changes in employee behaviour during the survey. Negative influences - apart from the discussions with the wife regarding the activated radio link in the bedroom - could not be determined.

The performance indicator system only unfolds their full benefits when they are embedded in a company. The benefit for a company can only be confirmed when the entire control loop has been activated. The control loop means that the management takes measures based on the results of the performance indicator systems and, after a certain period of time, carries out the measurements again to see whether the measures have been successful - a kind of "health benefit controlling".

5. CONCLUSION

The composite performance indicator system makes it possible to identify the tendency of chronic stress (using HRV parameters: SDNN, RMSSD and mean heart rate) of the employees on the basis of HRV analysis. It also offers the possibility of quantifying the number of employees affected. The performance indicator system does not require any human input and operate completely objectively. The system does not need to be calibrated by the wearer as it is based on changes in HRV parameters and does not work with absolute values.

The work has shown that the combination of already existing technologies and methods can be used to develop simple performance indicator systems. These systems of key performance indicators can be of great benefit to companies, as statements can be made about the current stress level of employees, on which suitable measures can be derived and implemented. Before a potential key performance indicator system can be implemented in the structure of a company, further clarifications are necessary. For example, the performance indicator system would have to be tested with further data, the optimum measurement period for determining trends and the optimum group size would have to be defined.

REFERENCES

- Annunzio, G. (2017). Die HRV-Analyse als Präventionsinstrument gegen Stress der Mitarbeitenden im Unternehmen (nicht publiziert). ZHAW School of Management and Law.
- Berger, M., Schneller, C., & Maier, W. (2012). Arbeit, psychische Erkrankungen und Burn-out: Konzepte und Entwicklungen in Diagnostik, Prävention und Therapie. *Der Nervenarzt*, 83(11), 1364–1372. <http://doi.org/10.1007/s00115-012-3582-x>
- Berntson, G. G., Quigley, K. S., Jang, J. F., & Boysen, S. T. (1990). An Approach to Artifact Identification: Application to Heart Period Data. *Psychophysiology*, 27(5), 586–598. <http://doi.org/10.1111/j.1469-8986.1990.tb01982.x>
- DAK. (2016, März). Arbeitsunfähigkeitsstage aufgrund psychischer Erkrankungen in Deutschland nach Geschlecht in den Jahren 1997 bis 2015. Abgerufen 19. November 2016, von <https://de.statista.com/statistik/daten/studie/254194/umfrage/au-tage-aufgrund-psychischer-erkrankungen-in-deutschland-nach-geschlecht/>
- Eller-Berndl, D. (2015). *Herzratenvariabilität*. Wien: Verlagshaus der Ärzte.
- Haurand, C., Ullrich, H., & Weniger, M. (2015). *Stressmedizin: Beratung, vorbeugung, behandlung*. Abgerufen von <http://site.ebrary.com/id/11015045>
- Hoos, O., Gläser, S., & Sommer, H.-M. (2006). Untersuchung zur Ursache von Artefakten in der mobilen Erfassung der Herzfrequenzvariabilität bei Laufbelastungen. Deutsche Vereinigung für Sportwissenschaft (Hrsg.) (Bd. 162, 188–196). Hamburg: Czwalina.
- Kaufmann, T., Sütterlin, S., Schulz, S. M., & Vögele, C. (2011). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. *Behavior Research Methods*, 43(4), 1161–1170. <http://doi.org/10.3758/s13428-011-0107-7>
- Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17(3), 354–381. <http://doi.org/10.1093/oxfordjournals.eurheartj.a014868>
- Mockenhaupt, J. (2009). Herzfrequenzvariabilität als Indikator mentaler Anspannung. In C. Zacharias, K. W. ter Horst, K.-U. Witt, V. Sommer, M. Ant, U. Essmann, & L. Mülheims (Hrsg.), *Forschungsspitzen und Spitzforschung* (353–363). Heidelberg: Physica. Abgerufen von http://link.springer.com/10.1007/978-3-7908-2127-7_31
- NIST/SEMATECH. (2012, April). Kolmogorov-Smirnov Goodness-of-Fit Test - NIST/SEMATECH e-Handbook of Statistical Methods. Abgerufen 1. Mai 2017, von <http://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm>
- Peltola, M. A. (2012). Role of editing of R–R intervals in the analysis of heart rate variability. *Frontiers in Physiology*, 3. <http://doi.org/10.3389/fphys.2012.00148>
- Sammito, S., & Böckelmann, I. (2015). Analyse der Herzfrequenzvariabilität: Mathematische Basis und praktische Anwendung. *Herz*, 40(S1), 76–84. <http://doi.org/10.1007/s00059-014-4145-7>
- Sammito, S., Thielmann, B., Seibt, R., Klüssmann, A., Weippert, M., & Böckelmann, I. (2014). AWMF Leitlinie: Nutzung der Herzschlagfrequenz und der Herzfrequenzvariabilität in der Arbeitsmedizin und der Arbeitswissenschaft. <http://doi.org/10.13140/2.1.5052.7684>
- Wittling, W., & Wittling, R. A. (2012). *Herzschlagvariabilität: Frühwarnsystem, Stress- und Fitnessindikator*. Heilbad Heiligenstadt: Eichsfeld.

TRUSTING IN THE SECURITY OF PRIVATE MESSAGES – A SURVEY-BASED ANALYSIS OF FINNISH INTERNET USERS

Jukka Vuorinen, Aki Koivula and Ilkka Koiranen
Economic Sociology, Department of Social Research, University of Turku, Finland

ABSTRACT

This paper analyzes the extent to which Internet users trust that the privacy of their private messages is protected and upheld in web-based messaging services such as Facebook Messenger and WhatsApp. The paper examines how users' trust for other people (social trust) and trust in institutional actors (such as the parliamentary system and police) predict a high probability to trust that their private and confidential messages are not disclosed. The analysis is conducted based on data ($n=1648$) that was gathered through a nationwide survey in Finland. Results show that high social trust is associated with high trust in the confidentiality and privacy of messages that are sent through messaging services on the web. Social trust is the strongest predictor even when controlling for institutional trust and demographic factors.

KEYWORDS

Privacy, Trust, Social media, Messaging, Communication

1. INTRODUCTION

Various social Internet platforms have vastly improved and extended opportunities for communication and interaction. For some, social media is a platform for affecting the opinions of thousands or seeking validation through modifying visual self-presentation and taste expression, while for others it is used for routine relationship maintenance and interaction (e.g. Keipi et al., 2018). In all cases, new communication technologies and services have vastly changed the ways of how people interact socially. This has generated a great deal of pressure towards ensuring the reliability, confidentiality, and integrity of these platforms.

Simultaneously, while a growing share of the population has embraced these novel means of social interaction, the societal significance of social media platforms has rocketed. Nowadays companies providing these new spaces of interaction mediate the most of the messages that are sent through the Internet. For example, Facebook Messenger and WhatsApp combined are mediating 60 billion messages per day (Tynan, 2016). This also means a great burden of responsibility for these social media giants in terms of securing the privacy of messages. In addition to the service provider, there is also another actor that can violate privacy, namely the receiver(s) of message. Other participants are able to redistribute sensitive content to third parties. Furthermore, third parties such as hackers, can get direct access to private content in question, which also presents a significant risk. In this respect, there are several actors who need to be trusted before private interaction can be established in a secure way.

Fundamentally, we are interested in users that communicate privately over the web using different messaging and communication services. Here, we do not solely refer to social media platforms, but also to more conventional means of interaction such as emails, discussion forums and chat messaging. Moreover, we are curious about users' trust in the strength of message privacy and confidentiality, i.e. how confident users are that this kind of messaging will stay private. As such, we ask the following research question:

1. How confident (i.e. trusting) are users that their confidential and private messages are kept private?

Furthermore, we compare the trust in message privacy to other forms of trust, namely institutional trust and social trust. Briefly, institutional trust refers to how citizens trust their governmental actors such as the parliamentary system and police whereas social trust designates general trust placed in other individuals.

On the one hand, these two forms of trust, social and institutional, have drawn significant attention in the field of social sciences. It is suggested that at the micro-level social trust promotes, for example, positive effects on happiness (Inglehart, 1999), health (Subramanian et al., 2002) and ethnic tolerance (Dinesen and Sønderskov, 2013). At the macro-level, nations reporting high levels of trust also tend to perform better in terms of social equality, levels of corruption and economic growth (Bjørnskov, 2012; Rothstein and Uslaner, 2005). Many studies have suggested that low corruption and trust in fairness of institutions are still essential in the accumulation of social trust (e.g. Dinesen and Sønderskov, 2013; Nannestad et al., 2014; Rothstein and Stolle, 2008). Overall, it can be argued that these forms of trust are crucial determinants in all relations making daily life easier and helping to meet desired expectations (e.g. Coleman, 1988; Möllering, 2006; Rothstein, 2005; Freitag and Traunmüller, 2009).

On the other hand, privacy issues have been a common concern in the field of information system sciences (e.g. Gross and Acquisti, 2005; Ellison, 2007). However, in these cases, privacy is mainly a concern at the platform level (how much a system gathers private information and information is exploited). Thus, it is important to examine trust in message privacy and compare it to institutional trust and general social trust in order to better gauge these contributors to users' wellbeing. In other words, we hope to show how social and institutional trust is associated with the perceived confidentiality of private messaging in digital media. We explore whether social and institutional trust are apparent in predicting citizens' concern about Internet security by asking:

2. To what extent does high social and institutional trust predict individuals' high trust in protection of the secrecy of their private messages (i.e. respondents trust that their messages will not go public)

In previous studies, the use of web services such as digital media has been shown to be linked to various contextual factors. For example, it has been shown that younger, highly educated and those living in urban areas are more likely to use social media compared to others (e.g., Schradie, 2011; van Deursen and van Dijk, 2014). In this respect, these active users have a better understanding and more experience on which trust (or distrust) can be formed (van Dijk 2005, 71-93). In addition, educated people tend to report higher trust than others (e.g., Helliwel and Putnam, 2007). Together, these results spur us to assume that the associations between trust dimensions (social and institutional) and trust in the confidentiality of private messages are partly dependent on the demographic factors of users. Accordingly, we also ask:

3. To what extent is the interplay of trust dimensions determined by demographic factors?

Before going into empirical testing, we present the framework of the study by explaining essential concepts and formulating a theoretical link between the different forms of trust. We then describe the data and methods of analysis. Finally, after the results section, we discuss the implications of our findings by considering potential limitations and further research avenues.

2. TRUST, EXPECTATIONS AND CONFIDENTIALITY OF PRIVATE MESSAGES

Messages that are sent through social message services can be assessed in terms of security through the three dominating concepts in the field, namely confidentiality, integrity, and availability (Vuorinen and Tetri, 2012). More specifically, confidentiality refers to actors who are allowed to access a specific piece of information (or a web based service). Integrity pertains to the order of the message and how it remains the same through the process of sending and receiving. For example, the order of letters should not mix in the process but stay exactly the same. Availability, in simple terms, means that the service (and thus the message) should be accessible when needed. In terms of our study, the dimension of confidentiality is the most significant as we are interested in how confident respondents are that their messages will remain private. In other words, the message will be seen solely by the receiver(s) and the message will not go public.

Trust can be a challenging property to measure directly in the context of socio-techno-material assemblages because “trust” is a strong social notion that is often related to a belief that someone (instead of something) will do something. As Ermisch et al. (2009) put it, “We trust when we trust that someone will do X—repay a loan; arrive on time; play fair; pay the fare; feed the cat; treat baby well; do his job as expected. The trust that we have in someone doing X does not necessarily extend to trust in that same person doing Y.” Here, trust is a human bonding word. Although we do not consider that agency would be a solely human property (Latour, 2005) we find it more relevant to approach trust through expectations (Mayer et al., 1995; Blomqvist, 1997; Rousseau et al., 1998; McEvily et al. 2003).

Sociologist Erving Goffman (1986) argues that every time an individual enters a new situation, they have to answer the question of “what is going on here?”. So, the answer can be “it rains”, “a child is playing”, “a computer crashes”. However, the frame does not merely cover the present but there is also a teleological dimension to it. This teleological frame is an expectation of what will happen, what will follow. For example, if one becomes surprised, it means that something unexpected took place. Expectation is also a teleological frame of trust; it designates how something X is believed to happen. Expectation is trust in terms of certainty (see Latour, 2005). In relation to technical systems, expectation can be used as a subtle way to measure trust. “I expect my messages to stay private (confidential)” is a teleological frame that refers to what is believed to happen and implies trust/non-trust.

In social action trust is not only based on social trust between the actors, but also on wider forms of trust (Lewis & Weigert, 1985, 973). Institutional rules, social norms, contracts and laws are defining what is legitimate and thus what outcome is expected (Rousseau et al., 1998, 400). In this sense, societal structures can be seen as important sources of trust that are streamlining all social processes and actions occurring in societal contexts. Accordingly, trust can be comprehended as feature adopted through socialization and additionally quality of the societal context itself (Kouvo, 2014, 22).

The most recent concerns have been raised about institutional privacy especially in terms of the use of personal data by institutions. This is relevant for policy discussions, as it firstly suggests that the collection, aggregation, and utilization of personal data for targeted advertisement has become an accepted social norm (Young and Quan-Haase, 2013). Secondly, different platform companies’, such as Google and Facebook, potential societal power has substantially accelerated during this decade. Nowadays big platform companies possess and process private information and messages from billions of people. In this respect, these platform providers have a special status in the discussion about privacy issues in the social Internet.

3. METHODOLOGY

The survey was based on random samples (N=4 001) of Finns aged 18 to 84 gathered from the Finnish population register database. The final data with a total of 1 648 respondents represent the Finnish population relatively well, but as is typical in survey research, there were age and gender related biases. In order to correct these biases, the data were weighted to meet the age and gender distribution of the Finnish population aged 18 to 84 years. As this study focuses on web-based messaging services, we have excluded the respondents who did not use the Internet at all from the analysis. Thus, our sample consists of 1452 respondents representing 88 percent of the total population which is exactly the same share as that of 16-89-year-old Internet users according to the most recent statistics (OSF 2017).

We use a variable elicited from the question “I worry that my personal messages will be revealed publicly without my consent” as a dependent variable. Answer options were given as a five point Likert scale from 1 ”Totally disagree” to 5 ”Totally agree”. At the beginning of the analysis, we turned the initial scale to measure how confident users are that their private message is kept in secrecy. In the multivariable analyses, we focused on those having high trust by transforming the variable into a binary by recoding the initial values into two categories as follows: 0 was given to those having either low or medium confidence (1 through 3 on the scale) and 1 was given to those having high confidence (4 through 5 on the scale).

Our main independent variables are social and institutional trust which were both combined from three single variables. Social trust was asked with similar questions used in European Social Surveys: “Most people can be trusted”, “Most of the time people are helpful” and “Most people try to be fair”. Institutional trust was based on respondents’ trust in the parliament, trust in the legal system and trust in the police. Regarding both trust dimensions, respondents were asked to report a score of 1-5 how much they personally

trust each of the institutions or how much they trust other people. In this sense, as we use mean variables instead of sum variables, 1 means respondents do not trust at all, and 5 means they have complete trust.

Our control variables included gender, age, education, domicile and Internet use frequencies. Age was measured in years and included as a continuous variable in the analysis. Education was recoded into four categories according to the ISCED classification. Residential area was used by following the NUTS 2 categorization to control the regional level effects. Finally, we controlled for respondents' subjective assessments of Internet and social media use frequency by standardizing on whether the user uses the Internet and social media several hours a day, daily or not more often than weekly. A descriptive overview of applied variables is shown in the Table 1. Mean and standard deviations are calculated only for internet users who had valid scores on dependent variable.

Table 1. Descriptive statistics for dependent and independent variables

	Obs.	Mean	Std. Dev.	Min	Max
<i>Dependent Variable</i>					
Confidence	1,393	3,0	1,3	1	5
<i>Independent variables</i>					
Institutional trust	1,392	3,6	0,8	1	5
Social trust	1,380	3,6	0,7	1	5
<i>Controllers</i>					
Gender	1,393	1,5	0,5	1	2
Age	1,393	45,7	16,2	18	83
Education (4 cat.)	1,393	2,6	0,9	1	4
Residence (4 cat.)	1,346	2,5	1,2	1	5
Internet use frequency (3 cat.)	1,393	2,2	0,7	1	3
Social media use frequency (3 cat.)	1,387	1,6	0,7	1	3

In order to find the extent to which social and institutional trust predict confidence, we used logit models in addition to descriptive methods. For the sake of clarity, we post-estimated the logit models and presented the main results as predicted probabilities in easily interpretative figures by utilizing the user-written packages developed in the Stata program (Jann, 2014; Bischoff, 2017). In order to do robust modelling in terms of nonlinear logit- models, we also use the khb-method developed by Karlson, Holm, and Breen (2012). The khb-method provides us with detailed information concerning the controlling process as it decomposes the total effect of the independent variables into direct, total and indirect effects. A direct effect may be found if the independent variable has an effect on the dependent variable by itself without that effect being mediated by some intermediary variable. For example, social trust may have a direct effect on the level of confidence but the effect can also be an indirect one through institutional trust or education.

4. RESULTS

We began the analysis by defining how confident Internet users are generally that their confidential and private message are kept in secrecy. The results of analysis are shown in the Figure 1. We can see that confidence is distributed quite evenly throughout the population. Approximately 35 percent of Finnish citizens reported low or not at all confidence. A slightly higher proportion (39 %) of respondents reported to being highly or completely confident.

An evenly distributed variable gives us a prolific starting point from which to conduct multivariable analysis. As noted before, we estimated to what extent social and institutional trust predict high confidence by combining those reported high and completely confident. As seen in Figure 2, the likelihood of confidence increased significantly according to the level of both trust dimensions. The average partial effect of social trust was 7.8 percent ($p<0.001$) and the detailed analysis revealed that the difference between the lowest and the highest deciles was approximately 19.7 percentage points ($p<0.001$). Interestingly, the effect of

institutional trust was not equally strong when compared to social trust. The average partial effect of institutional trust was 6.1 percent ($p<0.01$) as the marginal effect between the lowest and the highest deciles was 15.1 percentage points ($p<0.01$).

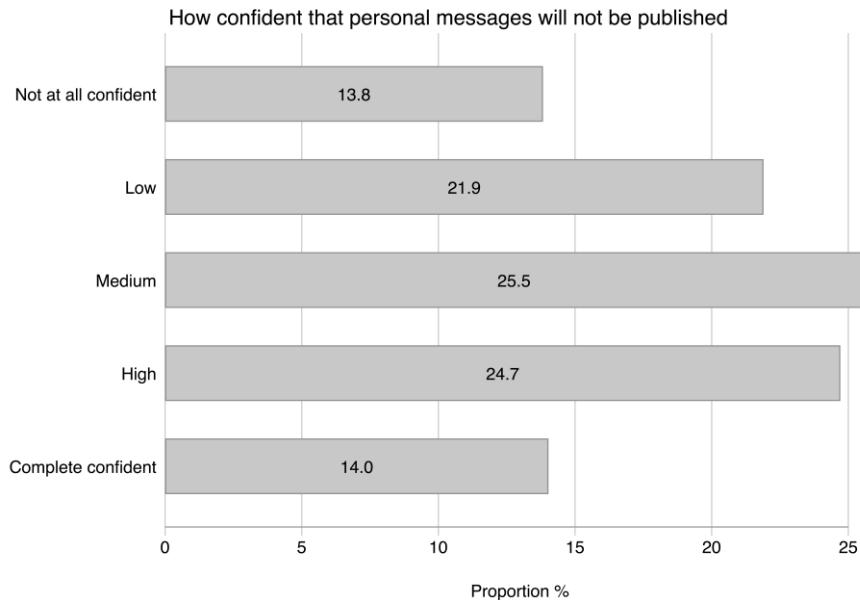


Figure 1. Finnish Internet users' confidence in that their private messages are kept in secrecy, proportions

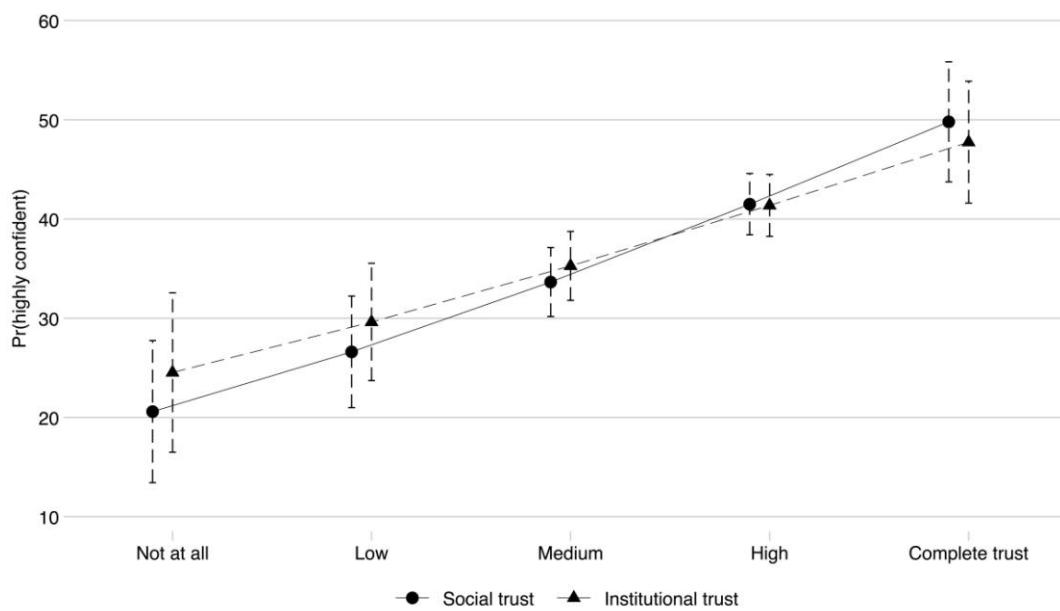


Figure 2. Predicting individual high trust in up-keeping of the secrecy of private messages according to social and institutional trust, logit estimated probabilities

We begin the next section by assessing the association between demographic factors and confidence. The results of analysis are presented in Figure 3. Notably, in contrast to our assumptions, the effects of demographic factors were extremely weak. However, according to findings, confidence was to some extent dependent on gender and age, since men and younger participants had a higher probability to be confident of their message privacy. Interestingly, education and domicile do not have any significant effect on confidence. Instead, it seems that those using the Internet several times a day (i.e. hourly) are more confident when compared those using only on a daily or weekly basis. Instead, frequency of social media use has no significant association with confidence.

After descriptive analysis, we also tested whether demographic factors mediate or confound the association revealed in Figure 2 by decomposing the effect between both forms of trust according demographic factors. First, the interplay between trust and confidence on private messages staying private was not explained or mediated by demographic factors. As a matter of fact, we found only one significant indirect effect between age and institutional trust. The detailed analysis of this association revealed that age did not mediate the effect of institutional trust, but rather it strengthened it.

In the final section of analysis, we tested the extent to which social and / or institutional trust mediated one another's effects. The most striking result here was that it is social trust that weakened the impact of institutional trust. Institutional trust does not weaken the impact of social trust. This was confirmed according to the indirect effects presented on the right-most columns of each figure. Here, we can see that the effect of institutional trust was insignificant ($b=0.064$; $p=0.054$) with regard to the effect of social trust, whereas social trust positively promoted ($b=0.102$; $p<0.01$) the effect of institutional trust. According to the detailed analysis of decomposition, social trust explained almost 38 percent of the institutional trust impact. On the other hand, only 19 percent of the impact of social trust was explained by institutional trust.

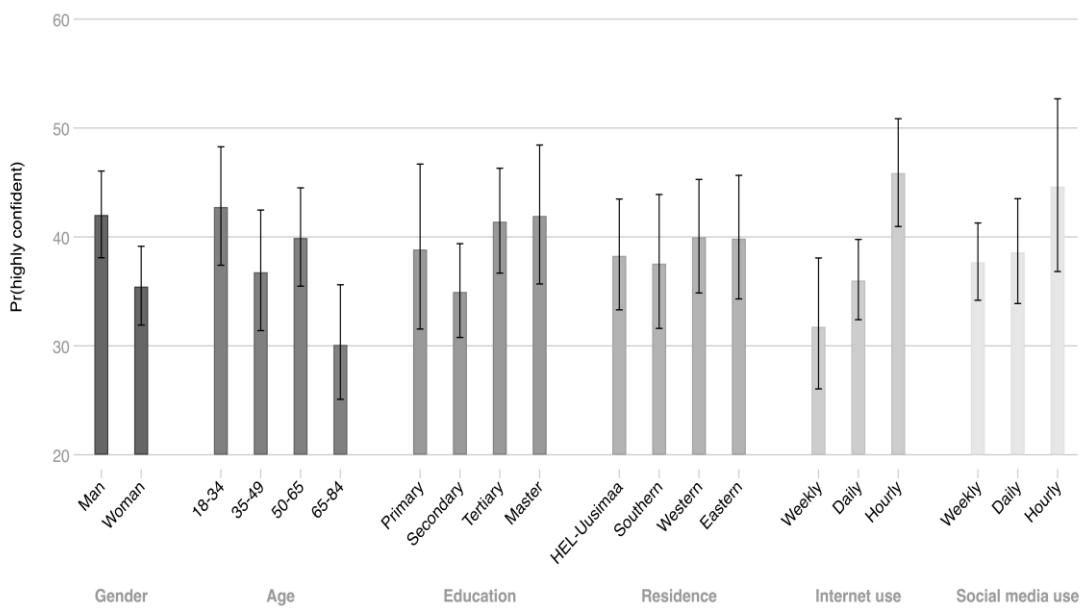


Figure 3. The effects of demographic factors on the confidentiality of private messages
Unadjusted predicted probabilities

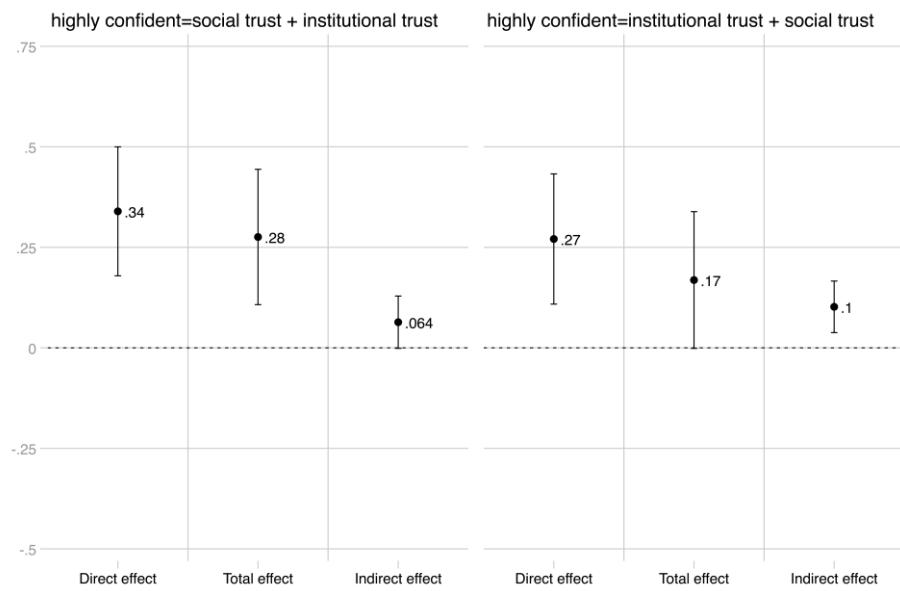


Figure 4. The effect of social and institutional trust on confidence of the protection of private messages. Decomposed (khb) logit coefficients

5. CONCLUSION

In this study, our goal was to determine how trusting the users of digital messaging and communication services are in relation to privacy expectations, i.e. how confident users are that their private messages are kept confidential. These results can help us to understand aspects of social and confidential communication in the digital environment. First of all, we need to emphasize that it was surprising that the extent of user trust could not be explained by factors of demographic background or Internet access. Instead, high estimates of social and institutional trust form a favorable basis for having high trust in the protection of private messages, i.e. the belief that private messages will stay private. Together our results merged with descriptive findings give us the confidence to put forth a conclusion that either trust in security of private messaging or the revealed trust effect were not explained by the demographic factors of citizens.

Our results are in line with earlier literature, which suggests that there is a close interplay between different forms of trust (Lewis & Weigert, 1985, 972; Kouvo, 2014). Those who trust other people are feeling high levels of trust towards institutions as well. However, our results show that social trust is in a more important position compared to institutionalized trust. This may indicate that people are primarily more worried about actions of other users when considering the security of their confidential content. Here, we need to bear our mind that, in some other countries, institutions are less trusted than in Finland. This may mean that when institutional trust is comparatively high on nation scale, different risks are more likely to be recognized at an individual level. To confirm this assumption, it would be reasonable to continue this study by comparing results from Finland with countries having a lower aggregate-level of institutional trust. In addition, our results should also be tested in other countries where institutions are perceived as equally trustworthy as in Finland.

In simple terms, it seems that if a respondent generally trusts other people, the technical devices in between do not diminish the level of that trust. In such cases, the message is expected to be kept in secrecy and the communicational and technological actors are trusted as well. In a sense, technology becomes invisible because it requires no attention in terms of trust. When technology works as expected, it hides its complexity. This means that in terms of the conventional sender-channel-receiver model (e.g. Shannon, 2001), overarching trust between sender and receiver inspires trust in the channel as well. Paradoxically, for

trusting users, other users make feelings of privacy possible. However, if a respondent does not trust other people generally, then the technological ways of communicating become suspicious. In fact, in such a case, privacy comes to designate muteness, silence, and non-communication. In terms of life quality, being constantly suspicious is a wearisome task. Everything raises doubts – both other people and technology.

Overall, it seems that trust remains an essential function in the use of communication technology and services. It would be important for future research to assess how different platforms' reliability and various levels of trust are channeling the use and avoidance of different communication platforms and services. It should also be noted that our data focuses exclusively on Finland and therefore these findings cannot be generalized internationally. Also, because trust in its different forms remains a highly important factor in all fields of society, it is important to research the more nuanced associations between trust and different aspects of digital media.

ACKNOWLEDGEMENT

This research was partly supported by the Strategic Research Council of the Academy of Finland (decision number 314250).

REFERENCES

- Bischof, D., 2017. New Graphics Schemes for Stata: Plotplain & Plottig. *Stata Journal*, Vol. 17, No. 3, pp 748–759.
- Bjørnskov, C., 2012. How does social trust affect economic growth? *Southern Economic Journal*, Vol. 78, No. 4, pp 1346–1368.
- Blomqvist, K., 1997. The many faces of trust. *Scandinavian journal of management* Vol. 13, No. 3, pp 271–286.
- Coleman, J.S., 1988. Social capital in the creation of human capital. *American Journal of Sociology*, Vol. 94, pp 95–120.
- Dinesen, P. T. and Sønderskov, K. M., 2015. Ethnic diversity and social trust: Evidence from the micro-context. *American Sociological Review*, Vol. 80, No. 3, pp 550–573.
- Ellison, N. B., 2007. Social network sites: Definition, history, and scholarship, *Journal of computer-mediated Communication*, 13(1), pp. 210–230.
- Ermisch, J. et al, 2009. Measuring people's trust. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 172, No. 4, pp 749–769.
- Freitag, M., and Traunmüller, R., 2009. Spheres of trust: An empirical analysis of the foundations of particularised and generalised trust. *European Journal of Political Research*, Vol. 48, No. 6, pp 782–803.
- Goffman, E. 1986. *Frame analysis: an essay on the organization of experience*. Northeastern University Press, Boston.
- Gross, R. and Acquisti, A., 2005. Information revelation and privacy in online social networks, in. Proceedings of the 2005 ACM workshop on Privacy in the electronic society, ACM, pp. 71–80.
- Helliwell, J. F. and Putnam, R. D., 2007. Education and social capital. *Eastern Economic Journal*, Vol. 33, No. 1, pp 1–19.
- Inglehart, R., 1999. Trust, well-being and democracy. In *Democracy and trust*, ed. Mark E. Warren, 88–120. Cambridge University Press, Cambridge.
- Jann, B., 2014. Plotting Regression Coefficients and Other Estimates in Stata. *The Stata Journal*, Vol. 14, No. 4, pp 708–737.
- Karlson, K. B. et al, 2012. Comparing regression coefficients between same-sample nested models using logit and probit: A new method. *Sociological Methodology*, Vol. 42, No. 1, pp 286–313.
- Keipi, T. et al, 2018. Assessing the social media landscape: Online relational use-purposes and life satisfaction among Finns. *First Monday*, 23(1).
- Kouvo, A., 2014. *Luottamuksen lähteet. Vertaileva tutkimus yleistynytä luottamusta synnyttävistä mekanismeista*. Sarja C 381. University of Turku, Turku.
- Latour, B., 2005. *Reassembling the social: an introduction to actor-network-theory*. Oxford University Press, Oxford.
- Lewis, J. D. and Weigert, A., 1985. Trust as a social reality. *Social forces*, Vol. 63, No. 4, pp 967–985.
- Mayer, R. C. et al, 1995. An integrative model of organizational trust. *Academy of management review*, Vol. 20, No. 3, pp 709–734.
- McEvily, B. et al, 2003. Trust as an organizing principle, *Organization science*, Vol. 14, No. 1, pp 91–103.

- Möllering, G., 2006. *Trust: Reason, routine, reflexivity*. Elsevier, Oxford.
- Nannestad, P. et al, 2014. Do institutions or culture determine the level of social trust? The natural experiment of migration from non-western to western countries. *Journal of Ethnic and Migration Studies*, Vol. 40, No. 4, pp 544–565
- Official Statistics of Finland [OSF], 2018. Use of information and communications technology by individuals [e-publication]. ISSN=2341-8710. Helsinki: Statistics Finland [referred: 2.3.2018]. Access method: http://www.stat.fi/til/sutivi/index_en.html
- Rothstein, B., and D. Stolle., 2008. The state and social capital: An institutional theory of generalized trust. *Comparative Politics*, Vol. 40, No.4, pp 441–459.
- Rothstein, B. and Uslaner, E. M., 2005. All for all: Equality, corruption, and social trust. *World politics*, Vol. 58, No. 1, pp 41–72.
- Rousseau, D. M. et al, 1998. Not so different after all: A cross-discipline view of trust. *Academy of management review*, Vol. 23, No. 3, pp 393–404.
- Schradie, J., 2011. The digital production gap: The digital divide and Web 2.0 collide. *Poetics*, 39(2), 145-168.
- Shannon, C. E., 2001. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, Vol. 5, No. 1, pp 3–55.
- Subramanian, S. V. et al, 2002. Social trust and self-rated health in US communities: a multilevel analysis. *Journal of Urban Health*, Vol. 79, No. 1, pp S21-S34.
- Tynan, D., 2016. Facebook's journey 'only 1% done' after surge in revenue, Zuckerberg says. *The Guardian* (27.7.2016) <https://www.theguardian.com/technology/2016/jul/27/facebook-ad-sales-growth-quarterly-results>
- Van Deursen, A. J. and Van Dijk, J. A., 2014. The digital divide shifts to differences in usage. *New media & society*, Vol 16, No 3, pp 507-526.
- Van Dijk, J.A., 2005. *The Deepening Divide*. Sage Publications, London.
- Vuorinen, J. and Tetri, P., 2012. The Order Machine - The Ontology of Information Security. *Journal of the Association for Information Systems*, Vol. 13, No. 9, pp 695–713.
- Young, A. L. and Quan-Haase, A., 2013. Privacy protection strategies on Facebook: The Internet privacy paradox revisited. *Information, Communication & Society*, Vol. 16, No. 4, pp 479-500.

CREATING COMPETENT TEAMS FOR EMERGENCY MANAGEMENT AND RESPONSE

Danilo Hasse¹, Fernando A. Ostuni Gauthier¹, Carlos Roberto De Rolt² and Gisiela Hasse Klein²

¹*Departamento de Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina - UFSC
Campus Universitário Reitor João David Ferreira Lima – Florianópolis – Santa Catarina, Brasil*

²*Centro de Ciências da Administração e Socioeconômicas, Universidade do Estado de Santa Catarina - UDESC
Campus Av. Madre Benvenuta, 2037 - Florianópolis – Santa Catarina, Brasil*

ABSTRACT

ICTs and social networks can contribute to the emergence of creating competent teams for emergency response. For example, during disasters such as Pakistan flood of 2010, Japan tsunami of 2011 and Thailand flood of 2011, On-Line Social Networks (OSNs) have served as a main technology for numerous people seeking to share information about their personal status, to request resources, or to report the status of their community. They can be used at least for three main functions: detecting emergencies, disseminating information, and managing emergencies. On the other hand, the combination of Internet and mobile technology generated smart devices with associated sensor technologies that are becoming crucial parts in delivering supports during disaster and emergency situation. However, such use of mobile devices usually requires a reliable support system from crowdsensing technologies and back-end intelligent systems. In this way, the mobile crowdsensing must be designed with focus on mechanisms to identify and localize survivors and first responders in an incident zone, as well as mechanisms for competence characterization provided by social networks to facilitate coordination for individuals as well as crowds. These main technologies, the crowdsensing and social networks combined with semantic web and ontologies can provide a complete emergency response system that is the purpose of the CO-SEMIWA platform. This platform permits an interaction among participants to create dynamically competent teams where all participants perform tasks and solve problems in a specific emergency context and situation.

KEYWORDS

Emergency Response, Social Networks, Crowdsensing; Ontologies, Expertise Location, Competent Teams

1. INTRODUCTION

During disasters such as Pakistan flood of 2010, Japan tsunami of 2011 and Thailand flood of 2011, On-Line Social Networks (OSNs) have served as a go-to-technology for numerous people seeking to share information about their personal status, to request resources, or to report the status of their community. In part, this is due to OSNs in providing an innovative platform that enables users to easily share information without the constraints imposed by distance, time, or space. Also, most platforms (e.g., Facebook, Twitter, LinkedIn etc.) are designed to work across devices (e.g., mobile devices, computers etc.).

Consequently, government, not-for-profit organizations, and citizens have found innovative ways to leverage OSNs tools to share information and coordinate action during and after disasters. OSNs can be used in at least three distinct ways to manage emergencies. They can be used while detecting emergencies, disseminating information, and managing emergencies (Carter et al. 2014).

ICTs and Social Networks can contribute to the emergence of creating competent teams for Emergency Response using emergent technologies such as mobile crowdsensing, Wearable Sensors Devices (WSD) and semantic web / ontologies. These technologies enable to provide a complete system and infrastructure to manage emergency situations in the response phase of the emergency management. Thus, with notion of competence the interaction among participants can create dynamically teams from a previously known emergency response platform where all participants perform tasks and solve problems in a specific emergency context and situation.

This paper presents a description of a framework for extraction competence profiles by social networks and creation of competent teams in the scope of CO-SEMIWA's project a semantic middleware and mobile crowdsensing platform for emergency response. For this purpose, a literature review was carried out on topics related to the goals of this proposed system and are grouped into two main sections: section 2 presents the Emergency Management and Response; and section 3, describes issues for Creating Competent Teams. The section 4 details our CO-SEMIWA's Framework to deal with team formation over a social network such as LinkedIn and the section 5, the Conclusion.

2. EMERGENCY MANAGEMENT AND RESPONSE

Every year, millions of people are affected by natural disasters such as earthquakes, tsunamis, volcano eruptions, hurricanes, tornados and floods, and governments all around the world spend huge amounts of resources on the reconstruction and the preparation for such calamities. In 2013 alone, 97 million victims were affected by 330 natural disasters and the economic damage amounted to 156.6 billion dollars (Rosas et al. 2016).

In 1999, The United Nation's International Strategy for Disaster Reduction (UNISDR) was established to coordinate efforts on disasters risk initiatives around the world. In 2009, the UNISDR published the UNISDR Terminology on Disaster Risk Reduction that aims to promote common understanding and common usage of disaster risk reduction concepts and to assist the disaster risk reduction efforts of authorities, practitioners and the public. According to UNISDR (2009) this document is still valid for reference, where the term Emergency Management was defined as: "the organization and management of resources and responsibilities for addressing all aspects of emergencies, in particular preparedness, response and initial recovery steps".

More recently, the third United Nations World Conference on Disasters Risk Reduction (WCDRR) was held in Sendai, Japan, in 2015 and the Sendai Framework for Disasters Risk Reduction (DRR) 2015-2030 was announced as a successor instrument to the Hyogo Framework for Action (HFA). The UNISDR is the key acting body in the United Nations and serve a focal point for coordination of disaster risk reduction as well has been tasked to support the implementation and review of de Sendai Framework.

Therefore, an emergency is a situation that poses an immediate risk to human health, life, and property, which requires urgent interventions to prevent its worsening.

In general, the emergency management as a cycle containing four phases: mitigation, preparedness, response and recovery. The mitigation phase provides redundancy in preparedness plans to avoid re-occurrence of a disaster or to diminish the destructive impact. Preparedness phase details on exact command and control actions to be performed by responsible Emergency Management Agencies (EMA) and first-responders in response for a disastrous situation.

The response phase puts in practice the plans conceived during the preparedness phase in order to restore order, help the affected people, and neutralize the most immediate hazards (life, health, property, and environment). For instance, the response phase immediately after the event includes multiple functions, such as damage assessment, response-needs assessment, response prioritization, coordination and mobilization of rescue operations, resource and logistic planning, evacuation planning, situation monitoring, and timely information dissemination to citizens and organizations (Cong et al. 2015).

We concentrate in this paper on the coordinating emergency assistance efforts for the response phase of the emergency management cycle. In this phase the interventions are organized as a process that is usually described in an emergency control plan, named an emergency response process.

In this process, all individuals, groups, and communities manage hazards in an effort to avoid or reduce the impact of disasters. It is based on the idea that an emergency response process is quite similar to a business process and, therefore, can be modeled as a workflow (Cong et al. 2015).

Therefore, the reliability, relevance and timeliness of data are crucial because they will foster effective coordination between the national and local governments, international organizations and non-governmental actors, and in particular with affected households and communities. Technology is believed to speed up the coordination efforts to relief the emergency situation and rescue victims for saving lives. The combination of internet and mobile technology generated smart devices with associated sensor technologies that are becoming crucial parts in delivering supports during the disaster and emergency situation. However, such use of mobile devices usually requires a reliable support system from crowdsensing technologies and back-end intelligent systems (Geumpana et al. 2015).

In this way, a mobile crowdsensing system is crucial for coordinating emergency efforts by a crowd management component with focus on mechanisms to identify and localize survivors and first responders in an incident zone, as well as mechanisms for competence characterization by social networks to facilitate coordination for individuals as well as crowds. Crowdsourced data from sources verified to be in the impact zone of an incident can be collected and integrated to the scene of the incident. On the other hand, several works relate the social networks as one the most promising applications in the emergency management and response (Geumpana et al. 2015). Finally, the semantic web technology and ontologies have been used in the emergency management and response area. Ontologies can also unify and normalize data from different resources, e.g., social networks, syntactically and semantically and associate it with emergency domain knowledge. They help to create meaningful relationships between information resources and to allow machines to process, infer, or combine the information from different sources automatically into a consistent body of knowledge (Moi et al. 2016). In this way, several works in the literature can be founded using Semantic Web and Ontologies for emergency management systems (Moi et al. 2016).

Thereby, the mobile crowdsensing and social networks combined with semantic web and ontologies are the main components in the CO-SEMIWA platform and this integration represents a novel approach for emergency response in the literature. This paper proposes a semantic middleware framework for the back-end intelligence of the CO-SEMIWA platform providing the possibility to enhance the creation of competent teams for emergency response as described in the next sections.

3. CREATING COMPETENT TEAMS

In this section, we reviewed several work to creating competent teams in the context of social networks and emergency management and response systems. We have concentrated on the most prominent works describing the main issues related to this effort in the next topics, as follows: 3.1 Team Formation with Social Networks; and, 3.2 Tasks Assignment and Delegation.

3.1 Team Formation with Social Networks

The social network of an individual plays a perfect role as a medium for the spread of information, ideas and influence among the team members. Awal and Bharadwaj (2014) present a comprehensive literature review of this area and propose a model for team formation.

Lappas et al. (2009) introduced the problem of team formation in the context of social networks, and thereafter, much work has been recently carried out in this area by considering various algorithms with different objectives (e.g., minimizing the coordination costs, covering all skills, etc.). The problem can be stated as: “given a particular task and a set of experts, the problem is to identify the right team of experts that can collectively perform the given task in an effective manner by covering all the required skills”. Li and Shan (2010) extended the work further by considering the problem of generalized tasks (the type of task where a designated number of experts is needed for each required skill), identifying a relevant set of experts for a particular task requires the discovery of an optimal configuration of a team. Wi et al. (2009) consider the two elements of the competence of a team in a social network setting, that is: i) knowledge competence: personal knowledge and knowledge from the social network; ii) collaboration competence: based on social network analysis measures (density, degree centrality and closeness centrality). Bonchi et al. (2011) show that Social Network Analysis (SNA) help in visualizing and understanding the roles and relationships that facilitate or hinder the collaboration and sharing of information and knowledge in an organization.

The Team Formation algorithms also have the first work with Lappas et al. (2009) in the team formation problem in the presence of a social network and presents greedy algorithms for minimizing the diameter and the cost of the minimum spanning tree (MST) induced by the team. Lappas et al. (2009) impose the strong assumption that a skill requirement of a task can be fulfilled by a single person. On the other hand, the solutions obtained by all these algorithms (including the MST algorithm) can be shown to be connected subgraphs if the underlying social graph is connected. Gajewar and Sarma (2012) explore the usefulness of the density based objective in finding strongly connected teams. Anagnostopoulos et al. (2012) considered an online team formation problem where tasks arrive in a sequential manner and teams have to be formed minimizing the (maximum) load on any expert across the tasks.

3.2 Tasks Assignment and Delegation

Assignment of tasks to people, following different kind of strategies and patterns has already been studied extensively in the literature. For example, Palen and Liu (2007) describe the emergence of information pathways for public participation in emergency response, and discuss the need to coordinate improvised activities between temporary organizations and formal response teams. They include examples of how the majority of victims in disaster are saved by local, ad hoc volunteer groups.

In general, these approaches can divided into three categories (Bessai and Charoy 2016): i) the mono-criterion approaches where the objective is to minimize only one criterion (generally the overall execution time); ii) the multi-criteria approaches where several objectives often conflicting are considered simultaneously; and iii) the constraint where the initial multi-objectives problem is transformed to the mono-criterion one (only one objective is considered and the rest of objective functions are transformed to constraints). These approaches have been proved to be effective in the case where all tasks are automated. A solution that is widely used to cope with this issue is to call for more resources to deal with the new load and to delegate tasks to these resources. These type of resources can be found using crowdsourcing platforms in an elastic way. However the problem with crowdsourcing is to find resources with the right skills and to trust the quality of output by unknown people. For instance, in case of a humanitarian crisis, it is usual to call to social networks and the crowd to execute task instead of responders (Bessai and Charoy 2016).

In this sense, the Dynamic Task Allocation is other important issue to the Task Assignment. Task monitoring and reassignment in large-scale disaster response, in rapidly changing environments, in the presence of data explosion is challenging, since many events may get lost in such volume of data. Intelligent decisions need to be made quickly, evaluated and changed depending on updated information. In emergency response, where the user's environment changes rapidly, mobility context is an essential element that needs to be part of the decision making process. However, there is a lack of systems that continuously monitor the mobility context of task owners to determine the quality or progress of the task, and whether additional resources should be added, removed, or reassigned to someone else who could do it faster and within the given time constraints (Luqman and Griss 2010).

Government's studies show that existing communication systems used for emergency response are ineffective and incompatible across teams. Recently, the interoperability of existing emergency response communications has been highlighted. Furthermore, the closed communication systems used by first responders and emergency response personnel are often incompatible with those used by volunteers. Current technologies are also inadequate for spontaneously formed disaster response teams. Well informed decisions must be made rapidly and reevaluated periodically to gauge their effectiveness (Luqman and Griss 2010).

Achieving automation in the task delegation process can be done by using a number of methods, most of them quite new and derived from the field of artificial intelligence. However, the chosen model was the usage of classic iterative algorithms that create automation with respect to task delegation by applying a set of instructions to properly stored, sorted and indexed data available on the project's back-end database. In general, project data is being distributed across multiple nodes via NoSQL databases where for this purpose are better suited than classic relational databases (Pop and Boian 2014).

4. CO-SEMIWA: A SEMANTIC MIDDLEWARE FOR EMERGENCY RESPONSE

The COLLEGA Semantic Middleware (CO-SEMIWA) is a semantic middleware for collaborative assistance in emergency response scenarios. The system is context-aware collecting emergency information from a crowdsensing platform and Wearable Sensors Devices (WSD) connected in the monitored people, extracting competence information from social networks such as LinkedIn to know potential participants by profiles and providing semantic information to build automatically teams for emergency response.

The CO-SEMIWA Architecture can be visualized in Figure 1, where its functional layers and modules are shown. In our research work, the architecture model is layered into two layers: the COLLEGA Middleware – the Application Layer (AL) and COLLEGA Semantic Middleware – the Semantic Layer (SL).

The COLLEGA middleware (AL) is our previous crowdsensing platform (De Rolt et al. 2017) to support mobile participatory healthcare communities and a reference middleware architecture, called by acronym

COLLEGA (COLLaborative Emergency Group Assistant), for handling anytime and anywhere emergency situations using a crowdsensing to manage the overall life cycle of an emergency situation, from alarm detection and understanding to the formation of virtual mobile communities formed by passing-by users.

Initially, the social network platform is composed of potential participants who do not know each other, and do not share a common interest in the virtual or physical world. Participants are physically in the same place at that time, they have no degree of previous cooperation and don't have any kind of communication synchronously or asynchronously with the WSD user but are users of mobile computing systems, such as smartphones and choose to install a crowdsensing APP (COLLEGA) of emergency response public service.

The Figure 1 also represents the user that consciously opted to use a remote monitoring system - WSD and, in doing so, joined a pre-existing virtual community with known and registered members. Participants of this pre-existing virtual community opted for some reason (family, friends, volunteers or EMAs) to join the community. Thus, the Application Layer (AL) is the core of COLLEGA's functionalities to form the overall structure with modules that manage the overall cycle of an emergency. The AL Layer encompasses multiple devices like smartphones. The "Applications" on the AL Layer are responsible for providing visualizations of ongoing processes or assist the users in sending control instructions.

The application description is an instance of our provided application model and holds meta-information about the executing device, the associated processes and a unified identifier that finds corresponding distributed applications on several devices. The Application Layer executes functions and provides services for the system.

The modules and respectively functionalities of the Application Layer are: a) Monitoring System (MOS): getting and analyzing data collected from sensors and receiving and sending emergency alerts by individuals; b) Emergency Context Analyzer (ECA): detecting and classifying emergency situations and choosing the most appropriate control plan; c) Potential Participant Analyzer (PCA): automatically identifying potential participants; d) Virtual Community Manager (VCM): ensuring that potential participants can properly access monitored emergency data and monitoring task implementation; and, e) SFW: providing information about the monitored emergency in a secure way to participants.

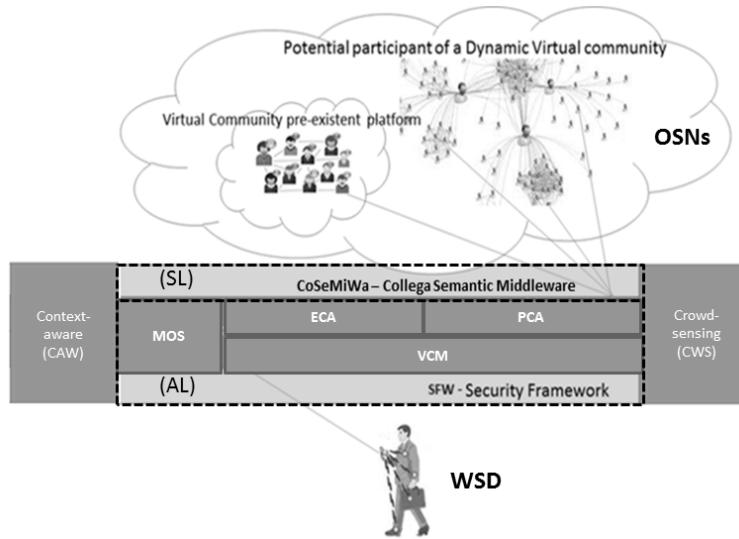


Figure 1. CO-SEMIWA architecture

The COLLEGA crowdsensing platform (CWS) interacts with the Application Layer through execution of functions to collect data of all available users in the emergency area. From technological point of view, a complete crowdsensing platform consists of an app running on smartphones and a web application on the back-end. Main support functions include: management of the crowdsensing requests (called tasks); sensing of data collected either passively (automatically via smartphone sensors), actively (with direct user participation), or with a mix of these modes; and evaluating assignment of tasks to users for future crowdsensing activities. The crowdsensing client is the component that takes care of receiving tasks, asking users whether they want to run them, managing data collection, and uploading results. Functionally, a

crowdsensing client comprises two main components: the task management component and the sensing management component. These components are responsible for both interacting with users and accessing smartphone sensors. The server side provides management, storage, and analysis of crowdsensed data. At the highest level comprises two main parts: the back-end and the crowdsensing manager (Corradi et al. 2015).

As explained above the WSD and CWS platforms connect sensors through mobile phones with relevant data about emergencies to be managed by COLLEGA application layer (AL). However, the selectively of OSN profiles is an important feature for creating competent teams to execute the control plans.

In this sense, our previous work (Hasse et al. 2017) studied relevant references of competence developing a new proposed approach for the competence in the professional social networks contributing with a model of “Competence Networks” with four dimensions represented by acronym KSAN (knowledge, skill, attitude and network) mapped on LinkedIn’s users profiles. This model is useful to implement our framework to create competent teams with social networks for emergency response. The table 1 shows this notion of competence mapped over the LinkedIn’s datasets (Hasse et al. 2017):

Table 1. Mapping the Competence Network with LinkedIn’s datasets

Categories	Datasets LinkedIn
Knowledge (K)	<i>Education</i>
Skill (S)	<i>Skill</i>
Attitude (A)	<i>Volunteer</i>
Network (N)	<i>Connections</i>

When assembling a team special attention is paid to its structure. The structure itself should be such that the knowledge, skills and character traits (attitude) of team members are complementary. Teams with the same or similar profiles of experts have not proved to be efficient in practice. Besides, practice imposes that, apart from skill and work experience, the criteria for team member selection should include the member character traits and relationships (network). This structure is another relevant data to implement an intelligent system with functionalities for team creation and tasks assignment/delegation.

Therefore, the proposed general architecture aims at integrating these three types of platforms: (1) Wearable Sensors Devices (WSD), (2) Crowdsensing System (CWS) and (3) Online-Social Networks (OSNs) to creating an emergency assistance service through an integrated mobile system and this requires the proposition of a new integrative semantic middleware – the Semantic Layer (SL). The Semantic Layer (SL) is formed by our proposed architecture to integrate all these functionalities and platforms in a semantic way. In general, the semantic layer is responsible for all interoperability between systems proposed as components of middleware. We designed these platforms with the same database in a semantic web W3C patterns (OWL 2, RDF, SWRL, RIF). The functionality of CO-SEMIWA seamlessly register all participating devices, aggregates the sensor semantic data, and provides channels to route semantic data to subscribed processes.

Thus, the context model management in emergency scenarios should be considered for ensuring the provision of the right information at the right time and place. As it was pointed in Evchina et al. (2012), ontologies play a crucial role in context-aware systems.

Ontology is a knowledge representation model, which combines the concepts and their relations among each other. Ontologies possess advantages, which make them preferable for use in context-aware systems. They have a set of advantages including human/machine understandable form, capabilities to infer new knowledge, and they are reusable among different domains. On the other hand, ontologies have already been used for discovery services and semantic interoperability among ubiquitous computing devices.

In this work we integrated context-aware extracting data from WSD, CWS and pre-existing OSNs such as LinkedIn, via API, by using extraction and analysis tools like R and Gephi that provide contextual emergency information. The COLLEGA Semantic Middleware (CO-SEMIWA) provides Context-Awareness (CAW) and the functionality to register all devices, acquire the sensor data, unify and aggregate it as semantic data, and route this data to subscribed processes.

The CO-SEMIWA holds all knowledge of the system in a semantic world model. The only requirement for maintaining interoperability with multiple different devices is to provide a common interface for data transmission to the associated process and an application description. Our proposal is to do this using the semantic layer. The CO-SEMIWA framework is shown in Figure 2.

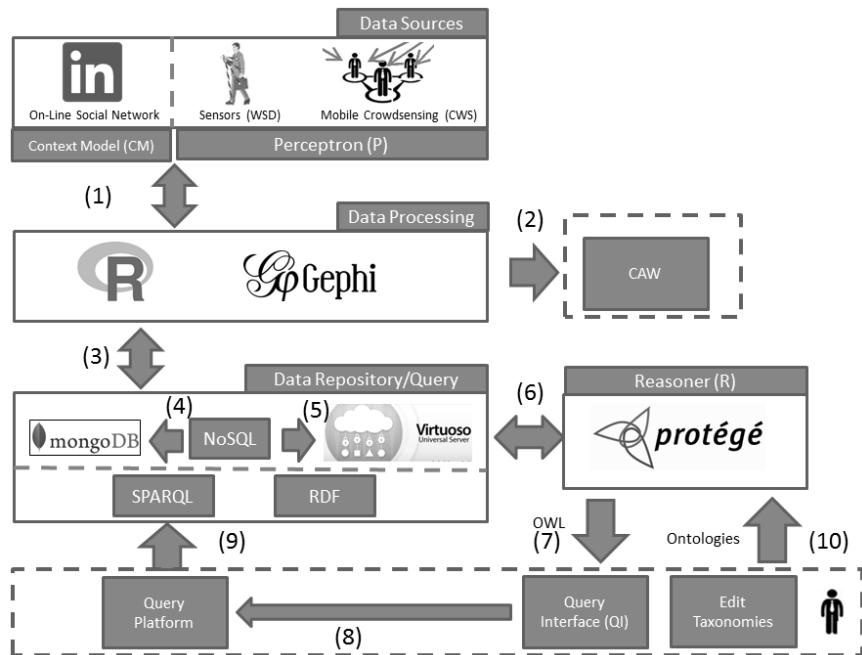


Figure 2. CO-SEMIWA framework

We have implemented the CO-SEMIWA semantic layer (SL) like a framework that uses several tools to extract, store and analyze context of competences from LinkedIn and relevant information of emergencies integrated with emergency data from WSD and the CWS to operate as a single emergency service session in COLLEGA middleware (application layer - AL).

Initially, the Data Sources are obtained from LinkedIn, the WSD and from the COLLEGA Crowdsensing System. They constitute data sources to the components of Semantic Layer Middleware: the Contextual Model (CM) with LinkedIn data and the Perceptron (P) with WSD and crowdsensing data (1).

These data sources are processed by the Data Processing framework that realizes the CAW system module of the CO-SEMIWA architecture (2), by detecting the context and triggering the ECA's module in the COLLEGA middleware (application layer) with the corresponded situation/status flag. Context detection lets ECA choose (select) the control plan, responding with emergency assistance for the event detected. After detection also are related tasks, profiles and geolocation of the first-responders and victims.

The Data Processing framework can be implemented using the RStudio platform - which is a complete R language environment to import and manage data and the Gephi tool. With Gephi tool we can provide all sort of SNA for graph visualization and retrieve metrics to understand the social network relations between users in the system. In this way, by using R we are developing a computer program that is able to implement the CM component as follows: 1) extracting competence profiles by LinkedIn's API; and 2) populates NoSQL databases storing data in a RDF format (3). After RDF data process, the NoSQL databases are the central component of the CO-SEMIWA's Framework.

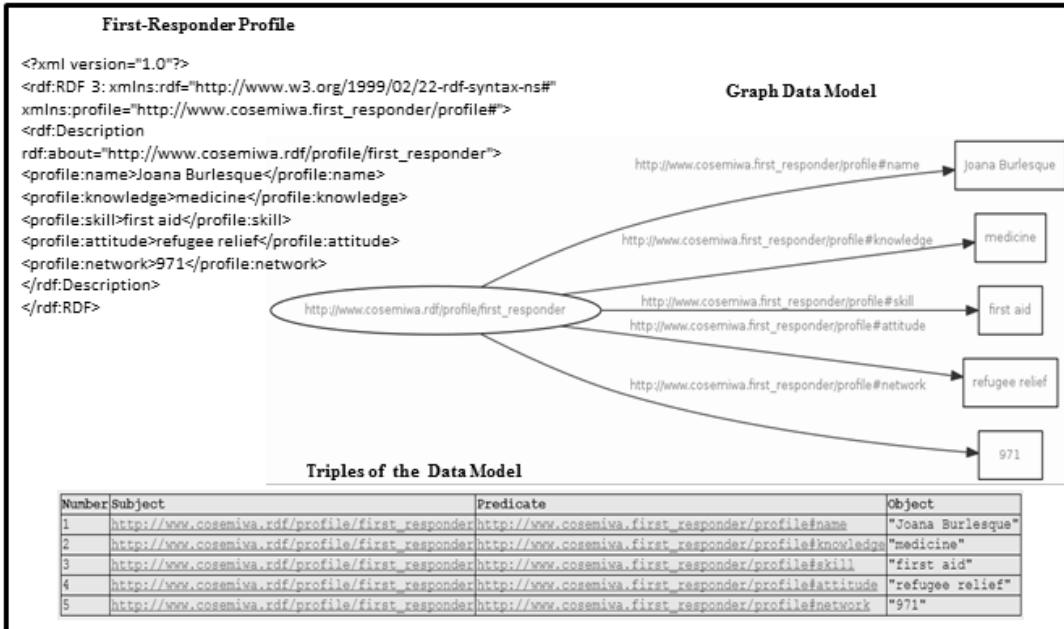


Figure 3. Example of CO-SEMIWA profile

We have deployed one prototype application in R, as shown in the figure 3 that is an example of LinkedIn's user profile (identification data) extraction and transformed in a RDF format (RDF/XML) and respective data model representations to storage the First Responder's profile in the NoSQL databases. The Data Repository/Query manages the processes of data storing and interlinking. The storage contains two different data repositories of NoSQL, RDF triple: MongoDB and Virtuoso databases. NoSQL repository takes charge of maintaining the big datasets that are harvested from various data resources from data on the Web including data generated from the COLLEGA client application in the CWS and WSD. LinkedIn's user information gathered in MongoDB database is constantly analyzed in a batch mode by the Reasoner (R) provided by Protégé environment that is a free, open-source ontology editor and framework for building intelligent systems with a built-in reasoner called HermiT. There are other reasoners available for Protégé, including Pellet and FaCT++ (plug-ins). Basically, new information from all users is searched in MongoDB database through Mongo sub-process (4). User-to-user, with its identification, OWLOnto sub-process (5) performs a search of its existing user profile ontology or a predefined ontological model within Virtuoso NoSQL database of RDF triples. Based on the extracted information, the aim of Interest Emergency and Event Detection sub-process (6) is to generate new emergency event relationships and concepts on the disaster ontology. Over these relationships and concepts, Inference sub-process – OWL (7) deduces new information about interest emergency event and disaster context by the reasoner HermiT in the Protégé to Query Interface (QI).

OntoVirtuoso sub-process (8) allows saving the updated user profile in the Virtuoso NoSQL database with the new information inferred by the reasoner HermiT (Protégé). Once generated or updated the ontological profile for all users, Mongo sub-process set in the MongoDB database that users have been processed. As a result, all relationships generated are stored and can be queried in Virtuoso database, where all sort of information can be accessed without performance problems. In this way, data also can be accessed using a SQL-like query language called SPARQL (9). Finally, to investigate how domain knowledge can help in the acquisition of disasters events we use ontology (by editing taxonomies) (10).

4.1 Related Work

The use of mobile software technology to create teams in the domain of disaster and emergency response is still in its infancy. Despite the fact that an emergency situation may develop very dynamically and people should have the right information in a very short time at a mobile location, few projects can be found in the literature developing solutions with all these concerns. For example, the First Aid (Surachat et al. 2013) is

developed to give some preliminary instructions for taking care of users in Android smartphones. Basically is a navigation system that uses Google API (maps) for searching an appropriate or suitable way or path to the nearest hospital. In the case of any emergency this function is activated on user's smartphone to navigate victims through the shortest path to the hospital. This application provides two types of services, first one is different sets of first aid guidance and second one is finding the appropriate and suitable hospital to go. The EMuRgency (Kalz et al. 2013) is a cross-border alarm system and a learning programme to teach people how to give cardiopulmonary resuscitation (CPR). The Volunteer-Notification-System (VNS) allows the notification of volunteers nearby a cardiac arrest. An E-Learning Open Educational Resource (OER) is developed on three different levels and in 4 languages allowing institutions to reduce face-to-face training time and to train large numbers in basic CPR skills and knowledge. The limitation of these approaches is the restrict use to creating and coordinating teams during an emergency situation and managing control plans with a huge number of sensors devices and people involved by the OSNs as the CO-SEMIWA aims to achieve. Another point is that they aren't using an intelligent back-end platform based on ontologies or other artificial intelligence technology. These aspects could in near future to limit the range of these solutions in dynamicity and complexity. In this way, there are some initiatives with the same CO-SEMIWA's approach, like for example in Jing et al. (2014) a mobile context-aware disaster response system called Skyguard that consists of a mobile client application and a disaster response server. The solution is differentiating itself from CO-SEMIWA by not adopting semantic web technology in the project. On the other hand, in (Cordeiro et al. 2011) the proposed system is using the Linked Data's approach in the design of information infrastructure for Collaborative Emergency Management System extensively adopting the semantic Web technology and creating teams to act in a disastrous scenario, but this solution doesn't have been focused for mobile applications and crowdsensing or using OSNs to amplify the scope of users scalability.

5. CONCLUSION

In this paper we identified and investigated the main issues related to handle with creating competent teams for emergency response using data collected by social networks and emergent technologies required such as Crowdsensing System, Wearable Sensors Devices, algorithms and intelligent systems (Ontologies).

As a result we propose a novel framework to integrate all requirements to provide a platform with capabilities to manage all life cycle of emergency assistance efforts for the response phase in the emergency management. For this, we presented the CO-SEMIWA architecture/framework, a semantic middleware for collaborative assistance in emergency scenarios. Semantic information is provided using semantic web patterns by addressing various Wearable sensors devices for users' interaction, through Crowdsensing System and required competences extracted from OSNs such as LinkedIn to create competent teams.

For future work this platform will be fully operational for conduct a practical use case with semantic data for testing and creating competent teams in a research environment of the emergency response area.

ACKNOWLEDGEMENT

This research was supported by the program CAPES 3º Cronograma – Chamadas de Projetos nº 09/2014, processo nº 8745-13-7 for Carlos Roberto De Rolt and nº 88887.122829/2016-00 for Danilo Hasse.

REFERENCES

- Anagnostopoulos A. et al, 2012. Online team formation in social networks. *Proceedings of the 21st international conference on world wide web, WWW '12*. New York, USA, pp 839–848.
- Awal, G.K. and Bharadwaj, K.K., 2014. Team formation in social networks based on collective intelligence – an evolutionary approach. In: *International Journal of Research on Intelligent Systems for Real Life Complex Problems*, Vol.41, No. 1, pp.627–648.
- Bessai, K. and Charoy F., 2016. Business Process Tasks-assignment and Resource Allocation in Crowdsourcing Context. *Proceedings 2nd International Conference on Collaboration and Internet Computing*. Pittsburgh, USA, pp. 11-18.

- Bonchi F. et al, 2011. Social network analysis and mining for business applications. *In: ACM Trans Intell Syst Technol*, Vol.2, No.3, pp.22.1–22.37.
- Carter, L. et al, 2014. Social Media and Emergency Management: Exploring State and Local Tweets. *Proceedings 47th Hawaii International Conference on System Science (HICSS)*. Waikoloa, USA, pp. 1968-1977.
- Cong L. et al, 2015. E-Net Modeling and Analysis of Emergency Response Processes Constrained by Resources and Uncertain Durations. *In IEEE Journals & Magazines: Transactions on Systems, Man, and Cybernetic*, Vol. 45, No. 1, pp. 84-96.
- Corradi, A. et al, 2015. Automatic Extraction of POIs in Smart Cities: Big Data Processing in ParticipAct. *Proceedings IFIP/IEEE International Symposium on Integrated Network Management (IM2015): Experience Session Paper*. Ottawa, Canada, pp. 1059-1064.
- Crowe, A., 2010. The social media manifesto: A comprehensive review of the impact of social media on emergency management. *In Journal of Business Continuity & Emergency Planning*. Vol 1, No. 5, pp. 409-420.
- De Rolt, C.R. et al, 2017. Competence-based mobile Community Response Networks. *Proceedings IEEE Symposium on Computers and Communications (ISCC)*. Heraklion, Greece, pp.994-1000.
- Evchina, Y. et al, 2012. Ontological Framework of Context-Aware and Reasoning Middleware for Smart Homes with Health and Social Services. *Proceedings IEEE International Conference on Systems, Man, and Cybernetics*. Seoul, Korea, pp.985-990.
- Gale, D., 2001. The two-sided matching problem. Origin, development and current issues. *In: International Game Theory Review*, Vol. 3, Nos. 2 & 3, pp. 237-252.
- Gajewar A. and Sarma A.D., 2012. Multi-skill collaborative teams based on densest subgraphs. *Proceedings of the 12th SIAM international conference on data mining, SDM'12*, Anaheim, California, USA, pp. 165–176.
- Geumpana, T.A. et al, 2015. Mobile Cloud Computing for Disaster Emergency Operation: A Systematic Review. *Proceedings IEEE International Symposium on Technology in Society (ISTAS)*. Dublin, Ireland, pp. 1-8.
- Hasse, D. et al, 2017. Formação de Redes de Competências a partir dos Perfis em Redes Sociais: Um Experimento com Datasets do Linkedin. *Proceedings Iadis Conferências Ibero Americanas WWW/Internet*. Algarve, Portugal, pp.12-20.
- Jing, N. et al., 2014. A Context-aware Disaster Response System Using Mobile Software Technologies and Collaborative Filtering Approach. *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design*. Hsinchu, Taiwan, pp.516-522.
- Kalz, M. et al, 2013. EMuRgency: Addressing cardiac arrest with socio-technical innovation in a smart learning region. *In Interaction Design and Architectures Journal*. Vol 17(2), pp.77-91.
- Lappas T., et al., 2009. Finding a team of experts in social networks. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '09*. New York, USA, pp. 467–476.
- Luqman F. and Griss M., 2010. Overseer: A Mobile Context-Aware Collaboration and Task Management System for Disaster Response, *Proceedings Eighth International Conference on Creating, Connecting and Collaborating through Computing*. La Jolla, USA, pp.76-82.
- Li C.T. and Shan M.K., 2010. Team formation for generalized tasks in expertise social networks. *Proceedings of the IEEE 2nd international conference on social computing SOCIALCOM '10*. Washington, DC, USA, pp. 9–16.
- Moi, M. et al, 2016. An Ontology for the Use of Quality Evaluated Social Media Data in Emergencies. *In IADIS International Journal on WWW/Internet*, Vol. 14, No. 2, pp. 38-57.
- Palen, L. and Liu, S., 2007. Citizen Communications in Crisis: Anticipating a Future of ICT Supported Participation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. San Jose, USA, pp. 727-736.
- Pop, B. and Boian, F., 2014. Algorithms for Automating Task Delegation in Project Management, *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems*. Warsaw, Poland, pp. 1191–1196.
- Rosas E. et al, 2016. Survey on Simulation for Mobile Ad-Hoc Communication for Disaster Scenarios. *In Journal of Computer Science and Technology*, Vol.31, No.2, pp. 326-349.
- Surachat, K. et al, 2013. First Aid Application on Mobile Device. *In International Scholarly and Scientific Research & Innovation*. Vol.7(5), pp.361-366.
- UNISDR, 2009. UNISDR Terminology On Disaster Risk Reduction. United Nations International Strategy for Disaster Reduction. Geneva, Switzerland.
- Wi H., et al., 2009. A Team Formation Model Based on Knowledge and Collaboration. *In Journal Expert Systems with Applications*. Vol. 36, No.5, pp. 9121–9134.
- Cordeiro, K.F. et al., 2011. Use of Linked Data in the Design of Information Infrastructure for Collaborative Emergency Management System. *Proceedings of the 2011 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. Lausanne, Switzerland, pp.764-771.

INTERNET OF THINGS, NEW WAVES TOWARD ORGANIZATIONS

Fateme Tavallaei¹ and Dr. Mohammad Ali Sarlak²

¹*Master Student of Public Administration*

²*Faculty Member of Management Department*

Payame Noor University, Tehran

ABSTRACT

New waves of technologies in 21st century, now comes in the name of Internet of Things. These technologies which are widespread in various contexts of use and different technical characteristics, have benefits and challenges. Organizations, as the important parts of the e-society, confront to the vast IOT usage in multi-level and multi-faced of their nature such as e-maintenance, e-procurement, e-business and so on. In this situation, a suitable adoption, e-readiness, implementation, and change plan is necessary to maximize the advantages and prevent failure of IOT implementation. Though lots of researches have done to measure the variables of organizations confronting new technologies, there is still need to investigate more in the reciprocal relations due to the nature of organizational levels. Technical and conditional factors should be considered in developing internet of things in organizations. This paper tries to have a review about these issues.

KEYWORDS

Internet of Things, Organizations, e-Readiness, Adoption

1. INTRODUCTION

Among new waves of the 21st era technologies, *Internet of Things* is a noticeable one which is going to penetrate all aspects of human lives, in e-health, smart cities, smart buildings, industry, smart environment, smart metering, smart retail, smart agriculture, smart automation, and so on. These inclusive types of technologies have passed the industries by diversity equipment and devices such as GPS in fleet tracking in logistics, smart metering for liquid flow and voltage, accelerometer for earthquake detection, LDRs for smart lightings, ECG, UVA, UVB for e-health and so on (Libelium Website, 2017). But in near future these technologies are not just for special usage and all objects in the world come to the network of IOT. Organizations, the complexities of human beings with unique structures is the part of the society which will get involved in a vast scale to IOT, due to their various and different roles among governances, people, visions, staffs, environments, regularity, and types of monitoring. E-procurement, e-infrastructure, e-business, e-commerce, e-learning, e-library, e-health are some of the applications of cloud systems in the organizational context. So considering the benefits and challenges of IOT around the organizations include of expertise, SMEs, public and private ones with different scope and goals is one of the most necessities which influence and be influenced vastly to the e-society of near future. To confront the challenges of organizational IoT, besides the technical issues, suitable adoption, e-readiness, implementation, and change plan should be considered, too. This paper tries to have a look on to the organizations in near IOT world. The paper can be suitable for researches, managers, stakeholders such as suppliers, customers, employees and investors, which are connected with organizations.

2. INTERNET OF THINGS FROM BIRTH TO NOW

Beginning from the last decades of 20th century in MIT, IoT was born by Kevin Ashton, an expert on digital innovation (Madakam et al, 2015). During the years, the technology have grown with different devices and usages such as Radio Frequency Identification (**RFIDs**) for transferring data, tracking tags and things,

supervision and controlling of things, and automation of electronic payments in trade markets (Suryadevara et el, 2017), Near Field Communication (*NFC*) designed for use by devices within close proximity (generally 5 cm or less) of each other for transition of connected living and interaction (Saha et el, 2016), Wireless Sensr Network (*WSN*) consist of spatially distributed autonomous sensor-equipped devices to monitor physical or environmental conditions (Lee et el, 2015), and *Clouds* which defined by NIST “a model that enables convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction” (Mell et el, 2011), and categorized into three part: IaaS that is the infrastructure as a service, Paas is the platform as a service, and SaaS is the software as a service (Hasan et el, 2015). Today the new generation of these systems become mature enough that can also analyze big data, too. Due to vast and diversity of components and usages, there are ambiguity and different definitions for IoT, but the latest ones defined it as: An open and comprehensive network of intelligent objects, system of interrelated computing devices, mechanical and digital machines, animals or people with the ability to uniquely recognize, signify and access things, sense, communicate, compute and potentially actuate, auto-organize, share information, data and resources, reacting and acting in face of situations and changes in the environment, by the use of intelligence, multi-modal interfaces, physical/ virtual identities and attributes, anytime and anywhere in an internetwork, with the minimizing the human-to-human or human-to-computer interaction (Madakam et el; 2015, Pawade et el, 2017). As said above, IoT is now penetrating to almost all aspects of human lives such as ***Smart Cities*** for “smart parking, structural health, noise urban maps, smartphones detection, electromagnetic field levels, traffic congestions, smart lighting, waste management and smart roads”, ***Smart Environment*** for “forest fire detection, air pollution, snow level monitoring, earthquake early detection, landslide and avalanche prevention”, ***Smart Water*** for “potable water monitoring, chemical leakage detection in rivers, swimming pool remote measurement, pollution levels in the sea, water leakage, and river floods”, ***Smart Metering*** for “smart grid, tank level, photovoltaic installations, water flow, silos stock calculation”, ***Security and Emergencies*** for “perimeter access control, liquid presence, radiation levels, explosive and hazardous gases”, ***Retail*** for “ supply chain control, NFC payment, intelligent shopping application, and smart product management”, ***Logistics*** for “quality of shipment conditions, item location, storage incompatibility detection, fleet tracking”, ***Industrial Control*** for “M2M application, indoor air quality, temperature monitoring, ozone presence, indoor location, vehicle auto-diagnosis”, ***Smart Agriculture*** for “wine quality enhancing, green houses, golf courses, meteorological station network, compost, hydroponics”, ***Smart Animal Farming*** for “offspring care, animal tracking, toxic gas levels”, ***Demotic and Home Automation*** for “energy and water use, remote control appliances, intrusion detection systems, art and goods preservation”, and ***e-Health*** for “fall detection, medical fridges, sportsmen care, patients surveillance, and ultraviolet radiation” (Libelium Website, 2017). According to Greenough (2014), stated that in 2019, IOT is going to add \$1.7 trillion value to the global economy and 26-30 billion objects shall be connected to the IoT in 2020. These estimates shows the serious roles of IOT devices in near future lives. But why IOT is becoming inclusive in such a way? Accuracy, efficiency, economic benefits, better quality of life, saving time & money, and automation of daily tasks through the remotely controlling and integrating devices are the opportunities of using IOT which are mentioned in researches (Saha et el, 2016). But these are just the one side of IOT. Like other cutting-edge technologies, IOT is not without any challenges. Security and privacy, technical issues, risks, testing analytics of IOT applications for developers and users, collecting data, regularity aspects, complexity of sensing environment, power supply, architecture, uniqueness IDs for objects, storing and representing exchanged information are the challenges mentioned in researches (Saha et el;2016, Weinberg et el, 2015, tii Website, 2014, Atzori et el, 2010, Weber et el, 2015, Ebersold et el, 2015, Folk et el, 2015, French et el, 2016, Hodgson, 2015). Though of lots of researches about IOT, there is still need to investigate more in a comprehensive perspective which integrate components of the complexities such as organizations which include goals, visons, infrastructures, structures, teams, staffs, processes and systems. So in continues of this paper, the organizational view of IOT is going to be discussed.

3. ORGANIZATIONAL INTERNET OF THINGS

Organizations are the part of the society that make the relationship between its members and environments to meet their needs. So it can be expected that whatever occur in the environment has some kind of effect on

organizations. Internet of things like other technologies is not far from this issue, not only penetrate in all human life aspects, but also has critical role in organizations' Information Communication Technology processes. By penetrating IoT in organizations, it engages with multi-dimensional network that can influence and be influenced by diversity elements such as goals, human resources, structures and infrastructures, and environments. Regardless of type, size, and industry, organizational internet of things becomes a term that has multi-level and multi-faced entity and includes most aspects in this network and should be reformed in order to become balanced with new conditions. Research orientations in this field goes toward the effectiveness and efficiency of IoT usage in order to respond the changing needs of the environment and clients among inside and outside elements in the best harmony way. Term definitions like quality or standards may be different in this context and the role of human resources changes from functioning to intelligent analyzing. Infrastructure hardware become more physical lightweight through clouds, but more vigilant due to security issues, and enable the organizations to have better choices and success. Due to better transferring data and transparency, organizational structures and environments also are going to have important changes. So because of the capabilities and challenges that IoT brings for organizations, these multi-dimensional entities should be aware enough to provide suitable plan for adoption, readiness, implementation, and change.

4. ADVANTAGES OF ORGANIZATIONAL INTERNET OF THINGS

From a whole view, IT and ICT are key elements in today's organizations. The infrastructure, core business functions, facing globalization, and expanded knowledge-based economy, and the importance of competitive advantage are the reasons to use ICTs (Cartelli, A 2007). ICTs come to make more effectiveness and efficiency in the operations, employees and managerial skills and decisions, better quality, reduce time consuming, man power, cost, and facilitating to access information (Maksoud et el, 2003). In the case of IOT, adding value to organizations (Helmo, P; 2017), cost reduction, service performance (Benlian et el; 2009), scalability of infrastructure, ease of implementation, accessibility, massive storage capacity, access to IT capabilities, elimination of procurement and maintenance, electricity consumption (Mohammed, et el, 2015), simplify the complex process, web-interfaced environments, ease of use for business processes and technical resolutions (Hunter et el, 2008), reduction in the economic downturn, in servers, in server maintenance and staffing costs (Gillen et el, 2008), and achieving long-term organizational sustainability, are the reasons for the organizations to get more interested to use IOT.

5. ORGANIZATIONAL E-READINESS, THE SOLUTION TOWARD IOT FAILURE

Though of lots of ICTs advantages, recent studies in the area of 'IT failure' have shown that 75% of IT investments did not meet their performance objectives (Alshawi, 2007). It can be understand that, the more complexity of the ICTs, need more precise adoption, decision making and operations to prevent huge failure such as Nike and HP (Venkatesh et el, 2008). As the researches show, open standards and interoperability, security and privacy, business continuity, internet dependency (Mohammad et el, 2015), lack of supporting resources, lack of understanding of the Cloud, departmental downsizing, uncertainty with new technology, deterioration of customer care and service quality, increased dependence on third parties, and decrease of satisfying work, are the issues which can influence the failure of IOT implementation (Khajeh-Hosseini et el 2010). Organizations toward cutting-technologies confront to new challenges in adoption, readiness, implementation and change processes. There are vast studies about various types of IT and ICT emergent technologies in the case of individual or societies, but not enough in organization level. One of the key solution of the problem is organizational e-readiness, which is defined as, "the ability of an organizational unit to be prepared, willing to adopt, use and benefit from e-innovations such as e-business, e-government, e-procurement, e-learning, etc (Lou, 2010)". Mutula et el (2006), achieved that e-readiness lead to less turn-around time, faster delivery of services, enhanced product qualities, international competitiveness, a broader market reach, convenience for customers, reduced procurements and transaction costs, and improve the efficiency in purchasing processes, profitability, communication and exchange of information. In the

context of organizations and the applications of IoT in organizations like e-procurement, e-business, e-maintenance, and others, there is need to study the models and reform them by adding the new dimensions according to the cutting edge technologies like IoT. Some studies in this area have been done by Naseebullah et el (2011) for preparing the framework of relevant factors impact e-procurement readiness in Malaysian organizations, or another one by Gilabert et el (2014) for optimizing e-maintenance through intelligent data processing systems which is about the developments in wireless, web-based, mobile and situated computing technologies.

6. APPLICATIONS OF ORGANIZATIONAL INTERNET OF THINGS

Regardless of type, size, and industry, organizations can use IoT for facilitating their functions, such as: **e-business** defined by IBM as “the transformation of key business processes through the use of Internet technologies” (Rosnafisah et el, 2010), so it can reach to the global marketplace and improve current business and competitive advantages (Bordonaba et el, 2012, Ang et el, 2012), **e-maintenance** defined by Crespo et el (2009) and Holmberg et el (2010) as e-technologies (ICT, web-based, tether-free, wireless), integrated with new trends such as e-monitoring, e-diagnosis, and e-prognosis, which will allow for remote and real time knowledge assessment of system performance, **cloud manufacturing** defined by XU as “a model for enabling ubiquitous, convenient, on demand network access to a shared pool of configurable manufacturing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction”, **e-science** defined by Watson et el (2008) as a platform for scientists to share their knowledge and science, **e-invoice** defined by Hernandez-Ortega (2011) and Lian (2015), as a kind of information system service that gathers transaction information, **e-portfolio** defined by Greenberg (2004) and Kim et el (2010) as “a network application that provides the author with administrative functions for managing and organizing work (files) created with different applications and for controlling who can see the work and who can discuss the work (access)”, **e-commerce** defined by Sohaib et (2017) as selling and buying of products or services over the Internet, allows a helpful analysis of the data which further aids in the business intelligence, product suggestions, and recommendations, and the fraud detection (Ngai et el, 2009), **e-library** defined by Trivedi (2010) as “a library in which collections resources are stored in digital formats (as opposed to print, microform, or other media) and accessible by computers, **e-procurement** defined by Wu et el (2007) as “The use of information technologies to facilitate B2B purchase transactions for materials and services”, it involves six forms of activities, including e-ordering/ e-maintenance repair operate, web-based enterprise resource planning (ERP), e-sourcing, e-tendering, e-reverse auctioning/e-auctioning and e-informing (de Boer et el, 2002), **e-learning** defined by Bing (2001) as “an Internet-based learning process, using internet technology to design, implement, select, manage, support and extend learning” which are used in universities or in similar conditions, in the context of IoT, cloud provider is responsible for building and maintaining e-learning cloud, and providing technical support to e-learning cloud, according to Laisheng et el (2011), it leads to reduce charges for schools or enterprises and achieve economies of scale for suppliers. As Gilalouche (2013) said, pioneers in using IoT, especially in cloud computing are: *Netflix*, an American entertainment company provides streaming media and video-on-demand online, *Xerox*, an American global corporation that sells print and digital documents, *Pinterest*, a web and mobile application company that operates software system design to discover information mainly by images, uses clouds to aggregate and analyze large amounts of data, *Etsy*, an e-commerce website focused on handmade items, uses clouds to sort big data by its predictive analytics, *Apple*, an American company that designs, develops, and sells consumer electronics, computer software, and online services, uses clouds to manage the release of Siri, a program that mimics a person listening ability and offers answers to users, *MediaMath*, a digital media buying platform that develops tools for marketing managers, uses clouds for big data analytics solutions to get deeper insight into audience segmentation. Mateo (2011), said that small and medium businesses, companies in financial sector, telecommunications, and educational sector are more interested in adoption of cloud computing, through the stronger agility, efficiency, and flexibility.

7. ORGANIZATIONS, MULTI-FACED AND MULTI-LEVEL ENTITIES

As said in the previous section, e-readiness studies are more around national or individual level and less in organizational level. One of the reason for this deficit, maybe relate to multi-faced and multi-level construct

of organizations, which lead to get involved in more complex situations in individual, group, and organizational level. One of the best theory that can explain the complex and triple relations in the organizations, is the socio-cognitive theory. Though socio-cognitive theory have diverse usage in different places, in IT concept researches mentioned to the mechanism of the triple theory, the relation among environment, behavior and cognitive characteristics, networks of subjective shared to varying degrees by their members (Allard-Poesi, 1998), a complex web of relationships or social activities among people(Akgün et al, 2003), integrating the individual innovator's cognition and the innovator's social system of collectives of organizations and individuals influences digital innovation processes and outcomes(Davidson, E. J.; 2002). Environment factors are: social pressure, unique situational characteristics, and cognitive factors like: personality and demographic characteristics (Compeau et al, 1995). The network of multilevel constructs leads to have different levels of analysis and at different organizational levels (Chen et al, 2002).

8. IMPORTANT MODELS AND VARIABLES IN THE CONTEXT OF ORGANIZATIONS

Due to emerging technologies in organizations, there are models which provide the tools to measure the important variables for cutting-edge technologies. Some famous models are: Theory of Reasoned Action (TRA) by Fishbein and Ajzen (1975) with Attitude toward behavior and subjective norm variables, Technology Acceptance Model (TAM) by Davis (1989) with perceived usefulness and perceived ease of use and subjective norm variables, Motivation Model (MM) by Davis (1992) with Extrinsic Motivation and Intrinsic Motivation variables, Theory of Planned Behavior by Ajzen (1991) with attitude toward behavior and subjective norm and perceived behavioral control variables, Combined TAM and TPB (C-TAM-TPB) by Taylor and Todd (1995) with attitude toward behavior and subjective norm and perceived behavioral control and perceived usefulness variables, Model of PC Utilization (MPCU) by Thompson et al (1991) with job-fit and complexity and long term consequences and affect toward use and social factors and facilitating condition variables, Innovation Diffusion Theory by Moor and Benbasat (1991) with relative advantage and ease of use and image and visibility and compatibility and results demonstrability and voluntariness of use variables, Social Cognitive Theory (SCT) by Compeau and Higgins (1995) with outcome expectations performance and outcome expectations personal and self-efficacy and affect and anxiety variables, Technology-Environment-Organization (TOE) by Tornatzky and Fleischner (1990) with Industry characteristics and market structure and technology support infrastructure and government regulation for external task environment variable, formal and informal linking structure and communication processes and size and slack for organization variable, and availability and characteristics for technology variable, Institutional Theory by Scott and Christensen (1995) with mimetic and coercive and normative variables, Lacovou et al Model (1995) with perceived benefits and organizational readiness and external pressure variables. As Hasan et al (2015) said, in the case of cloud computing the variables which are more mentioned in the researches are sorted in this way: security, perceived usefulness, perceived ease of use then cost, compatibility, top management support, then competitive pressure, relative advantage, complexity, attitude toward technology, innovation, organizational competency, and at last need for cloud, reliability, organization type, marketing effort, social influence, lack of interoperability, convenience, entrepreneurial orientation, trust in supplier, training and education, perceived connectedness, service and system quality, Satisfaction, perceived mobility, and firm size. But there is need to more investigate in reciprocal way and use variables that can reflexes the multi-faced and multi-level nature of organizations like self-efficacy or collected self-efficacy.

9. FUTURE STUDY OF ORGANIZATIONAL INTERNET OF THINGS

In the context of organizations, there is need to have more investigation about internet of things technical and non-technical issues. Researches should focus on technical issues such as business intelligence, information security, UI and UX design, mobile and system application design, hardware engineering, networks, and also skills in programming like Python, Raspberry pi, Node.js, Java-Script, C++, Micro-Control, Arduino, IOT

consulting and analyzing, skills in embedding nodes, quality control department and laboratory, data center management, standards for applications test, and so on. Besides technical issues, some conditions like provide training courses, implementation in small scale, support creativity in designing objects, providing guidelines for users, customer services, enabling for customizing applications, feedback of users, provide security standards merged with ethical regulations should be considered too. In the other side, about the models of adoption and e-readiness, there is need to use other contexts models in organizational view and reanalyze them, so they can be reformed due to new cutting edge technologies conditions. Because the organizations are multi-faced and multi-leveled entities, the models are designed more in individual levels with static variables, but there are less models in organizational levels and also no models observed about the group level both inter group and between group relations. These kind of suggested models may be important for organizations with vigilant human functions such as hospitals or fire stations or other kind of critical tasks that needs high and strong group relations. Among the variables of adoption and e-readiness models, *coordination*, and *ethics* can be measured too by the use of internet of things nodes in the suggested group model. Combining models, increasing dimensions due to technologies content, and considering the demographic and cultural variables should not be forgotten too.

10. CONCLUSION

This paper was about the internet of things emergence, introduction of its technologies like RFIDs, WSNs, Tags, and Clouds, and also the fast widespread usage in all human aspects lives such as smart cities, smart metering, smart environment, smart water, security and emergencies, industrial control, smart retails, logistics, smart agriculture, and smart buildings. Besides the benefits of the IOT like cost reduction and vast capacity of gathering data, there are serious challenges in most usage contexts like security and privacy. Organizational studies is one of the area which needs to be considered more due to special reciprocal entities. Organizations engaged vastly with IoT in the faces of human resources, structures, infrastructures, goals and environments. IoT can be used in many applications in the context of organizations like e-procurement, e-business, e-commerce, e-library, e-infrastructure, e-learning and so on. In confronting new technologies, organizations should be prepared by e-readiness and adoption mechanisms to increase the maximum use of IOT devices. In the body of the paper, important models and the most studied variables was reviewed, including: security, perceived usefulness, perceived ease of use then cost, compatibility, and top management support. At last, the needs of organizational internet of things was presented in technical and non-technical view to cover the lack of reciprocally area of the current models, and improving nodes and objects capacities due to IoT conditions.

REFERENCES

- Ang, S.K. and Husain, W., 2012. A study on implication of adopting e-business technology by SMEs.)^(Eds.):‘Book *A study on implication of adopting e-business technology by SMEs*’(2012.), pp.5-6.
- Tornatzky, L.G., Fleischer, M. and Chakrabarti, A.K., 1990. *Processes of technological innovation*. Lexington books.
- Ajzen, I., 1991. The theory of planned behavior. *Organizational behavior and human decision processes*,50(2), pp.179-211.
- Alshawi, M., 2007. *Rethinking IT in construction and engineering: Organizational readiness*. Routledge.
- Allard-Poesi, F., 1998. Representations and influence processes in groups: Towards a socio-cognitive perspective on cognition in organization. *Scandinavian Journal of Management*, 14(4), pp.395-420.
- Akgün, A.E., Lynn, G.S. and Byrne, J.C., 2003. Organizational learning: A socio-cognitive framework. *Human relations*, 56(7), pp.839-868.
- Atzori, L., Iera, A. and Morabito, G., 2010. The internet of things: A survey. *Computer networks*, 54(15), pp.2787-2805.
- Benlian, A., Hess, T. and Buxmann, P., 2009. Drivers of SaaS-adoption—an empirical study of different application types. *Business & Information Systems Engineering*, 1(5), p.357.
- Bing, L., 2001. E-learning and modern education reform. *Education Information*, 10(25), p.21.
- Bordonaba-Juste, V., Lucia-Palacios, L. and Polo-Redondo, Y., 2012. Antecedents and consequences of e-business adoption for European retailers. *Internet Research*, 22(5), pp.532-550.

- Cartelli, A., 2007. ICT and knowledge construction: Towards new features for the socio-technical approach. *The Learning Organization*, 14(5), pp.436-449.
- Chang, V., Walters, R.J. and Wills, G.B., 2015. Cloud computing and frameworks for organizational cloud adoption. *Delivery and Adoption of Cloud Computing Services in Contemporary Organizations*, 1.
- Chen, G. and Bliese, P.D., 2002. The role of different levels of leadership in predicting self and collective efficacy: Evidence for discontinuity. *Journal of Applied Psychology*, 87(3), p.549.
- Compeau, D.R. and Higgins, C.A., 1995. Computer self-efficacy: Development of a measure and initial test. *MIS quarterly*, pp.189-211.
- Compeau, D.R. and Higgins, C.A., 1995. Computer self-efficacy: Development of a measure and initial test. *MIS quarterly*, pp.189-211.
- Crespo Márquez, A., Moreu de León, P., Gómez Fernández, J.F., Parra Márquez, C. and López Campos, M., 2009. The maintenance management framework: A practical view to maintenance management. *Journal of Quality in Maintenance Engineering*, 15(2), pp.167-178.
- Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pp.319-340.
- Davis, F.D., Bagozzi, R.P. and Warshaw, P.R., 1992. Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of applied social psychology*, 22(14), pp.1111-1132.
- Davidson, E.J., 2002. Technology frames and framing: A socio-cognitive investigation of requirements determination. *MIS Quarterly*, pp.329-358.
- De Boer, L., Harink, J. and Heijboer, G., 2002. A conceptual model for assessing the impact of electronic procurement. *European Journal of Purchasing & Supply Management*, 8(1), pp.25-33.
- Ebersold, K. and Glass, R., 2015. The Impact of Disruptive Technology: The Internet of Things. *Issues in Information Systems*, 16(4).
- Fishbein, M. and Ajzen, I., 1975. Belief, attitude, intention and behavior: *An introduction to theory and research*.
- Folk, C., Hurley, M.D.C., Kaplow, A.W.K. and Payne, J., 2015. The Security Implications of the Internet of Things. *Fairfax: AFCEA International Cyber Committee*.
- French, A.M. and Shim, J.P., 2016. The Digital Revolution: Internet of Things, 5G, and Beyond. *CAIS*, 38, p.40.
- Gillen, A., Grieser, T. and Perry, R., 2008. Business value of virtualization: Realizing the benefits of integrated solutions. *IDC, July*, 7.
- Greenberg, G., 2004. The Digital Convergence. Extending The Portfolio Model. *EDUCAUSE review*, 39(4).
- Hasan, L.M., Zgair, L.A., Ngotoye, A.A., Hussain, H.N. and Najmuldeen, C., 2015. A review of the factors that influence the adoption of cloud computing by small and medium enterprises. *Scholars Journal of Economics, Business and Management*, 2, pp.842-848.
- Helmiö, P., 2017. Open source in Industrial Internet of Things: a systematic literature review.
- Hernández-Ortega, B., 2011. The role of post-use trust in the acceptance of a technology: Drivers and consequences. *Technovation*, 31(10-11), pp.523-538.
- Hodgson, K., 2015. The Internet of [Security] Things. *SDM: Security Distributing & Marketing*, 45, pp.54-72.
- Holmberg, K., Adgar, A., Arnaiz, A., Jantunen, E., Mascolo, J. and Mekid, S. eds., 2010. *E-maintenance*. Springer Science & Business Media.
- Hunter, J., Little, S. and Schroeter, R., 2008. The application of semantic web technologies to multimedia data fusion within escience. In *Semantic multimedia and ontologies* (pp. 207-226). Springer, London.
- Joseph, N., Kar, A.K., Ilavarasan, P.V. and Ganesh, S., 2017. Review of discussions on internet of things (IoT): Insights from twitter analytics. *Journal of Global Information Management (JGIM)*, 25(2), pp.38-51.
- Iacovou, C.L., Benbasat, I. and Dexter, A.S., 1995. Electronic data interchange and small organizations: Adoption and impact of technology. *MIS quarterly*, pp.465-485.
- Libelium - Connecting Sensors to the Cloud, Date Accessed: 2017-11-23, URLs: www.libelium.com/
- Madakam, S., Ramaswamy, R. and Tripathi, S., 2015. Internet of Things (IoT): A literature review. *Journal of Computer and Communications*, 3(05), p.164.
- Mell, P. and Grance, T., 2011. The NIST definition of cloud computing
- Maksoud, S.S. and Youseff, M.A.A., 2003. Information and communication technology for small and medium enterprises in Egypt. *SME Development Unit-Ministry of Foreign Trade, Cairo*, pp.1-15.
- Mohammed, F. and Ibrahim, O., 2015. Models of adopting cloud computing in the e-government context: a review. *Jurnal Teknologi*, 73(2).
- Moore, G.C. and Benbasat, I., 1991. Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information systems research*, 2(3), pp.192-222.
- Mutula, S.M. and Van Brakel, P., 2006. An evaluation of e-readiness assessment tools with respect to information access: Towards an integrated information rich tool. *International Journal of Information Management*, 26(3), pp.212-223.

- Naseebullah, S., Dominic, P., & Khan, M., 2011. A framework of organizational e-readiness impact on e-procurement implementation. *Communications in Computer and Information Science*, 155, 240–245.
- Ngai, E.W., Xiu, L. and Chau, D.C., 2009. Application of data mining techniques in customer relationship management: A literature review and classification. *Expert systems with applications*, 36(2), pp.2592-2602.
- Pawade, P.A. and Wasukar, A.R., 2017. A Review on Importance of Internet of Things for Next Generation Computing. *International Journal of Research In Science & Engineering, Volume: 3, Issue:1*
- Rosnafisah, S., Salbiah, M.S. and Sharifuddin, A.M., 2010. Factors affecting the e-business adoption among the home-based businesses (HBBs) in Malaysia. *World Academy of Science, Engineering and Technology*, 42, pp.1236-1240.
- Suryadevara, N. and Mukhopadhyay, S., 2017. Internet of Things: A Review and Future Perspective. Space.
- Taylor, S. and Todd, P.A., 1995. Understanding information technology usage: A test of competing models. *Information systems research*, 6(2), pp.144-176.
- Thompson, R.L., Higgins, C.A. and Howell, J.M., 1991. Personal computing: toward a conceptual model of utilization. *MIS quarterly*, pp.125-143.
- Venkatesh, V. and Bala, H., 2008. Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, 39(2), pp.273-315.
- Weber, R.H., 2015. Internet of things: Privacy issues revisited. *Computer Law & Security Review*, 31(5), pp.618-627.
- Weinberg, B.D., Milne, G.R., Andonova, Y.G. and Hajjat, F.M., 2015. Internet of Things: Convenience vs. privacy and secrecy. *Business Horizons*, 58(6), pp.615-624.
- Wu, J.H., Wang, S.C. and Lin, L.M., 2007. Mobile computing acceptance factors in the healthcare industry: A structural equation model. *International journal of medical informatics*, 76(1), pp.66-77.
- Silva, B.N., Khan, M. and Han, K., 2017. Internet of things: a comprehensive review of enabling technologies, architecture, and challenges. *IETE Technical Review*, pp.1-16.
- Trivedi, M., 2010. Digital libraries: functionality, usability, and accessibility. *Library Philosophy and Practice (e-journal)*, p.381.
- Kim, P., Ng, C.K. and Lim, G., 2010. When cloud computing meets with Semantic Web: A new design for e-portfolio systems in the social media era. *British Journal of Educational Technology*, 41(6), pp.1018-1028.
- Lee, I. and Lee, K., 2015. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, 58(4), pp.431-440.
- Lian, J.W., 2015. Critical factors for cloud based e-invoice service adoption in Taiwan: An empirical study. *International Journal of Information Management*, 35(1), pp.98-109.
- Suryadevara, N. and Mukhopadhyay, S., 2017. Internet of Things: A Review and Future Perspective. Space.
- Xu, X., 2012. From cloud computing to cloud manufacturing. *Robotics and computer-integrated manufacturing*, 28(1), pp.75-86.
- GilAllouche, 2013. Well-known companies who have moved to the cloud, *smart data collective website*, available on: <https://www.smartdatacollective.com/7-well-known-companies-have-moved-cloud/>
- Gilabert, E., Jantunen, E., Emmanouilidis, C., Starr, A. and Arnaiz, A., 2014. Optimizing e-maintenance through intelligent data processing systems. In *Engineering Asset Management 2011 (pp. 1-9). Proceeding of the sixth world congress on engineering asset management*. Springer, London.
- Khajeh-Hosseini, A., Greenwood, D. and Sommerville, I., 2010, July. Cloud migration: A case study of migrating an enterprise it system to iaas. In *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on* (pp. 450-457). IEEE.
- Laisheng, X. and Zhengxia, W., 2011, January. Cloud computing: a new business paradigm for E-learning. In *Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on* (Vol. 1, pp. 716-719). IEEE.
- Lou, E.C., 2010, September. E-readiness: how ready are UK construction organizations to adopt IT. In *Proc. 26th Annual ARCOM Conference* (pp. 6-8).
- Mateo, L., 2011. Which businesses are adopting cloud computing first?. In *IBM website*, available on: <https://www.ibm.com/blogs/cloud-computing/2011/12/which-businesses-are-adopting-cloud-computing-first/>
- Nasir, U. and Niazi, M., 2011, June. Cloud computing adoption assessment model (CAAM). In *Proceedings of the 12th International Conference on Product Focused Software Development and Process Improvement* (pp. 34-37). ACM.
- Saha, H.N., Saha, N., Ghosh, R. and Roychoudhury, S., 2016, October. Recent trends in implementation of Internet of Things—A review. In *Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016 IEEE 7th Annual* (pp. 1-6). IEEE.
- Sohaib, O. and Naderpour, M., 2017, July. Decision making on adoption of cloud computing in e-commerce using fuzzy TOPSIS. In *Fuzzy Systems (FUZZ-IEEE), 2017 IEEE International Conference on* (pp. 1-6). IEEE.
- Watson, P., Lord, P., Gibson, F., Periorellis, P. and Pitsilis, G., 2008, May. Cloud computing for e-science with carmen. In *2nd Iberian Grid Infrastructure Conference Proceedings* (pp. 3-14). Citeseer.

THE POSTAL SERVICE IN THE SMART CITY

Maria Pavlovskaya and Olga Kononova

*St. Petersburg National Research University of Information Technologies, Mechanics and Optics (ITMO University)
197101, St. Petersburg, Russia*

ABSTRACT

These days the popular Smart City concept focuses on the quality improvement of a citizen's life by using the information and communication technology. Frequently Smart City projects are stalled at the initial stages due to insufficient data volume. Meanwhile, State Postal Services remain outside the Smart City concepts and models, despite the available resources to handle big data. The purpose of the study is to highlight the significance of the involvement of Postal Services into effective data collection, storage and aggregation. The study refers to the international sources and considers data-driven approach; gives an analysis of research results and case studies of International Postal Services in the field of socially significant initiatives; describes the opportunities, services and assets of the Russian Post. The paper provides an overview of the international experience of the Smart Technology usage such as Big Data, Internet-of-things, and other technologies by the Posts. It also assesses the preparedness level of the Russian Post, city state administration and citizens to use these technologies. It is suggested to examine the business and data architecture, the postal electronic services development, technological tendencies, affecting the provision of electronic services by the Russian Post. The integration of urban information systems and the Postal Services' assets could provide more efficient and ubiquitous data collection which entails additional quality to public services as well as improves the city authorities results and ensures the safety of information resources, helping all participants to achieve their goals. Such approach is named as Internet-of-Postal-Things. The paper describes the research results which purpose was to identify the opinions of the citizens from different regions and age groups about Smart projects and the opportunities of the Russian Post participation in these projects. The paper also provides recommendations on activities that promote the involvement of the Russian Post into innovative projects.

KEYWORDS

Big Data, Data-Driven City, IoT, Internet-of-Postal-Things, Postal Electronic Services, Smart City Architecture

1. INTRODUCTION

The majority of the currently implemented Smart Cities does not include Postal Authorities in the list of participants. Nevertheless, the justification for the relevance of the Post participation in such projects is confirmed by studies conducted in different countries of the world. One of the important factors explaining this position is the high density of presence and significant material assets of the Postal Services (Universal Postal Union, 2015; Oxford Strategic Consulting Report, 2017; Postal Services in the Digital Age, 2014). The use of the modern technologies underlying the digital economy by the Postal Services to provide a variety of digital services, the preparedness of the Post to become a full participant in the Smart City projects is confirmed by the practice of the Universal Postal Union, consulting and IT companies, as well as the practical activities of the postal operators themselves.

Analysis of the current Russian Post services and initiatives, the state of the business and data architecture have revealed the absence of the significant activity in the interaction with the city authorities for collecting and processing of the data of the urban performance, including involving Big Data and Internet-of-Things concepts. The dynamics of the Russian Post development in recent years has led to the conclusion about the possibility of changing of the role of the Russian Post services in the society. According to the conclusion of the Russian Post prepared for the Universal Postal Union, among the technological trends that will affect the provision of electronic services in the coming years, Big Data and Internet-of-Postal-Things are particularly emphasized. By analogy with the Internet-of-Things technology, the described approach could be named as the Internet-of-Postal-Things (IoPT). IoPT is instrumenting the postal infrastructure with low-cost sensors to

enable them to collect, communicate, and act upon a broad variety of data. It could help the Postal Service to generate operational efficiencies, improve the customer experience, and develop new business models.

There are many reasons why the Postal Service would become an equal partner in Smart City projects, but the first postal network characteristic of great importance is its omnipresence. The postal assets can be broadly divided into three main categories, such as stationary assets, transport fleet and carriers. In Russia, the category of stationary objects includes post offices, collection boxes and home mailboxes which have been installed throughout the country. The transport fleet of the Russian Post presented by various vehicles including automobiles, trains, plains, and others. The total length of main and internode mail routes exceeds 2.8 million km. The structure of the Russian Post assets is shown in Table 1.

Table 1. The Russian Post assets

Postal assets' categories	Postal assets	Number of objects
Stationary objects of the Russian Post	Post offices	42.000
	Collection boxes	over 140.000
	Home mailboxes in the residential sector	Moscow – 3.5 million; St. Petersburg – 1.5 million; Nizhny Novgorod – 0.5 million
Transport fleet of the Russian Post	Automobiles	14.000
	Other vehicles	3.000
Couriers	Couriers and postmen	about 100.000 (The Russian Post, 2015)

The next important characteristics of the postal network are frequency and consistency which are more relevant for dynamic types of the assets. Since the Postal Service is considered as a universal service in the most countries, Post offices are situated even in remote and sparsely populated communities; their vehicles pass through almost every road, including the roads that bus routes may not cover. Such wide coverage by one enterprise allows Smart City projects a degree of flexibility in their scope. Data could be collected nationally or along specific areas (Ravitzky, 2011). Accordingly, the data collection potential of a large number of vehicles overcoming significant distances is enormous and create a powerful information network.

The strategy of the Russian Post implies digital transformations that will make the Russian Post a profitable, customer-oriented and technologically advanced company. Digitalization of the Post allows to conclude that the company is ready to use the results of world experience in applying modern technologies to expand the spectrum of the provided products (Pavlovskaya and Kononova, 2017).

The relevance of the work is further conditioned by government activity in the development of Smart Cities in Russia which is caused by "Russian Digital Economy" document release, where the Smart City direction is allocated. In accordance with the document city authorities, scientific and business communities started the "Smart Saint Petersburg" urban innovation program, where the research results could be used.

Recommendations on the involvement of the Russian Post in Smart City projects, developed on the basis of an architectural approach, taking into account the opinions of the citizens and employees of the enterprise, will facilitate the organization of the effective interaction with participants in smart urban programs.

2. POSTAL SERVICES AND SMART TECHNOLOGIES

2.1 Postal Services and Technologies used in Smart City projects

Postal operators around the world already actively participate in the providing of public services to citizens and in the development of socially significant initiatives. The most advanced Postal Services try to turn the social responsibility of the enterprise into one of the most important components of its effective work and innovations. Relative to the sustainable development objectives set by the United Nations, the postal sector is to become one of the major players and over the time contributes to the implementation of initiatives. Introducing innovations, the postal sector constantly undertakes obligations to support social and environmental changes, expresses the preparedness to place the experience in the service of society.

It is obvious that big data guarantees the good result from adoption of administrative decisions. A data-driven city is characterized by the ability of municipal authorities to use data collection, processing and

analysis technologies to improve the social, economic, environmental situation (PWH, 2016; Kupriyanovsky and Martinov, 2014; Narmeen and Jawwad, 2015; Robinson, 2012; U.S. Postal Service, 2017).

Therefore, some of the Postal Services have already begun to explore sensor-based data collection. In 2014, Spanish Post was involved in developing air quality monitoring sensors for postal vehicles (Grupo Correos, 2014). Finnish Postal Service has begun to conduct experiments on how sensor-based collected data (for example, road conditions, traffic flow) could be used (Posti, 2015). French postal operator is taking a tack, aiming to be a platform where sensor-based data from variety sources can be housed together securely. Under this model, La Poste plays the role of data broker, offering storage and analytics services (Docapost, 2016).

Smart mailboxes are becoming more and more popular. They are equipped with Wi-Fi, work on solar panels, inform the recipient at the time when the delivery has been done to his mailbox and use an intelligent locking system. Moreover, these mailboxes are connected to the mail processing center and receive data on the mailman's schedule and send information about empty boxes for optimization daily routes for collecting items. A user with a mobile application can remotely monitor what is happening with the mailbox.

Mobile technology has already helped redefine the role of carriers, expanding the variety of their tasks. US Post has equipped carriers with mobile devices to facilitate scanning packages at delivery and communicating with the office. Handheld devices could become a platform for a variety of other activities, such as collecting sensor data and interacting with citizens (U.S. Postal Service, 2015).

Postal operators also implement various monitoring services. The "lost and found" service implemented in Denmark, where Postal vehicles help identify stolen bicycles. A sensor embedded in the bicycle automatically registers its location through the closest Postal connected device in the vicinity. This approach could be extended to monitoring the status of the city infrastructure such as road conditions or street lights.

Several Posts have created passive and active "check on" services, whereby carriers regularly visit elderly people. As a part of new Japan Post/IBM/Apple partnership, these clients will receive iPads with apps to be connected with services, health care, community, and their families (Etherington, 2015). The interconnection of sensor data from the elderly citizen's and the smart device carrier could be a key to the effective provision of innovative check of services. For instance, the system schedules the visits, alerts the client that the carrier is on his way, enables the timely delivery of medication, or reports to family or local healthcare authorities.

Table 2 presents a general view of the world experience of the involvement of Posts in socially significant initiatives. The data indicates high levels of the Post involvement in the state services provision, acceptance of payments, issuing identifiers for accessing to electronic services through the post offices (OIG, 2015; Poste Italiane, 2017). Among environmental initiatives of cities, where the Postal Services participate, one of the most important issue is transportation with fuel efficient vehicles, such as electric vehicles or bicycle couriers, public transportation system that uses driverless buses to carry passengers, bike share programs etc. (Chourabi et al, 2012; Giffinger and Pichler-Milanović, 2007; Post and Parcel, 2015; Swiss Post, 2015).

Table 2. World experience of the involvement of Posts in socially significant initiatives

The Post	Service / Initiative						
	State services	Acceptance of payments	Environmentally friendly transport	Data collection	Storage and analysis of the data	Monitoring services	
Italy	+	+	+				
Switz.	+	+	+				
Spain	+	+	+	+			
Finland	+	+	+	+	+		
France	+	+	+	+	+		
Norway	+	+	+				
Germany	+	+	+	+	+		
Denmark	+	+	+	+		+	
USA	+	+	+	+	+	+	
Japan	+	+	+				+
Russia	+	+					

The whole range of socially significant initiatives has not been implemented even in the USA, which Postal Service is recognized as the best in the top 20 largest economies in the world (Postal Services in the Digital Age, 2014). At the same time, the least attention is paid to the collection, storage and analysis of data in the various spheres of the city economy and the main activity of the Post Services.

2.2 The Russian Post Readiness Level to use Smart Technology

The practice of the Russian Post enterprise architecture management is based on the product-service model, where the modeling of the current and target architecture is performed in the context of products (the activity of the enterprise in providing products to customers) and services (the internal activity of the enterprise). The product landscape of the Russian Post on a conceptual level is shown in Figure 1.

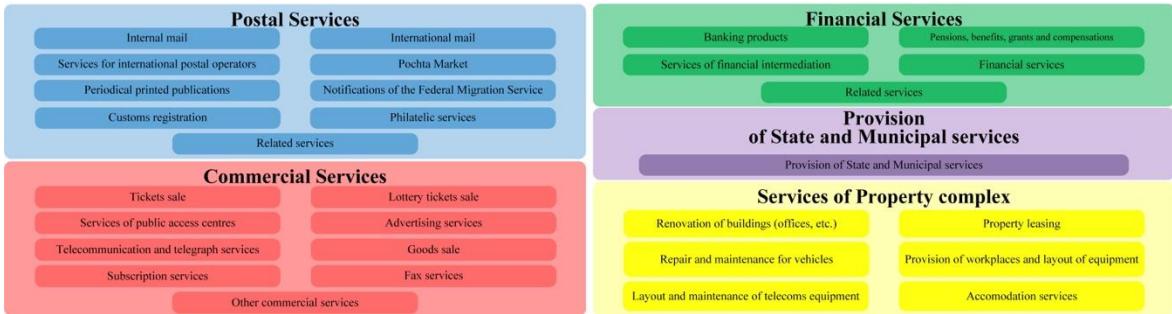


Figure 1. The Product Landscape of the Russian Post

The revenue structure of the enterprise shows that the most profitable issues are the postal (domestic and international written correspondence, parcels and EMS) and financial products as well as commercial services (sale of goods and subscription services). Internal activity of the enterprise is also extremely important primarily because of the activities related to production and logistics. This is the main activity of the Post, they spend the main production capacities and resources of the enterprise, the productivity of the Postal Service as a whole depends on the quality of their provision (The Russian Post, 2015).

Currently, the data architecture of the Russian Post is under development. At the same time, both scattered descriptions of data generated and received by some information systems or organizational units of the enterprise and data flows identified as a result of modeling products and services of the enterprise are available. An album of data formats and data models of subject domains is created. The data identified during the implementation of information systems and collected in the corporate data warehouse is classified.

While basing on the information about the current state of the Russian Post is impossible to give an unambiguous assessment about the impossibility of significant changes in the status and the role of the enterprise in the society. Table 3 presents the conclusion of the Department of Enterprise Architecture of the Russian Post in terms of technological trends that will influence on the provision of electronic services in the enterprise, prepared as a part of the Universal Postal Union report. For each factor, the impact upon the scale 1 to 5 is given, where 1 is the absence of influence, and 5 is a significant influence.

Table 3. Technological trends that will affect the Russian Post

Technological trends	1	2	3	4	5
Big data, data analytics and cloud computing technologies					x
The new generation of portable terminals for postmen				x	
Sensors for postal infrastructure (postal vehicles, mailboxes) – Internet of Things					x
New payment technologies				x	
New improvements in e-health and services for the elderly			x		
Drones for delivery	x				
Unmanned vehicles or autonomous robots for delivery	x				
Crowdshipping	x				
Security standards and technologies in cyberspace				x	
3D Printing Technologies			x		
Blockchain technology (identification information, logistics, virtual currency)			x		
Augmented reality or virtual reality, for example, in sorting centers or post offices			x		

We could conclude that the greatest prospects for use in the Russian Post are Big data, Internet of Things, mobile devices of a new generation for employees, new payment technologies, security in cyberspace, augmented reality or virtual reality. The Russian Post and various government agencies are also constantly announcing the launch of new Postal Services that support new technologies. At the moment, the Russian Post is working on a technology for identifying customers by the face at the entrance of the post office.

3. "SMART CITY AND THE RUSSIAN POST" RESEARCH

3.1 Purposes and Issues

The influence of public opinion on the government activities is constantly increasing, and the need to adopt socially-oriented management decisions, taking into account the interests of various social groups, is beyond doubt. Sociological analysis of public opinion contributes to raising the level of conceptual interpretation of socio-economic reforms, optimizing the activities of power structures. One of the most important components of Smart City which can make the existence of such cities possible is Smart Citizens (Allessie, 2016; Forbes, 2014). Moreover, the problem of the need to change the mentality of people living in cities claiming to be "smart" is crucial. Citizens should be ready to use the initiatives introduced by the city authorities, actively participate in the formation of needs for the introduction of such initiatives and in management (Vidyasova, 2017). This challenge pushes to the background the problems of choosing and implementing technologies. Initiatives should also have upward character, because the approach of the 'ideal' Smart City with a downward orientation, which is ubiquitous today, destroys democracy and often minimizes the involvement of citizens (Allessie, 2016). Thus, public opinion is an important factor in the decision making of city authorities. The study of its influence should be the first stage of the research on the development of proposals for involving the Russian Post in such projects.

The purpose of the study is to identify trends in relation to the Smart City projects and the participation of the Russian Post in such projects of citizens of different regions and age groups. At the same time, special attention is paid to identifying the relationship, assessing the potential of the Russian Post in the implementation of "Smart City" and "smart technology" projects from the staff of the Russian Post. The study can be considered as exploratory and aimed at the confirmation of hypotheses and propositions. Based on the results, priorities for the further development of ongoing research will be determined. Within the framework of the research, an online interviewing of Russian citizens including employees of public authorities and subordinate organizations, representatives of the scientific and business communities, employees of the Russian Post have been conducted. The research tool is the online questionnaire that contains 30 questions and consists of 4 parts.

The rationale for the sample is based on the ideas of the approach formulated by Everett Rogers (2003). The general population is hyper-digital users – students, employees of the Russian universities, employees of IT companies (IT divisions of the various companies). An important role is also given to the views study of the transport/logistics employees, in particular, employees of the Russian Post and representatives of the public administration and managers at different levels. Quotas have been maintained by sex, age, position level and profession. The sample has been purposive and formed by "the snowball method" through the social connections of the researcher.

3.2 Survey Results

Among the respondents the age categories up to 35 years prevail, which is determined by the specificity of the target group. The age categories up to 35 years prevail, 15% of respondents classifies themselves into the age categories of 35-44 and 45-54. 21% of respondents refer the organization they work/study in to the professional field of postal activity, 20% – to the IT field, 19% – to the science and education, 10% – to the public service. The majority of respondents (69%) highly assesses their skills in using information technology in their daily and professional activities.

In the survey the global trends in the digital technologies, Russian digital technologies, and their use in various sectors of the Russian economy have been researched. The respondents' meanings in the fields mentioned above have been assessed and an average level of citizens' awareness has been demonstrated. The respondents suppose that the implementation of the Smart City project in their region will improve the quality of their family life rather noticeably (85%). Slightly less than a half of the respondents (46%) refer the organization they work or study in to the participants of the city socially significant projects based on digital technologies, and 20% of respondents state that their organization is a direct participant in the projects "Smart City". 23% of respondents say that their organization uses modern technologies, but there is no interaction with participants in urban life. The respondents primarily consider that the main participants in the

"Smart City" projects are the government bodies and their subordinate structures (84%); the next three positions are occupied by business, scientific and public organizations. It should be noted that the missed position in the questionnaire - the "Citizens", which should certainly be presented, is the most mentioned by the respondents' open answers.

The leading position among the new technologies that can serve as the basis for the services provided by the Russian Post is occupied by new payment technologies (57%), analysis of big data (45%), augmented reality (35%), Internet of Things (35%), face recognition technologies (35%), autonomous robots (28%), unmanned vehicles (27%) and drones (22%) for delivery of mail items (Figure 2). It is important to note that most of the mentioned technologies are currently considered by the Russian Post as promising. New payment technologies, big data analytics, augmented reality technologies and face recognition technologies are already being used in the projects developed by the Post. Moreover, the opinion of the Russian Post authorities regarding the importance of the technologies corresponds to the respondents' assessments.

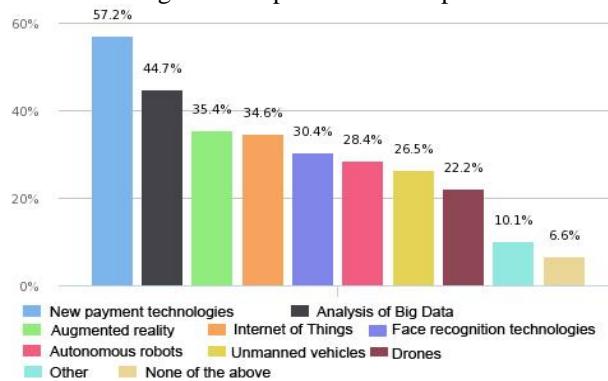


Figure 2. Distribution of the answers to the question "Which modern technologies, in your opinion, can form the basis of the services provided by the Russian Post?"

The majority of the respondents agree that the involvement of the Post in the Smart City projects is possible, but the opinion regarding the appropriateness of such projects is equally divided. Respondents suppose that the postal offices, vehicles and employees are the most demanded assets of the Russian Post in the Smart City projects (Figure 3). According to results, the Russian Post should organize data collection primarily for the postal sphere. Close indicators are mentioned for transport/housing and utilities (Figure 3).

Respondents who refer to the organization they work in, to the professional field of postal activities, are asked several additional questions. 90% of them sure that participation in Smart City projects corresponds to the strategic goal of the Russian Post - to become a profitable, customer-oriented, efficient and technological company, a modern provider of postal, logistics and financial services. Moreover, 83% of respondents are ready to take part in the activities within the framework of the "Smart City" projects.

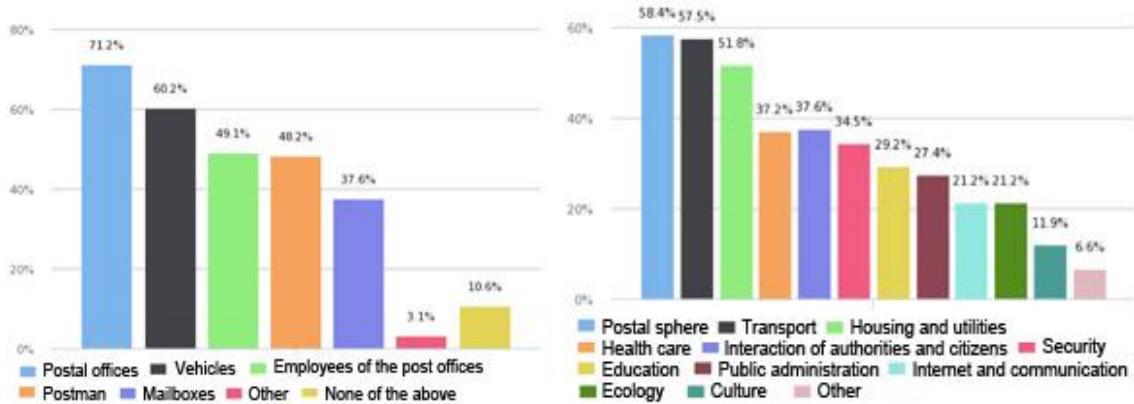


Figure 3. Distribution of the answers to the questions "Which assets of the Russian Post, in your opinion, can be used to implement the Smart City projects in Russia?" and "For the development of which areas Russian Post should organize the data collection in the frames of the "Smart City" project in your region?"

4. CONCLUSION AND RECOMMENDATIONS

The revision of the urbanization opportunities provides updated information on city growth and global transformation processes. Recognizing the importance of information technologies, the revision expands the Smart City architecture, takes into consideration the postal assets and services. The study has showed the Russian Post mission corresponds the Smart City vision. The Russian Post also support the trend of data-driven enterprises that allows considering it as a potential partner in the construction of data-driven cities.

It can be concluded that there are many opportunities to involve the Russian Post in the processes of building and managing Smart Cities, which are mainly related to the use of postal assets to collect large amounts of urban data. The Internet of Postal Things technology is able to provide the postal infrastructure with low-cost sensors to enable them to collect, communicate and act upon a broad variety of data. It could help the Postal Service, generate operational efficiencies, improve the customer experience and develop new services and business models. Smart City projects exist in the interests of cities and citizens. Having become involved in these projects, the Russian Post could translate this interaction into the cost savings of a city, improving its activity efficiency, promoting its sustainability plans, strengthening the role as a national public infrastructure and service provider. The Russian Post will accomplish its mission and achieve its goals.

It is also important to apply the architectural approach to the construction of urban agglomerations. It allows looking at the city as a complex system. Then the participation of new enterprises, such as the Postal Service, in the city projects will become easier and be in accordance with the interests of all stakeholders. At present, there is no need for significant changes in the Postal Services and its functional capabilities for taking part in urban initiatives. It is enough to add the capabilities of the Post to the existing city architecture and to update the communication between the key city domains by connecting them to the Postal Services.

The relevance of the study is further conditioned by increased government activity in Smart Cities development in Russia, which is caused by "Digital Economy of the Russian Federation" release, where exist the Smart City direction. In accordance with the document, the city authorities, scientific and business communities started "Smart Saint Petersburg" program (2017), where the research results could be used.

Hypotheses of the study are confirmed. It has been revealed that the Smart City projects have ambiguous support from the population. The attitude of respondents to Smart Cities and technologies depends on their level of awareness about the latest technological trends, the level of competence of respondents in the field of information and communication technologies. The attitude of citizens towards the possibility of the Russian Post participation in innovative projects and initiatives varies from "restrained" to "critical". At the same time, a high level of interest of the expert group (employees of the Russian Post) is established to participate in large-scale, high-tech city initiatives. As the recommendations made on the results of the study, it is necessary to emphasize the importance of raising the awareness of the population and employees of the Russian Post about smart projects, concepts and technologies. Employees of the Post should also be informed about the possibilities of the enterprise involving in Smart City projects in order to increase their efficiency.

Correctly selected recommendations on the involvement of the Post in the Smart City projects, developed within the framework of the architectural approach, taking into account the opinions of citizens and employees of the enterprise, can facilitate the organization of the effective interaction between stakeholders.

The world experience in the creation of systems in various spheres of human activity says that any initiative to create an Intelligent City must be accompanied with the set of basic documents, in particular the architecture of the Smart City. Involving the Russian Post in the Smart City projects leads to the change in the different levels of its architecture. The importance of architectural principles, which are one of the main components of the architecture, as well as the need to develop them at the initial stages of architecture creation, is confirmed by leading methodologies and standards. In connection with this, the important recommendation is the development of the principles of the Smart City projects with the participation of the Post. It is planned to develop such principles, primarily affecting the business and data architecture domains, as a part of the further study in accordance with the methodology of Griffhorst&Proper and TOGAF.

For the Russian Post, there are many opportunities to expand the range of its products and services that support socially significant initiatives and initiatives of city authorities, following the example of world practice. It is necessary to understand the potential of the Post to meet the state interests and develop proposals for strengthening the role of the Post in the public administration system. As a direction for future research, we should highlight a more detailed study of the data required to collect in the framework of Smart City projects and the possibility of its collection and aggregation using the assets of the Russian Post.

REFERENCES

- Allessie, D., 2016. *Only Smart Citizens can enable Smart Cities*. KPMG Innovative Startups. Available at: <https://home.kpmg.com/nl/en/home/social/2016/02/only-smart-citizens-can-enable-true-smart-cities.html> (Accessed 2018/02/21).
- Chourabi, H. et al, 2012. Understanding Smart Cities: An Integrative Framework. *Proceedings of the 2012 45th Hawaii International Conference on System Sciences*. Hawaii, USA, pp. 2289–2297.
- Docapost (2016) *Digital Hub and IOT*. Available at: <http://en.docapost.com/solutions/digital-hub-iot> (Accessed 2018/02/21).
- Etherington D. (2015) *Apple and IBM Team with Japan Post to Address the needs of an aging population*. TechCrunch. Available at: <http://techcrunch.com/2015/04/30/apple-ibm-japan-post/#.appdc9:DFfD> (Accessed 2018/02/21).
- Finger, M. et al, 2014. *Postal Services in the Digital Age*. IOS Press. Available at: <http://ebooks.iospress.nl/volume/postal-services-in-the-digital-age> (Accessed 2018/02/21).
- Forbes (2014) *Smart Cities – A \$1.5 Trillion Market Opportunity*. Available at: <https://www.forbes.com/sites/sarwantsingh/2014/06/19/smarter-cities-a-1-5-trillion-market-opportunity/#289b18576053> (Accessed 2018/02/21).
- Giffinger, R. and Pichler-Milanović, N., 2007. *Smart cities: Ranking of European medium-sized cities*. Vienna University of Technology, Vienna, Austria.
- Grupo Correos (2014) *2014 Integrated Annual Report: Innovation and Technology, Our Present, Our Future*. Available at: <http://www.correos.es> (Accessed 2018/02/21).
- Kupriyanovsky, V. and Martinov, B., 2014. “Smart planet” – how to do it?. *ArcReview*. No. 68. Available at: https://www.uspsoig.gov/sites/default/files/document-library-files/2015/rarc-wp-13-003_0.pdf (Accessed 2018/02/21).
- Narmeen, Z. and Jawwad, A., 2015. Smart City Architecture: Vision and Challenges, (*IJACSA International Journal of Advanced Computer Science and Applications*, Vol. 6, No. 11, pp 246-255.
- OIG (2013) *E-Government and the Postal Service*. Available at: https://www.uspsoig.gov/sites/default/files/document-library-files/2015/rarc-wp-13-003_0.pdf (Accessed 2018/02/21).
- Oxford Press (2017) *Oxford Strategic Consulting Report*. Available at: <http://www.oxfordstrategicconsulting.com/wp-content/uploads/2017/03/Delivering-the-future-16-03c.pdf> (Accessed 2018/02/21).
- Pavlovskaya, M. and Kononova, O., 2017. The usage of the postal assets in the secure information space of a Smart city. *Proceedings of SPOISU 2017 conference (Regional informatics and information security)*. S.Petersburg, Russia, pp. 22-24.
- Post and Parcel (2015) *Norway Post to add 330 Electric Vehicles to Fleet*. Available at: <http://postandparcel.info/65771/news/environment-news-2/norway-post-to-add-330-electric-vehicles-to-fleet/> (Accessed 2018/02/21).
- Poste Italiane (2017) *Sportello amico*. Available at: <http://www.poste.it/ufficio-postale/sportello-amico.shtml> (Accessed 2018/02/21).
- Posti (2015) *Annual Report 2015*. Available at: http://annualreport2015.posti.com/filebank/1229-POSTI_Annual_Report_2015.pdf (Accessed 2018/02/21).
- PWH (2016) *Data-Driven City: from the concept to application-oriented decisions*. Available at: https://www.pwc.ru/ru/government-and-public-sector/assets/ddc_rus.pdf (Accessed 2018/02/21).
- Ravnitzky, M., 2011. *Offering Sensor Network Services Using the Postal Delivery Vehicle Fleet. Reinventing the Postal Sector in an Electronic Age*. Edward Elgar, USA.
- Robinson, R., 2012. *The new architecture of Smart Cities*. Available at: <https://theurbantechologist.com/2012/09/26/the-new-architecture-of-smart-cities/> (Accessed 2018/02/21).
- Rogers, E., 2003. *Diffusion of innovations (5th ed.)*. Free Press, New York.
- “Smart Saint Petersburg” program (2017) Available at: <https://www.petersburgsmartcity.ru> (Accessed 2018/02/21).
- Swiss Post (2015) *The Pleasure of Simple Solutions: Annual Report*. Available at: http://annualreport.swisspost.ch/15/ar/downloads/geschaeftsbericht_konzern/en/E_Post_GB15_Geschaeftsbericht_WEB.pdf (Accessed 2018/02/21).
- The Russian Post (2015) *The development strategy of the federal state unitary enterprise "The Russian Post" for the period until 2018*. Available at: <https://www.pochta.ru/about-documents> (Accessed 2018/02/21).
- U.S. Postal Service (2015) *The Internet of Postal Things*. Available at: https://www.uspsoig.gov/sites/default/files/document-library-files/2015/rarc-wp-15-013_0.pdf (Accessed 2018/02/21).
- U.S. Postal Service (2016) *The Postal Service and Cities*. Available at: <https://www.uspsoig.gov/sites/default/files/document-library-files/2016/RARC-WP-16-017.pdf> (Accessed 2018/02/21).
- Universal Postal Union (2015) *Measuring postal e-services development*. Available at: http://www.upu.int/uploads/tx_sbdownloader/studyPostalEservicesEn.pdf (Accessed 2018/02/21).
- Vidyasova, L., 2017. Conceptualization of the “smart city” concept: socio-technical approach. *International Journal of Open Information Technologies*. Vol. 5, No. 11, pp 52-57.

IMPROVING BIM WITH VR IN CONSTRUCTION

Alcínia Zita Sampaio

University of Lisbon, IST, Dep. Civil Engineering, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

ABSTRACT

Building Information Modelling (BIM) is defined as the process of generating, storing, managing, exchanging, and sharing building information. The potential of BIM methodology to support a transformation of the processes of design and construction has been evident in the construction industry. A current topic that requires attention is the integration of BIM with Virtual Reality (VR) where the user visualizes a virtual world through interactive devices or a total immersion. VR combines several devices for interaction, creating virtual environment, and this must be followed by studies concerning how to use devices or how to establish links for the presentation of information contained in a BIM model. By adding VR, the BIM solution can address retrieving and presenting information and increasing efficiency on communication and problem solving in an interactive and collaborative project. BIM + VR allow two main capacities: walkthrough and consulting data, and currently BIM tools allow links to VR plugins in order to achieve both capacities. The text presents a review of actual perspective of the VR use applied over 3D/BIM models to support multi-dimensional BIM applications, namely, 4D/BIM and 7D/BIM models. The objective of the study is to report the improvement of BIM uses with the addition of interactive capacities allowed by VR technology.

KEYWORDS

BIM methodology, VR technology, 4D/BIM construction, 7D/BIM maintenance, VR+BIM applications

1. INTRODUCTION

Building Information Modelling (BIM) methodology enables project stakeholders to create information-rich virtual models that help better visualize building projects. BIM is defined as the process of creating building information, stored in a centralized virtual model, that can be managed allowing exchanging, and sharing building data in an interoperable and reusable way (Azhar *et al.*, 2008). BIM provides an opportunity to analyse a model namely for: constructability conflicts; ways to reduce energy usage; validating energy reducing design ideas; or improving life safety. BIM represents the procedure of development and use of a computer generated model to simulate the planning, design, construction and operation of a facility (Eadie, *et al.*, 2013). The resulting product, a centralized BIM model, is a data-rich, intelligent and parametric digital representation of the building project. So, BIM can first be considered as a digital representation of a building, an object-oriented three-dimensional (3D) model, but also a repository of project information to facilitate interoperability and exchange of information with related software applications. Therefore, BIM data-rich model allows views and data, appropriate to various users' needs. BIM data can be extracted and analysed to generate information that can be used to make decisions and to improve the process of delivering the project or the facility.

BIM can generate and maintain information produced during the whole lifecycle of a building project, from design to maintenance, and can be applied to various fields (Chen and Luo, 2014). Due to the consistency of design data, with quality data and construction process with quality control process, the potential of BIM implementation in quality management lies in its ability to present multi-dimensional data including design data and time sequence. BIM and its applications in project management are considered nD/BIM models, namely:

- 3D/BIM model refers to all 3D building components (architectural, structural, mechanical, electrical, etc.) and it incorporates all the building aspects, including geometry, spatial relationships, properties and quantities;

- 4D/BIM model concerns the construction process that can be visualized by building the 3D product model through time according to the critical path network (the model supports dynamic construction site safety management, preparation of schedules and estimates, tracking and managing changes, and managing site logistics) (Sampaio and Mota, 2016);
- 5D/BIM model is related with costs (take-off material quantities, cost planning and estimating, safety checking integration for dynamic safety analysis);
- 6D/BIM model is associated to support management facilities and maintenance along the post occupation lifecycle of the building;
- 7D/BIM model is created to support management facilities and maintenance along the post occupation lifecycle of the building (Sampaio and Simoes, 2017).

The nD directions of BIM use must be based in an adequate relationship between the team members improving an adequate collaborative project, supported in an efficient interoperability of specific software. As per the reports, companies began seeing faster project approvals, increased positive team interactions and higher product quality. A current BIM topic that require state of the art technology, is the integration of BIM with Augmented Reality (AR), combining the real world with virtual objects (partial immersion), and with VR, where the user visualize a virtual world that is not real (total immersion), through special glasses. This field is interrelated with other fields that can make use of the visualization allowed by the BIM model, such as facilities management related with the visualization of data included in a BIM model, in real time following the interaction made possible by VR technology. As such, it is expected to be further explored in the near future. The objective of this text is to report the improvement of BIM uses with the addition of interactive capacities allowed by VR technology. The text analyses the degree of achievement allowed by the actual software to perform each aspect combining BIM and VR, mainly in two important applications: 4D model supporting construction activity; 7D model concerning maintenance. A bibliographic and software research was made to support the study.

2. VR AND BIM CAPACITIES

Virtual Reality (VR) technology has been improving BIM allowing interaction with 3D/BIM models in two essential ways:

- **Walkthrough** and is the most popular as the user can view the 3D model in a virtual environment in real time from multiple perspectives, from inside or outside of the building;
- **Consulting data** concerns the possibility to consult the information centralized in a BIM model.

The VR technology leads to a better communication for key players across the building sector, based on greater design visualization, contributing to a better understanding of the project. The knowledge about the different type of data that can be associated to the set of parameters that identifies each parametric object, base of the BIM modelling process, is of great importance when the team members wants to analyse problems and to discuss alternative solutions.

The VR context is interrelated with other fields that can make use of the visualization allowed by the BIM model, such as facilities management related with the visualization of data included in a BIM model, in real time following the interaction made possible by VR technology.

2.1 Walkthrough

Currently there are various applications of virtual reality tour applied in the field of construction. To allow a tour around or inside of a BIM model, first it is necessary to generate the VR tour contents, and for that BIM software like Autodesk Rabbit and 3ds Max for visualization and rendering works can be used (BIM Grid, 2017). Revit files as *fbx* format can be imported into 3ds Max, and render with appropriate materials and settings for virtual space can be applied. An experience VR content can be non-immersive, by using just a desktop or tablet PC, or immersive, by using a head-mounted display like Gear VR (Rho, 2015) (Figure 1) or Oculus Rift (Rift, 2017).



Figure 1. Improving BIM with Gear VR glasses in a desktop and in a work place

Oculus can be used to replicate an entire building and give a better perspective of the actual scale of the project, which is impossible using traditional monitor oriented systems and offers a wide viewing angle of up to 110° (Rho, 2015). VR glasses are handling to visualize and experience 3D model, 360-degree panorama picture, and virtual mock-up over a BIM model.

In construction activity the use of VR capacities brings great potentials. Presenting BIM models of projects in VR environment redefines communication and collaboration in the field and in the office. BIM technology and VR have the ability to innovate the building industry. At a first glance, many feel the benefits of using a BIM model with VR are purely for marketing leveraging the 3D model for visual aids, but taking a deeper look at a BIM model it will reveal many practical reasons to adopt BIM with VR. Collaboration in VR can be the future of VR BIM.

Currently architects, engineers and other construction specialists are already exploring the length and breadth of this joint technology (Figure 2). There are almost 50+ VR software and hardware tools available right now, that can redefine the way BIM is implemented and interpreted (Viatechnik, 2017). The checklist can be prepared smoothly so as to save valuable time and money for the company as well as the client.



Figure 2. Virtual reality tour inside and outside a BIM model and supporting a collaborative meeting

CAVE-like platforms have been developed for immersive VR experience as they track user's head and control wand usually with 6 degrees of freedom, to navigate inside the virtual environment and interact with the contents. But also VR head-mounted displays (HMDs) such as the Oculus Rift have the capacity to improve the way architects design and communicate buildings before they are built. The wearer is instantly immersed in a true three dimensional environment that gives an incredible sense of scale, depth and spatial awareness that simply cannot be matched by traditional renders, animations or physical-scale models. The sense of presence was overwhelming.

In addition, there are applications like Smart Reality (SmartReality, 2017) that can be downloaded for iOS and Android and used to check and analyse distinct aspects of the project in any place and with any member of the team, in an easy and collaborative way. It is designed for the AEC (Architecture, Engineering and Construction) sector and works with BIM tools like Autodesk's Revit. The app helps to make the communication process easier and more accurate leading to better efficiency and cost savings for AEC projects. The VR capacity of this app is beginning to introduce a new level of certainty for the design and construction process as well, helping stakeholders to get a more full experience of the BIM model (Figure 3). The viewing experience can further be enhanced by using VR headsets like Oculus Rift and Epson's Moverio BT-200 smart glasses (Sims, 2015). This app brings a whole new realism of tools to interact with clients, engineers, contractors, and other professionals in the building industry.

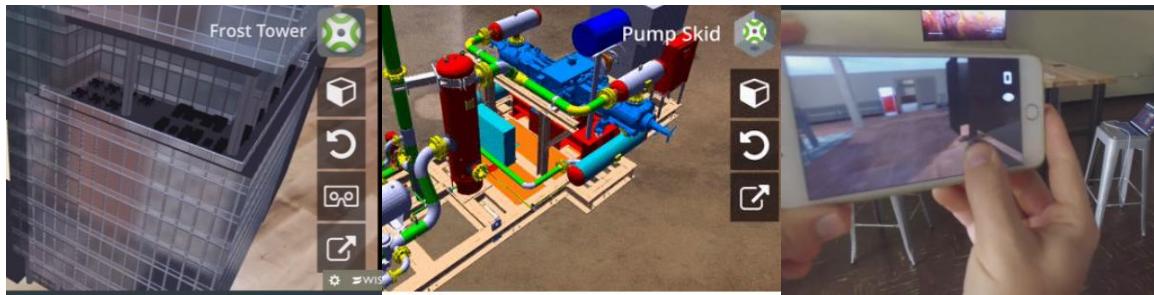


Figure 3. Smart Reality application (SmartReality, 2017)

2.2 Consulting Data

The basic concept of BIM is a consistent 3D model of buildings containing all main data as base for collaboration for all disciplines. Compared to dedicated AEC design review software like Navisworks or Tekla BIMsight, equivalent tools inside a VR environment are still very much in their beginning. This is particularly true of game engine VR experiences, where the onus tends to be on presenting a polished vision of a proposed building, rather than delivering practical tools for solving real world design and construction problems. When moving from Revit to Autodesk Stingray (Stingray, 2017), data is not only retained, but users can click on objects and view the underlying attribute information. In VR mode, users can click on an object to view information, including metadata.

Fully interactive VR software also has extremely high performance demands, so some form of model optimisation is required when bringing BIM data into a VR environment. This is one area where specialist VR consultancies earn their keep with finely tuned processes for tasks like simplifying geometry, adding lighting, fixing gaps in the model and culling objects that will not be visible in the scene. Once the model is inside the VR environment, things like materials, lighting, furniture and other small details that make the VR experience feel real are added:

- An accessible high end visualisation and virtual reality of a BIM model is obtain using the VR plugin of Revit, the **Enscape** (Enscape, 2017). Enscape is a VR and real-time rendering plugin for Revit. Inside Revit is possible to access the plugin Enscape and start to walk through the fully rendered project, without uploading to cloud or exporting to other programs. The user can observe both models in Revit and in Enscape (Figure 4). So, all changes in Revit are immediately available to evaluate in Enscape. With Enscape, the user is able to quickly explore different design options supporting a collaborative project team. Enscape will immediately show the alternative solutions or changes in the design, the designer makes in the project. With the option to create standalone files, is possible to send an Enscape file to colleagues who allow the engineer to quickly demonstrate the project and the alternative options. As the Revit allows the user can work over the model applying all the capacities of modelling, consulting the information linked to the parametric objects used in the model process, and obtain cuts over perspectives allowing the analyses of the composition of all elements of the BIM model. So the aspect of linking the consulting capacity and the VR ability of walking around is a very important improvement in the use of BIM methodology. To improve VR experience Enscape can also be used together with Oculus Rift;
- Some progress has been made on VR techniques such as registration on tracking and display hardware, but only recently the link to BIM methodology have being made. However, a VR system should be more convenient and combined with in-use applications to support multi-disciplinary users throughout construction lifecycle. Jiao *et al.* (2013) presents a pilot cloud framework regarding an environment utilizing **web3D**, **BIM** and **BSNS** (Business Social Networking Services). Technical solutions to key issues such as authoring, publishing, and composition are designed. The proposed environment is seamlessly integrated into in-use information systems and therefore enjoys greater usability.

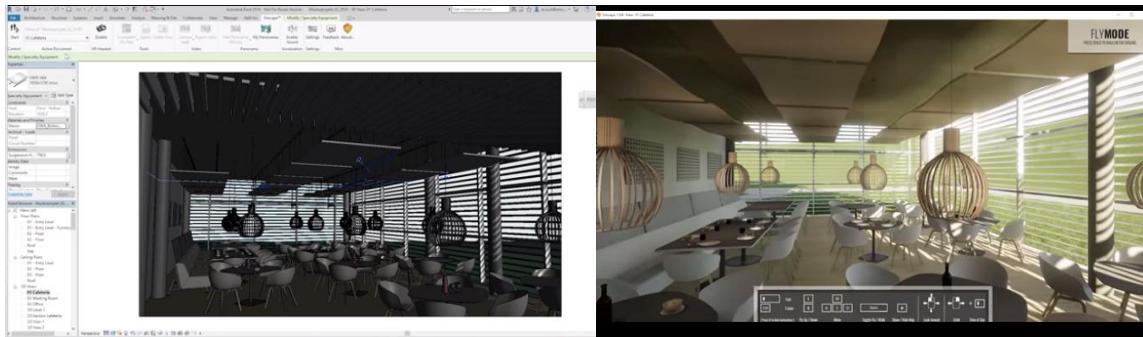


Figure 4. Enscape plugin of Revit

3. 4D/VR/BIM MODEL – CONSTRUCTION PLANNING

4D CAD models, that integrate physical 3D elements with time, have been used to visualize construction processes in several projects worldwide. 4D models have been used during the construction phase and have been shown to have benefits over processes that span the entire lifecycle of a project such as collaboration with stakeholders, making design decisions, assessing project constructability, identifying spatial conflicts in construction. Concerning the creation of 4D applications, as a support to follow construction planning, several studies can be found, linking VR technology to 3D geometric models and reports related with the generation of 4D/BIM models (Figure 5):

- Sampaio *et al.* (2012) implemented a prototype based on VR technology applied on construction planning. The geometrical AutoCAD 3D model of distinct steps of the construction activity is linked to the construction planning schedule, defining a 4D model. VR technology allows the visualization of different stages of the construction, and the interaction with the construction activity, resulting in a valuable asset in monitoring the development of the construction activity. The prototype makes use of MS Project, AutoCAD and **EON Studio** VR software). The 4D/VR application clearly shows the constructive process, avoiding inaccuracies and building errors, and so improving the communication between partners in the construction process;
- Sampaio and Mota (2016) created a 4D/BIM model using Autodesk software (Revit and Navisworks) and MS Project. The 4D model allows the addition of time associations to the 3D elements, grouped in sets related with each task, allowing the visualizations and analysis of the activity sequence for the construction. The **Navisworks** software allowed the interconnections amongst the 3D models created in the planning where this solution built the interconnection of 3D models with planning in MS Project. The ability to navigate through the model allows analyzing each corner and each location of the model. Navisworks also has the potential to provide support in the analysis and detection of conflicts amongst the specialty projects.

Recent developments in VR have encouraged the utilization of interactive architectural visualization in the design, construction and facility management of building BIM projects. Du (2016) find that the interpersonal interaction in the VR environment is more critical to the effective communication in a building project, as it creates a shared immersive experience, and developed a BIM-enabled VR environment to realize multiplayer walkthrough in virtual buildings. The multiplayer virtual walk allows real time interactions of remotely located project stakeholders in the same environment, with a shared immersive walkthrough experience.

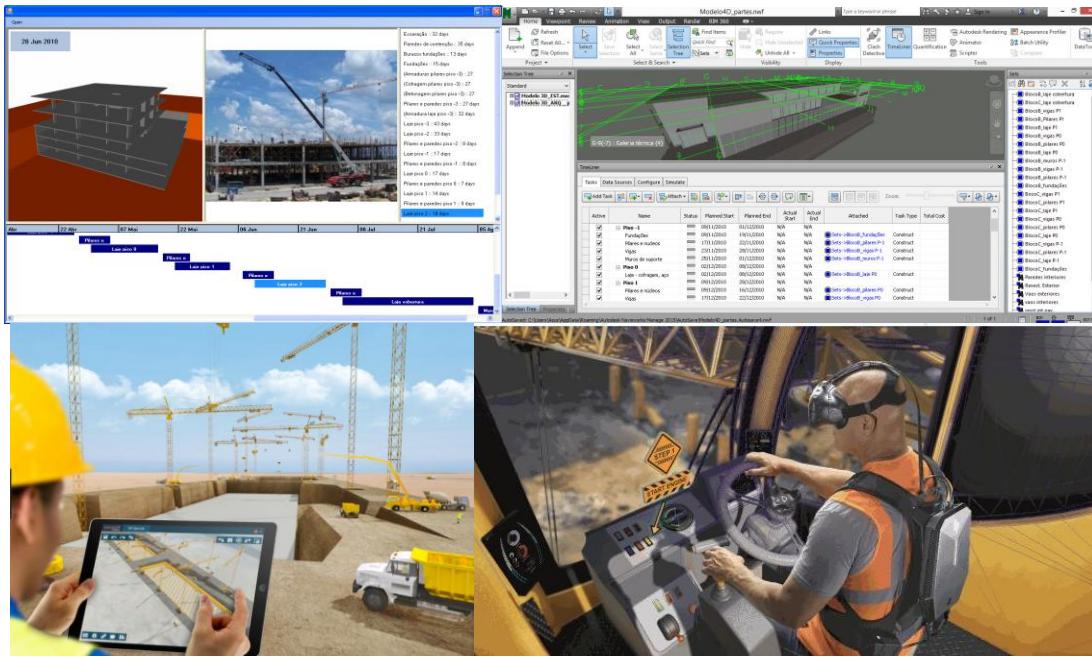


Figure 5. Screen shots of 4D/VR/BIM simulation processes

VR combines a device for interaction creating virtual environment and the current aim is study how to use device or establish links for the presentation of information contained in a BIM models. Following that perspective, commercial software houses have been developing advanced BIM/VR products. The use of BIM technology on construction projects has the potential to improve the process by allowing all team members to collaborate in an accurate and efficient way. Some of the VR software adapted to BIM and the main capacities are listed below:

- **Enscape** is Revit plugin that creates a VR walkthrough with one click, based on BIM data and can be also used in a 4D context (Enscape, 2017). There is no need to download or learn how to operate additional software. All materials, geometry and entourage come from the Revit construction project and can be changed during the VR simulation;
- Combined with the **Oculus Rift**, customers can virtually walk through the Enscape project and experience it as if it were already built. In the construction process the engineered visualise the correspondent building 4D/BIM step and consult all the information the user need to compare the progress building stage, the delay or the advance realized in the construction place. Enscape has become a standard application in projects worldwide namely as a construction planning support;
- **Augment** is software that can be used in smartphone or tablet allowing the user to view the 3D model in the real environment in real time from multiple perspectives. BIM with VR enables project stakeholders to create information-rich virtual models that help better visualize building projects. Data-rich virtual information is available on site through connected devices in real time. Site managers can overlay the BIM model on the project site. Construction issues can be addressed in real time, saving time and resources (Augment, 2017) (Figure 6);
- **Samsung Gear VR** is a virtual reality device that allows exploring virtual worlds at the construction site or during meetings. To use Gear VR during construction phases and facility management purposes two paths are needed: BIM model and construction site picture based, and users must be familiarized with Revit software and 3ds Max software for visualizing and rendering, or to navigate inside game engines like Unity3D with Android Studio are need (Rho, 2015).



Figure 6. Screen shots of Augment software

4. 7D/VR/MODEL - MANAGEMENT AND MAINTENANCE

Combining BIM with VR is expected to envision efficient collaboration, improved data integrity, intelligent documentation, distributed access and retrieval of building data and high quality project outcome through enhanced performance analysis, as well as multidisciplinary planning and coordination. In BIM, while the potential benefits of the technologies may seem evident, the industry adoption rate of BIM varies and now the importance of adding VR benefits must be a research point (Advenser, 2016). The acceptance and expansion of VR has been growing exponentially in the construction field. A VR tour can be applied, from a facility management perspective, in the maintenance control (7D/BIM model), or from a project team view, to review constructability (4D/BIM model) and to support decision making. Ever since embracing the idea of VR, the AEC sector has benefited a great deal (BIM Grid, 2017) (Figure 7):



Figure 7. Screen shots of 7D/VR/BIM applications

- A virtual tour together with **Enscape** allows the facility manager to look around the facility and check the conditions of equipment and obtain relevant information from the BIM model. Augment tool can also be used to view and consult the 3D model data, in the real environment and in real time, to facilitate installation and layout;
- The traditional EON software has been used in 4D/VR models to support the construction activity. Currently the **EON Icube** is a multi-walled system where the user is completely surrounded by images and sound. The EON Icube uses light-enhancing rigid wall material, provides users with an impressive speed of immersive content creation and deployment, and has built-in collaborative capabilities. EON maintains a virtual reality library that already has many sample AEC environments built (Eon, 2017);
- The **Autodesk LIVE** allows users to transform Revit models and step inside their own designs to share and present. AEC professionals can upload BIM models to the cloud with just one click and interactive visualizations that can be published to mobile. This gives collaborators a better sense of the design early on so changes can be made with minimal disruption or delay. It allows the consult of BIM data and so it supports management and maintenance within a collaborative team.

5. CONCLUSIONS

BIM+VR enable architects and designers to improve the level of communication in a collaborative design. The advantage in using VR is in the communication of ideas, concepts and the vision for the building in an interactive environment.

When everyone shares a common understanding of the design, the project can be executed supported in an adequate and clear platform, improving efficiency when compared to a VR absence project work. This current BIM with VR topic require dissemination; application in real cases and pointed out, in reports, achievements and limitation; following the technologic advances that supports the BIM use and the visualization of data, in real time while the interacting with the model made possible by VR technology. BIM+VR provide an opportunity to analyse and explore BIM models within virtual environments.

Introducing VR interactive capabilities into 3D/BIM models in the construction process is a main way to test virtually and correct a construction project before the realization, as the walkthrough is available as well the visualization of data linked to each parametric object, improving the necessity collaboration within the design team. So, BIM/VR applications can contribute to reduce costs due to the construction of a real mock-up, and to avoid mistakes on site that generates material wastes. Software houses have been demanding the integration of BIM with VR plugins in order to support the development of nD/BIM tasks. The fundamental base of BIM+VR concerns the collaboration as a first step, but the possibility of consult data while walking through the building, improves significantly the use of BIM in design, construction, maintenance and management.

REFERENCES

- Advenser, 2016. Integrating BIM and Virtual Reality – The Next Big Thing?, <https://www.advenser.com/2016/11/23/integrating-bim-and-virtual-reality-the-next-big-thing/>
- Augment, 2017, <http://www.augment.com/blog/enhancing-bim-with-augmented-reality/>
- Azhar S. et al, 2008. Building information modeling (BIM): benefits, risks and challenges, <http://ascpro.ascweb.org/chair/paper/CPGT182002008.pdf> 2008
- BIM Grid, 2017. Collaboration in VR, The Future of VR BIM. <https://www.linkedin.com/pulse/collaboration-vr-future-bim-bim-grid>
- Chen L.J. and Luo H., 2014. A BIM-based construction quality management model and its applications, *Automation in Construction* 46 (2014) 64–73 <http://dx.doi.org/10.1016/j.autcon.2014.05.009>
- Eadie R. et al, 2013. BIM implementation throughout the UK construction project lifecycle: An analysis, *Automation in Construction* 36 (2013) 145–151. <http://dx.doi.org/10.1016/j.autcon.2013.09.001>
- Enscape 2017, <https://enscape3d.com/>
- EON icube 2017 <https://www.eonreality.com/eon-icube/>
- Jiao Y. et al., 2013. Towards cloud Augmented Reality for construction application by BIM and SNS integration, *Automation in Construction* 33 (2013) 37–47. <http://dx.doi.org/10.1016/j.autcon.2012.09.018>
- Rho J., 2015. CS10427: Virtual-Augmented Reality for Construction, Utilizing Gear VR, Autodesk University <http://au.autodesk.com/au-online/classes-on-demand/class-catalog/2015/revit-for-architects/cs10427#chapter=0>
- Rift, 2017 Oculus Rift, <https://www.oculus.com/experiences/rift/>
- Sampaio A.Z. and Mota C., 2016. BIM model of structures used in construction planning, *Proceedings of the National meeting of Structural Concrete*, Coimbra, Portugal, 2 - 4 November 2016. ID: 4, 10 pgs, <http://be2016.dec.uc.pt/>
- Sampaio A.Z. and Simões D., 2017. Maintenance of buildings supported on BIM methodology, *Proceedings of XIV DBMC - Durability of Building Materials and Components*, Ghent, Belgian, May 2017, e-ISBN: 978-2-35158-159-9
- Sampaio A.Z. et al, 2012. Construction and maintenance planning supported on virtual environments, Book: *Virtual Reality - Human Computer Interaction*, Ed. Xin-Xing Tang, Publisher: InTech, ISBN 978-953-51-0721-7, DOI: 10.5772/46409, ch. 07, 125 – 152, <http://www.intechopen.com/articles/show/title/construction-and-maintenance-planning-supported-on-virtual-environments>
- Smart reality 2017, The ultimate BIM interface, <http://smartreality.co/>
- Stingray, 2017. <https://knowledge.autodesk.com/support/stingray/getting-started?sort=score>
- Viatechnik, 2017, 50 Virtual Reality Technologies in Architecture and Engineering, <https://www.viatechnik.com/resources/50-virtual-reality-technologies-in-architecture-engineering-and-construction/>

SCALING E-HEALTH: AN ANALYSIS OF THE FUNCTIONS OF THE TECHNOLOGICAL INNOVATION SYSTEM AND THE ROLE OF GEOGRAPHIC PROXIMITY

Annemijn van Gorp and Bert Mulder
eSociety Institute, The Hague University of Applied Sciences, The Netherlands

ABSTRACT

Even though a wide variety of e-health technologies is available, uptake is only slowly taking place. This study analyze factors that influence the uptake and scaling of e-health initiatives in the Netherlands. Taking a Technological Innovation Systems (TIS) approach, through interviews, observation and document analysis the seven functions of the innovation system are analyzed and barriers to system development are identified. The study finds that financial resource mobilization, resistance to change, and to some extent ‘guidance of the search’ are hampering the scaling of e-health, even though the latter is increasingly taking shape through formation of active networks of health professionals and managers, and ICT professionals. In addition, it was found that both organizational and geographical proximity matter in the collaborative development of e-health.

KEYWORDS

E-health, scaling, technological innovation system, geographical proximity

1. INTRODUCTION

In the light of rising health care costs, e-health is seen as a driver for enabling more cost-effective health (e.g. Home et al., 2015; Ross et al, 2017). In addition, positive aspects like patient-empowerment are attributed to e-health.

Besides large global players like Google, Apple, and Amazon entering the field of e-health – through services ranging from the offerings of online communities, software developer kits, AI applications and cloud storage - at the national level e-health is gaining commercial interest as well: telecom operators, large ICT vendors and small software development companies alike have ventured into the e-health business.

Nevertheless, in the Netherlands the percentage of patients using e-health services remains small. For example, in 2015, online access to medical information was still less than 1% (Nictiz, 2015); and only 15% of people was able to ask online for repeat prescriptions (Nictiz, 2015). While already 19% of people tracks certain physical activity, only 7% has digital access to information about doctor’s visits or treatment (Nictiz, 2015).

Given the interest in e-health by ICT players, and acknowledging the potential e-health has, while at the same time observing limited use in the formal health domain so far, this study started with the question: how can we further stimulate the scaling of e-health in the Netherlands? To this end, in this paper we aim to identify factors that influence the uptake and scaling of e-health initiatives in the Netherlands.

For purposes of this study, e-health may concern various types of applications where digital and internet technologies are being used to provide citizens information and services related to healthcare, cure, and prevention, and includes applications ranging from ‘basic’ apps that provide information about certain types of diseases and their treatment, to apps that enable caregivers to communicate patient needs with each other, to digital watches that provide medication notifications or wearables that track physical condition, to name just a few.

2. EHEALTH ADOPTION AND IMPLEMENTATION

Research in the field of e-health is fairly wide ranging, spanning from for example studies concerning the potential of emerging technologies such as the rise of wearables (e.g. Piwek et al., 2016) or big data (e.g. Kuo et al., 2014; Liu & Park, 2014), to the evaluation of factors that influence the use of (new) applications (e.g. Garcia-Gomez, 2013, Redfem et al., 2014, Zhang et al., 2015), or adoption by certain target groups such as the elderly (Bujnowska-Fedak & Pirogowicz, 2014).

While most often application-specific factors that stimulate or impede adoption are identified, some more general barriers to adoption at the individual level have been found as well. These include e.g. insufficient communication from care provider to patient about the service, lack of perceived value of the service to the patient, patients' preference for oral rather than digital communication with care provider, low levels of internet literacy by the patient, lack of availability of devices, etc. (Zhang et al., 2015).

Although individual level adoption studies are valuable, sustainable health innovation asks for a system that enhances value to all stakeholders. Users, health professionals, service providers, and innovators need to work together in the same direction. It asks for understanding how value is generated at different levels: the individual, organizational, and societal, while taking into account the regulatory systems (Home et al., 2015). Studies at a more aggregated level of analysis into adoption and market growth remain scarce.

Nevertheless, a number of studies on implementation and scalability of e-health have been conducted as well (e.g. Dehzad et al., 2014; Devlin et al., 2015; Home et al., 2015). A variety of challenges to successful implementation have been identified, including issues of practical nature such as interoperability and information governance (Dehzad et al., 2014; Devlin et al., 2015) as well as privacy and security problems (Dehzad et al., 2014). Branding and marketing have been found challenging as well (Devlin et al., 2015). Further, challenges of economic nature have been found in lack of opportunity for finding a suitable business case (Dehzad et al., 2014; Van Gorp, 2016) and challenges with payment systems due to absence of funding from insurance companies (Dehzad et al. (2014)).

E-health development is by nature a networked process, as various challenges related to the network have been found as well. Development tends to happen through co-creation, often with the patient at the center (Uruena et al., 2016). But often also various business partners play a role (Uruena et al., 2016). Establishing and maintaining these relationships can be challenging (Devlin et al., 2015), but also due to the networked nature, at times a tension can be observed between embracing innovative co-design and achieving delivery at pace and scale (Devlin et al., 2015), as various parties taking decisions can be time consuming. Also, stakeholder-networking capability is used for attracting new knowledge. Hence, it is considered important for the success and sustainability of projects (Uruena et al., 2016).

Finally, challenges of organizational and management nature have been found to influence e-health development. Organizational capabilities contribute in transitioning e-health innovation projects from pilots to real implementations (Uruena et al., 2016). Culture in organizations may impede development; conservative culture in the field of health has been found to play a role (Dehzad et al., 2014), but also leadership is important -- visionless development has been found to contribute to lack of development (Dehzad et al., 2014).

3. EHEALTH AS AN INNOVATION SYSTEM

The factors at the organization and societal level influencing e-health implementation and scaling remain rather scattered: few studies have taken an integrated look at e-health development from the perspective of an innovation system.

The role of innovation systems has long been focus of research in various sectors of the economy. It builds on the idea that technological change comes with changes in the social dimension: changes in the behavior of the user, in regulation, etc.. Such changes can also be viewed as a 'system-level' change. Hence, an innovation system consists of all societal subsystems, actors, and institutions, related to innovation. Innovation systems co-evolve along with technology (Hekkert et al., 2007).

TIS highlights how platforms and technologies emerge and the role legitimization processes and the collective dimension play when taking an innovation to the growth phase (Hallingby, 2016). The emphasis on emergence fits the fairly new field of e-health well. The Innovation Systems approach stems from a

combination of institutional theories and evolutionary theories. Its main idea is that innovation and diffusion of technology is both an individual and collective act. The network of institutions and organizations, through their activities, develop and diffuse technologies. The dynamics is thus the key, not just structure.

The basic building blocks of (all) innovation systems are actors, institutions (e.g. legislation, technology standards), networks (linkages between organizations) and technology. Technology enables and constrains activities of actors in the innovation system (Hekkert et al., 2011).

TIS focuses on a number of key processes that are important for the performance of the innovation system (the so-called “functions of innovation systems”). While all innovation systems consist of similar components, they function differently. The functioning of the innovation system determines its performance. The following seven functions of innovation systems have been identified in research: 1) entrepreneurial activities; 2) knowledge development; 3) knowledge exchange; 4) guidance of the search; 5) formation of markets; 6) mobilization of resources; 7) counteracting resistance to change.

Although traditional innovation systems research tends to focus on national innovation systems, the role of local, regional and global systems are also known to have significant impact on innovation and market development. For example, it has been found that external contacts outside of local industry can be crucial in innovation processes. Further, innovation processes in certain instances have been found to be regional phenomena, as regional resources and collaborative networks often have significant influence on individual organizations' activity. Regions have been found to bring place-specific, contextual knowledge (Asheim & Isaksen, 2002).

Indeed, regional innovation systems literature and theory stress the role of learning and knowledge exchange, for which geographic proximity is important for innovation to take place. In that realm, geographical proximity has been viewed as a key driver for regional innovation due to the possibilities it has for tacit knowledge transfer, which is more difficult at a global level. However, advances in ICT have made global communication and knowledge transfer much easier. Therefore, Rallet and Torre (1999) pose that organization proximity might be more influential than geographical proximity when it comes to technology transfer.

Finally, global knowledge flows also play an important role in innovation, through channels such as international R&D collaboration, FDI, virtual communities, conferences, etc. Organizations use both the global and more local or regional communities; both multinational corporations and SMEs (Martin et al., 2017).

4. RESEARCH APPROACH

This study will use the TIS approach to analyze performance of the e-health innovation system in the Netherlands, and in particular the influence of geographical proximity on the functioning of the system.

The TIS approach prescribes five steps to be followed to monitor the development of the emergence of technologies. Step 1 describes the mapping of structure. The next step describes the phase of development in relation to the state of diffusion: pre-development, development, take-off, acceleration or stabilization phase. In the third step the functions of the system and their performance levels are described. As many structures are often not yet in place for emerging technologies, in this step, processes that contribute to the development, diffusion, and use of innovations are described. In the fourth and fifth step key failures/barriers and handholds for improvements are described.

Given the focus on the role of geographical proximity, the study has focused on one particular region: the Delft region in the Netherlands. A variety of data collection methods have been used to gain insight into e-health developments: interviews, observation, and document analysis. Approximately twenty stakeholders have been interviewed and/or observed in meetings and workshops, namely board and line managers, health professionals, and ICT professionals (mainly app developers in the e-health domain). These participants were also selected to represent various health sectors, including primary and secondary care, long term care, and both small and large health facilities, in order to provide for a wide variety of perspectives. Interviews were recorded and transcribed. Next, through open coding they were analyzed.

In line with the TIS approach, this study will first identify the structural components of the Dutch e-health system. Next, the phase of development of e-health in the Netherlands will be determined. Then the system functions will be analyzed, including the structural causes for functional barriers (hampering).

5. RESULTS

5.1 Structural Components of the eHealth Innovation System

The structure of the innovation system consists of actors, institutions, networks, and technological factors.

The e-health system in the Netherlands consists of a large interconnected network of organizations and institutions. Besides well-known organizations like hospitals, clinics, physicians, mental health clinics, etc., a range of other actors is involved in the field. For example knowledge institutes include universities and other higher education institutes that educate students in health and/or ICT, as well as research subjects related to (e-) health. Market actors include new players in the field like ICT companies and consultants that develop health services, as well as traditional players like pharmacies and retailers of medical instruments.

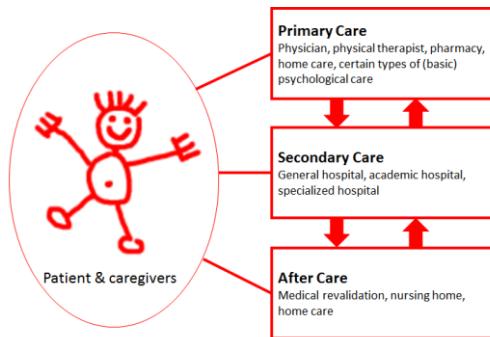


Figure 1. Traditional patient journey through the healthcare system

The patient is at the center of the system, and a customer journey through the health care system may go from primary, to secondary, to after care (see Figure 1). Primary care refers to care that anybody can use without needing formal referral, in which the physical takes a central position. Secondary care refers to those types of health care services that require a referral, such as seeing specialists in the hospital or medical rehabilitation, and certain psychological help. Third line care refers to situations in which the patient needs in-house care – e.g. hospitalization or nursing home.

Given the relation between primary and secondary care, networks are often created at the regional level: physicians from certain areas will direct their patients if needed to specialists (secondary care) in their regional hospital. There is no strict demarcation of regions however, e.g. because sometimes if particularly specialized experts (surgeons, specialists) are needed, patients are sent to a hospital or other health facilities further away.

In the Delft region for example, the Delft E-Health Academy is active. It is a network where stakeholders commit themselves to collaborate in a e-health learning network. It is comprised of patient/citizens, health care providers, municipality, developers of e-health, etc. Its goals are awareness creation, motivation, knowledge exchange, development of skills, and networking/meeting. Activities are centered around needs of stakeholders.

Government bodies also play an important role in the e-health landscape. At the national level the Ministry of Health, Wellness and Sports plays a role in promoting the scaling of e-health. At the municipal level government is involved in health as well, in particular for youth and elderly care. In 2014 the Minister of Health in the Netherlands has announced a number objectives related to e-health in the areas of 1) digital access to medical information; 2) self-monitoring and measurement; and 3) 24/7 access to a doctor through teleconsulting. By 2019 80% of chronically ill people should have direct access to medical information, including information about medication, vital functions and test results, possibly for use in mobile apps or web-based applications (Schippers, 2014). Further, early February 2017 MKB Nederland (small and medium sized companies in the Netherlands) together with VNO-NCW (employers' organization) and a number of other organizations including health insurance organizations, GGZ (branch organization mental healthcare providers) and the Dutch Association of Hospitals advised the Dutch Minister to stimulate ehealth; VNO-NCW and MKB-Nederland indicating that they believe that 'smart health' can lead to significant cost savings (VNO-NCW & MKB-Nederland, 2017). The Council for Health and Society also advises the Minister to

further stimulate e-health adoption together with health provider and health insurance organizations. In particular, they advocate for the government to develop a so-called “e-health highway” to enable easier and cheaper data exchange, as currently data exchange between different information systems is difficult, and patients do not have access to their own files. Further, new apps are stand alone applications that cannot communicate with existing health information systems.

Other important actors include patients and patient federations.

5.2 Phase of Development

The phase of development of the TIS determines to a large extent how the structure and functioning of an innovation system should be built up: when technology is in early phase of development the innovation system has a different structure and certain functions of the systems are then more relevant than others than for a more mature technology (Hekkert et al., 2011). The e-health market, because of its size, with many different (applications of) technologies making up the market, could be argued to be in the development phase, although for different parts different phases could be discerned.

On an aggregate level, the pre-development phase would be too basic – since availability and use of e-health is available beyond just a number of pilots. The development phase is characterized by the availability of commercial applications, which is indeed the case for the Netherlands. However, market growth is still limited, and market saturation is still far away. For example, during interviews physicians express to barely receive requests from patients about e-health use; and while hospitals are starting to develop their own apps, their application areas are highly scattered, ranging from generic informative ‘how to prepare for surgery’ apps to dermatology apps that are very specific in use and help with diagnosis.

5.3 System Functions of the e-health Innovation System

5.3.1 Entrepreneurial Experimentation and Production

Seemingly quite some entrepreneurial experimentation and production takes place in the Dutch e-health market. Many large and small app developers exist. Together they have developed hundreds of e-health apps. Large scale uptake of e-health has taken place in only few large health organizations. Much e-health development happens through small development programs, where independent developers make an app, and try to reach consumers through Apple and Google app stores. During interviews many of them indicate that they experience that through selling in these stores they do not easily gain extra customer base; therefore as a next step they try to sell their services to (large) health organizations. During this step many developers experience a lengthy decision process and overall not much interest yet in e-health uptake.

Limited activity by the health organizations, rather than entrepreneurial activity itself, therefore seems a barrier to entrepreneurial activity.

5.3.2 Knowledge Development

E-health research gains attention in higher education institutes (both ‘regular’ Dutch universities as well as a number of universities of applied sciences) but also at think tanks and other (commercial) research institutes (e.g. Nictiz). The Medical Delta, a network of various stakeholders which stimulates cross boundary research, serves as an example of the perceived importance of technological innovation in health in the Netherlands. An increasing number of publications on e-health is available.

Research studies often times evaluate certain applications or information systems. Such evaluations often concern pilots or fairly new services that have not been ‘tried and proven’.

There are not so many explicit knowledge gaps that hamper the development of e-health.

Nevertheless, even though there is interest in the academic and ICT community in the topic, and an increasing knowledge base about the added value of e-health is available, it all concerns very new activity while long term effects of e-health cannot (yet) be tested because it is a young and emerging field. Moreover, given the quick developments in the field, by the time extensive testing has been done, chances are high a new innovation is already outdated – an often heard critique from practitioners when it comes to e-health research in academia. This is problematic to the field of medicine which is largely built on “evidence based”

practice. Indeed, in the medical profession the need to test to ensure informed decisions before investing, has often been voiced (see also e.g. Prof. Chavannes, from Leiden UMC, in a recent interview¹).

Thus, generally speaking there is no lack of knowledge development. However, possibly we could say that the practice of evidence based medicine at times is incompatible with the often quick to market push of new apps by ICT developers.

5.3.3 Knowledge Exchange

Generally, for the development of IS it is important that multiple, and various, types of networks exist: to enable knowledge exchange between science and industry, between users and industry, across geographical borders. Of course, there are national conferences, for both professionals and practitioners. Universities and research institutes organize public lectures etc. An annual national e-health week is organized in January. In addition, regular professional publications like “Healthcare & ICT” (“Zorg & ICT”) magazines are published.

Knowledge exchange is also taking place through the DEHA. It is unique in that it brings together users and health professionals from a particular (Delft) region, as well as to some extent the academic/science community (learning from others). Sessions with end-users (citizens) are held to assess needs and wants; sessions with health providers are held to learn about their problems, experiences, and needs.

Interviewees indicate that generally they find that there are many opportunities for some knowledge exchange (small conferences or networking events). DEHA participants indicate that they generally enjoy being able to hear from experiences of others, and being able to learn from that, or for example to be able to exchange policy documents. The specific regional advantage that is mentioned ranges from: “What works in one place, doesn’t work in another”. “we have to take into account what people in the region find important”, “It shouldn’t be only about sharing ideas and knowledge. You have to be able to do something with it”. “I think that at the end we need to go towards more collaboration. The neighborhood approach, personal empowerment. That’s where you have to collaborate”. However, this collaboration is not only seen as a positive thing: “you need to watch out with sharing, in order to remain a unique player”. Nevertheless, not everybody thinks the local/regional approach is important: “I go to meetings because of the subjects of discussion. Not so much for the network”.

Related to this, a number of professionals from health providers with multiple locations throughout the country indicate that internal collaboration with colleagues from other locations but also sometimes collaboration with other organizations in the same ‘sub-field’ of health care (e.g. mental health) is particularly helpful; more so than regional collaboration with different types of health providers.

Thus, knowledge exchange increasingly takes place. Depending on context and need of the care professional, regional or national level collaboration is preferred.

5.3.4 Guidance of the Search

Guidance of the search relates to the existence of a clear vision on how the market should develop (Hekkert et al., 2011). A number of interviewees indicate that they would like to see more formal guidance and direction. They are waiting for formal policy and targets by the government. At the same time, a networking organization like DEHA provides for the exchange of ideas and thereby the development of this vision. Interviewees indicate that this gives them guidance as to where e-health is going and how it will affect their practice.

5.3.5 Market Formation

Market formation concerns primarily the market size (Hekkert et al., 2011). Question is whether market size forms a barrier for further development of the TIS, and whether current or expected future size are sufficient.

In the case of e-health in the Netherlands, this is difficult to say. There is no clear ‘e-health market’ yet. E-health is in its initial stages of development. Given that e-health entails such a broad array of services, there is potential for a large market to develop. Market development will depend on viable business cases for the various services. This turns out to be difficult, also due to very limited support by insurance companies (see also next section).

¹ See <https://www.hcc.nl/kennis/1221-weinig-onderzoek-naar-resultaten-e-health>

5.3.6 Resource Mobilization

Resource mobilization refers to the availability of physical, human, and financial resources (Hekkert et al., 2011).

Physical resources: physical resources do not seem a barrier – e.g. citizens tend to have high quality internet access. Standardization or compatibility of systems is an issue however.

Human resources: education and training programs on the use of e-health in the health care sector -by and large- still need to be developed. Nurses, doctors, etc., do not yet receive this as part of their formal education. Training thus is necessary. Organizations that have started to implement e-health have experienced the need for training at multiple levels: how to use certain software, how to communicate with clients through text rather than F2F, etc. Organizations are figuring out these issues through trial-and-error. Nevertheless, often with limited amount of training employees can start to use e-health. From interviews it was inferred that the culture shift of being willing to work with e-health is the more important hurdle. Incorporation of e-health in formal education programs will of course stimulate a culture of e-health acceptance.

Financial resources: although generally increasing start-up capital is available, a key problem is the system through which health care is payed. Health care organizations are paid by insurance companies on the basis of treatment carried out. All activities that patients undergo are according to pre-specified treatments. E-health is barely included in the insurance system (to date this is mostly restricted to blended combination therapy in mental (e-)health). Therefore, doctors do not necessarily have an incentive to start using e-health (except when they can be directed by the hospital because of other certain benefits such as more efficient use of operating rooms).

5.3.7 Counteract Resistance to Change/Legitimacy Creation

Usually a number of employees have an interest in the new technology, and are among the first to start experimenting, or doing a pilot with the technology. A number of organizations indicated that these employees are key in spreading awareness and enthusing colleagues to start using the technology as well. The extent of resistance on the other hand seems to depend on how much impact the new technology has on work processes. For example, when organizations start using online video calling instead of paying house visits to clients, significant resistance is experienced. But as an interviewee indicated, for blended combination therapy in mental healthcare this is much less the case because the face to face contact remains and digital communication is actually perceived as an added value. Nevertheless, overall the level of resistance is a personal matter, and tends to differ across people in an organization.

5.4 Barriers to Optimal Performance

A number of barriers have been identified. Financial resources and a number of social/cultural aspects play a role. The social and cultural aspects relate to the culture of evidence based practice that seems incompatible with quick developments in ICT. In addition, resistance due to unknown or new work processes (legitimacy creation) play a role. Limited guidance of the search (vision) plays a role, however increased knowledge exchange in networks may solve this problem.

At the same time, regional and organizational proximity were found to play a positive role in taking away barriers: they support knowledge exchange.

6. CONCLUSIONS

This paper investigated what factors play a role in the uptake and scaling of e-health in the Netherlands by analyzing the seven system performance functions developed by Hekkert et al. (2007, 2011) and the role of geographical proximity. The functions lagging behind in more or lesser extent are: financial resource mobilization, resistance to change, and to some extent guidance of the search, even though the latter is increasingly taking shape through formation of active networks of health professionals and managers, and ICT professionals. With regard to the latter, it was found that both organizational and geographical proximity matter. Organizational proximity matters: organizations with strong linkages to similar organizations, even when not geographically closely located, were found helpful collaborators in e-health development.

Nevertheless, regional proximity plays an important role as well, as e-health asks for collaboration between various organizations in primary, secondary and home care.

Recommendations that follow from these findings include: 1) establishment of stronger regional networks may help spur vision creation and knowledge exchange; and 2) Stronger networks in general; national or regional - may also help counter resistance to change due to increased awareness of what e-health can do.

REFERENCES

- Asheim, B. and Isaksen, A. (2002). Regional Innovation Systems: The Integration of Local ‘Sticky’ and Global ‘Uniquitous’ knowledge. *In the Journal of Technology Transfer*.
- Autio, E. and Thomas, L.D.W. (2014). Innovation ecosystems: Implications for innovation management. In Dodgson, M., Philips, N. and Gann, D.M. (eds.). *In The Oxford Handbook of Innovation Management*.
- Bujnowska-Fedak, M.M. and Pirogowicz, I. (2014). Support for e-Health Services Among Elderly Primary Care Patients. *In Telemedicine and e-Health*, pp 696-704.
- Dehzad, F., Hilhorst, C., de Bie, C. and Claassen, E. (2014). Adopting Health Apps, What’s Hindering Doctors and Patients? *In Health*, Vol. 6, pp 2204-2217.
- Devlin, A.M., McGee-Lennon, M., O’Donnell, C.A., Bouamrane, M.-M., Agbakoba, R., O’Connor, S., Grieve, E., Finch, T., Wyke, S., Watson, N., Browne, S., and Mair, F.S. (2015). Delivering Digital Health and Well Being at Scale: Lessons Learned during the implementation of the Dallas Program in the United Kingdom. *In Journal of the American Medical Informatics Association*, Vol. 23, pp 48-59.
- Garcia-Gomez, J.M., de la Torre-Diez, I., Vicente, J., Robles, M., Lopez-Coronado, M. and Rodrigues, J.J. (2013). Analysis of Mobile Health Applications for a Broad Spectrum of Consumers: A User Experience Approach. *In Health Informatics Journal*, Vol. 20, No. 1, pp 74-84.
- Hekkert, M., Negro, S., Heimeriks, G., and Harmsen, R. (2011). *Technological Innovation System Analysis: A Manual for Analysts*. Universiteit Utrecht, Faculty of Geosciences, Copernicus Institute for Sustainable Development and Innovation.
- Hekkert, M.P., Suurs, R.A.A., Negro, S.O., Kuhlmann, S., Smits, R.E.H.M. (2007). Functions of Innovation Systems: A New Approach for Analysing Technological Change. *In Technological Forecasting and Social Change*, Vol. 74, pp 413-432.
- Home, R., Bell, J.I., Montgomery, J.R., Ravn, M.O. and Tooke, J.E. (2015). A New Social Contract for Medical Innovation. *Comment in The Lancet*, Vol. 385, March 28.
- Kuo, M-H., Sahama, T., Kushniruk, A.W., Borycki, E.M., Grunwell, D.K. (2014). Health Big Data Analytics: Current Perspectives, Challenges and Potential Solutions. *In International Journal of Big Data Intelligence*, Vol. 1, No. 1/2, pp 114-126.
- Lio, W., and Park, E.K. (2014). Big Data as an e-Health Service. *In Computing, Networking and Communications*, pp. 982-988.
- Martin, R., Aslesen, H.W., Grillitsch, M., and Herstade, S.J. (2017). Regional Innovation Systems and Global Knowledge Flows. *Papers in Innovation Studies: Working Paper*. Paper no. 2017/07.
- Piwak, L., Ellis, D.A., Joinson, A. (2016). The Rise of Consumer Health Wearables: Promises and Barriers. *In PLOS Medicine*, pp 1-9.
- Rallet, A. and Torre, A. (1999). Which Need for Geographical Proximity in Innovation Networks at the Era of Global Economy? *Presented at the 39th Congress of the European Regional Science Association*, August 23-27, Dublin, Ireland.
- Redfem, J., Usherwood, T., Harris, M.F., Rodgers, A., Hayman, N., Panaretto, K., Chow, C., Lau, A.Y., Neubeck, L., Coorey, G., Hersch, F., Heeley, E., Patel, A., Jan, S., Zwar, N., and Peiris, D. (2014). A Randomised Controlled Trial of a Consumer-Focused e-Health Strategy for Cardiovascular Risk Management in Primary Care: the Consumer Navigation of Electronic Cardiovascular Tools Study Protocol. *In BMJ Open*, pp 1-8.
- Ross, J., Stevenson, F., Lau, R. and Murray, E. (2017). Exploring the Challenges of Implementing e-Health: a Protocol for an Update of a Systematic Review of Reviews. *In BMJ Open*, pp 1-5.
- Thomas, L.D.W. and Autio, E. (2012). Modeling the ecosystem: a meta-synthesis of ecosystem and related literatures. *Paper presented at the DRUID 2012*, CBS, Copenhagen, Denmark.
- Uruena, A., Hidalgo, A., Arenas, A.E. (2016). Identifying capabilities in innovation projects: Evidences from eHealth. *In Journal of Business Research*.
- VNO-NCW & MKB-Nederland (2017). *Vooruit met de Zorg: Beter, Slimmer, Menselijker*. Available at www.nl-nextlevel.nl
- Zhang, X., Yu, P., Yan, J. and Spil, T.A.M. (2015). *In BMC Health Services Research*, Vol. 15, No. 71, pp 1-15.

AN IMPACT ASSESSMENT FRAMEWORK FOR ELECTRONIC PATIENT RECORDS MANAGEMENT: A CASE OF NAMIBIA

Ronald Karon and Nomusa Dlodlo
Namibia University of Science and Technology, Namibia

ABSTRACT

The government of Namibia has invested significantly in Health Information Systems for purposes of quality healthcare. The investments can however not be substantiated, as it is not clear whether Information and Communication Technologies (ICT) have an impact in the healthcare. Therefore a development of an Electronic Patient Records Management (E-PRM) framework to assess the impact was the aim of this research paper. The title of the developed framework was Ronald's Patient Records Impact Assessment (R-PRIA) Framework. This framework can be used as a tool that would aid in measuring the change brought about by ICTs on Patient Records Management. The application of the framework may yield results that could aid Namibian healthcare decision makers on whether or not further investments on ICTs, specific to E-PRM, would be necessary. The drive of the Ministry of Health and Social Services is on ensuring that public healthcare, which was designed for those who cannot afford private healthcare services, is of the right quality. E-PRM automates the process of admission, discharges and billing to mention a few. A number of reviewed impact assessment frameworks ranging from Sustainable livelihoods approach, Sen's capability framework and Rural ICT comprehensive evaluation framework feed into the development of the E-PRM impact assessment framework, which was the result of this research paper. The case study was a public hospital and data collection methods consisted of semi-structured interviews and questionnaires.

KEYWORDS

Health Information Systems, Impact Assessment, Impact Assessment Framework, Electronic Health Record

1. INTRODUCTION

The presence of Health Information Systems (HIS) increasingly became visible in developing countries, over a period of time. Its use is supported by governments and non-governmental organisations with the objective of improving healthcare services (Nyella, 2009). The combination of technology users, technology and processes enhancing healthcare work related tasks defines HIS (Tossy, 2014). Additionally, HIS aids in the management of health related events, such as patient records (Fabbri et al., 2011; Lee, 2013). The management of patient records is often achieved by the use of Electronic Health Records (EHR) (Fabbri et al., 2011). Accessing and storing patient records electronically is referred to EHR, which in this context can be also called Electronic Patient Records Management (E-PRM) (Jafari et al., 2011).

The Ministry of Health and Social Services (MOHSS) is the main body or custodian of health services in Namibia. Health services consist of public and private healthcare (Khan and Edwards, 2012). The public healthcare is of interest to the general population for reasons such as affordability.

Much investment in Information and Communications Technologies (ICT), including E-PRM systems, for healthcare purposes had been made by the Namibian government (GRN) (Tossy, 2014). However the impact that these technologies, such as E-PRM systems, have is lacking to support investment efforts by the MOHSS (Economist, 2016). Additionally an impact assessment framework for this purpose is non-existent (Economist, 2016). Patient information and the management thereof is a crucial aspect of healthcare. Health personnel including health practitioners are reliant on well managed and current patient information, in other words effective Patient Records Management (PRM), at times for critical decision making regarding a patient's health. The effective management of patient information is additionally essential for executives in running the healthcare as it relates to decision making.

The aim of this research paper was to develop an impact assessment framework specific to E-PRM, in the Namibian public healthcare.

2. OBJECTIVES

The objectives outlined below were defined to aid in achieving the aim of the research paper:

1. Examine how ICTs are utilised in the Namibian healthcare facilities for patient records management.
2. Determine the criteria for an impact assessment framework to assess the effectiveness of E-PRM systems through examining existing impact assessment frameworks.
3. Assess the effectiveness of E-PRM.
4. Develop a framework that assesses the impact of E-PRM in Namibia.

3. LITERATURE REVIEW

3.1 Impact Assessment Definitions

Impact assessment is known to focus on “identifying and evaluating change” (Streatfield and Markless, 2009, p. 134). The International Federation of the Red Cross and Red Crescent Societies (IFRC) explains that impact assessment refers to a “means of measuring the effectiveness of organizational activities and judging the significance of changes brought about by those activities” (IFRC, 2005, p. 1). Key factors such as measuring activities and change are therefore fundamental concepts in understanding impact assessment. For the context of this paper, the underlying definition would be health events by health personnel causing a change when using the E-PRM system.

3.2 Impact Assessment Frameworks

Existing impact assessment frameworks were identified and briefly explained in this section. The frameworks were reviewed in achieving objective two of this paper. Criteria from these frameworks were identified which fed into the questionnaires that would be used in the development of an E-PRM impact assessment framework, aligned to the Namibian healthcare setting. The selected frameworks were constrained by the scope of the paper. The first framework is the Sustainable livelihoods (SL) approach. This framework had been linked to issues such as rural development and poverty reduction (Wilson, 2009). SL approach is described to be as “people's capacity to access options and resources and use them to make a living in such a way as not to foreclose options for others to make a living, either now or in the future” (Wilson, 2009, p. 6). Secondly, Sen's capability framework. This framework has aspects of functionings and capabilities related to livelihood in a community (Walker, 2015). An individual possesses the freedom to choose the life he/she wants or would want for him/herself, which can refer to an individual's well-being. This freedom in Sen's term links to capabilities of individuals to attain certain functioning. The third and final framework is the Rural ICT comprehensive evaluation framework (RICT-CEF), which focuses on the use of ICTs as a tool to arguably improve communities. The use of ICTs in community projects is known to have its successes and failures (Mthoko and Khene, 2015). This lead to the development of the RICT-CEF which is a comprehensive framework that spans throughout the life cycle of a project.

4. RESEARCH METHODOLOGY

4.1 Research Design

For this research study a case study approach was employed. A public hospital called Brown (pseudonym for anonymity), situated in Windhoek, was the case study for this paper. The hospital is known to be a national

referral treating patients from across the country. As a research paradigm an interpretivist method was adopted. The method enabled for the study of things as is, for sense making of phenomena linked to meaning by or from people. Additionally qualitative data gathering processes were supported with this method.

4.2 Participants' Sample

A random sampling of hospital personnel was done. The participants consisted of nurses and secretaries as the primary users of the PRM system and the support staff to the users of the PRM system, namely an IT technician, a senior systems administrator and a nurse quality health improvement officer. 49 participants were selected, of which 38 were nurses and eight secretaries. A total of 43 questionnaires were distributed to the primary users of the PRM system. Data saturation was reached validating the sample size. The participants used the same system within the hospital, thus at some point data collection feedback provided the same answers reaching a point of saturation.

4.3 Data Collection

Semi-structured interviews and paper based questionnaires were used as data collection methods. The criteria necessary to produce the E-PRM impact assessment framework, was extracted from the questionnaires. The questionnaires were a compilation of open and close ended questions. The target participants for the questionnaires were the nurses and secretaries, as the main users of the PRM system.

5. FINDINGS

5.1 Using the E-PRM system

In attaining objective one of this research, the E-PRM system used at the hospital is called Integrated Healthcare Information Management System (IHCIMS), according to interviewed nurses and secretaries. The system is web-based and also known as an E-health system (Wilson, 2009). Managing patient records is fundamental for determining patient health needs and producing a health work plan accordingly (Poulymenopoulou et al., 201). To use the system, login is a requirement for authentication which acts as a security measure. Users are provided with a username by the system administrators and a password must be created by the users. Patient admissions, discharges and billing are some tasks the secretaries use the E-PRM system for. Nurses primarily use the system for pharmaceutical and ward stock orders. Additionally HIV administration and management of such information can be done by the system. Examples of systems that could be used for E-PRM are MedVault, Indivo, Google Health and Microsoft HealthVault (Economist, 2016).

Paper filling is however still used simultaneously with the IHCIMS. Patient demographics are captured both on the IHCIMS and on paper files. Patient files which are usually completed by nurses are handed to the secretaries to execute discharges or admissions. This however can be a setback as these files can get lost. Nurses nonetheless can also process admissions or discharges on the IHCIMS but only when it is absolutely necessary, citing double work. One of the users, a nurse, said that the use of the IHCIMS is impractical as they record patient health reports on paper and are expected to re-capture the same information on the IHCIMS.

5.2 Effectiveness of the E-PRM according to the Users

In achieving objective three, this section (5.2), 5.3 and 5.4 presents findings on questionnaire feedback from participants. The following results are shown in percentages as it relates to the effectiveness of the IHCIMS, for this section (5.2). Out of 43 participants 72% of the users agreed that the IHCIMS was effective for capturing patient medical records. Furthermore, 69% agreed that the IHCIMS is effective for managing patient records. Regarding effectiveness of accessing patient medical records, 65% agreed that it is effective with 30% undecided and 2% in disagreement. Storing patient medical records is effective according to 70%

of the participants whereas 26% were undecided and 2% disagreed. As such, capturing patient medical records was ranked the highest as it relates to effectiveness of the IHCIMS, followed by storing, managing and finally accessing. Capturing patient information electronically is increasingly the focus of healthcare institutions (Fabbri et al., 2011).

Effectiveness of the E-PRM according to the patients as reported by the users

Patients views as reported by the IHCIMS users. Out of the 43 participants 51% of the participants agreed that patients find the existing IHCIMS system effective. Therefore over 50% of the users stated that patients viewed the existing IHCIMS system to be effective. Wellness is a factor perceived to be the benefit in using ICTs in healthcare (Ruxwana et al., 2010). Additional benefits included easier access and storage of information and that medical practitioners could be assisted in diagnosis of ailments.

5.3 Essentiality of the E-PRM

From the 43 participants 91% of the users concluded that there is a need to use an E-PRM system. 91% agreed that the use of an E-PRM will help improve their PRM. Additionally the use of an E-PRM has been indicated to make work simpler, quicker and effective for the users according to 88% of the participants. Quality, time and communication enhancements are seen to be elements that could be improved by the use of ICTs in healthcare (Fabbri et al., 2011).

5.4 Primary and Secondary Factors impacting/contributing to the use of E-PRM

On the primary elements, the questionnaire results showed that from the 43 participants 70% agreed that patients will benefit from the use of the E-PRM system. According to 91% of the participants finances must be allocated to service the E-PRM system. Equally 91% indicated that the E-PRM must be relevant or customised to the Namibian healthcare setting. Furthermore, 61% of the participants agreed that the infrastructure in the Namibian healthcare is suitable for the E-PRM and related systems. 21% disagreed on the matter. From the participants, 47% of the users agreed on adequate staffing in the healthcare to use the E-PRM, 9% were undecided and 42% disagreed accordingly. Training in using the E-PRM was perceived to be essential according to 96% of the participants. Furthermore 95% of the participants agreed that the E-PRM must be user friendly.

As it relates to secondary elements, the analysis of data shows that from the 43 participants 54% users agreed to an existing culture in using E-PRM in Namibia. 16% of the participants were undecided and 30% disagreed on the matter. 72% of the participants supported the notion of applying ICT policies in the healthcare when utilising the E-PRM. According to 61% of the participants, the Namibian health sector motivates the use of E-PRM, however 21% were undecided on this question and 19% disagreed. 52% of the participants agreed that the GRN is willing to invest in E-PRM, whereas 44% of the users were undecided and 4% disagreed. In conclusion, 74% of the participants agreed that the internet infrastructure in the healthcare was on an acceptable level and that the Namibian health sector supports the use of E-PRM.

5.5 An E-PRM Impact Assessment Framework

The above findings from the questionnaires provided the elements and factors making up the framework as seen in Figure 1, in attaining objective four of this research paper. The title of the framework is Ronald's Patient Records Impact Assessment (R-PRIA) Framework. The framework is specific to the public hospital, consisting of E-PRM users and patients as key features. The central part of the framework linking the two features is the E-PRM system. As such both features have a relationship with the E-PRM. The users have a direct relationship with the E-PRM system, whereas the patients are affected by the use of the system denoting an indirect relationship with the E-PRM system. The three E-PRM system elements are the effectiveness of the E-PRM system, essentiality of the E-PRM, primary and secondary factors. These elements are further discussed as follows.

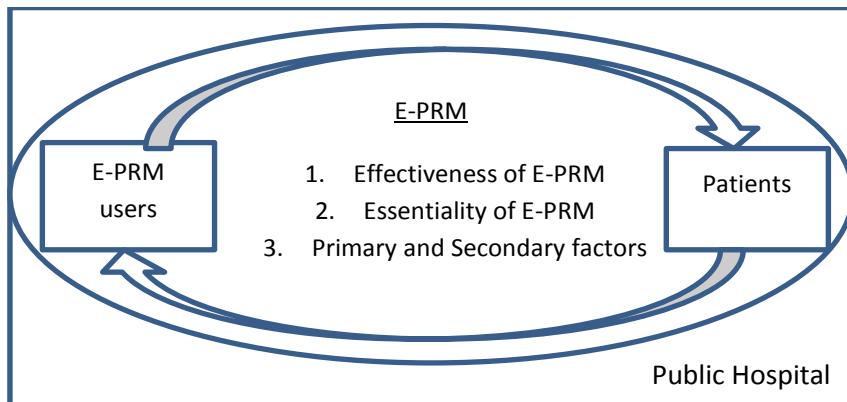


Figure 1. Ronald's Patient Records Impact Assessment (R-PRIA) Framework

5.6 Effectiveness of E-PRM System

The effectiveness of the system consist of two viewpoints, namely the effectiveness of existing E-PRM as viewed by the users and the effectiveness of existing E-PRM according to the patients as perceived by the users. The latter is discussed in the next section. As for the former, effectiveness of the E-PRM consists of entering, organizing, retrieving and storing patient medical records. When these processes are achieved easily with no obstacles, it is classified as effective. Obstacles can refer to incompetent users or technical problems in using the E-PRM system. In the study entering patient medical records using the E-PRM system was rated highest effective by secretaries and nurses, thereafter storing, organizing and lastly retrieval. These aspects are thus a fundamental part of the E-PRM system which should be assessed.

5.7 Effectiveness of the E-PRM

The users of the E-PRM were asked how the patients viewed the system. 51% of the participants agreed that patients found the system beneficial. Thus it can be said that the effectiveness of the E-PRM system indirectly is a gain to the patients, in that admissions and discharges could be done faster. This contributed to patients being a feature in the framework towards impact assessment of the E-PRM system. Questions such as how does the use of the E-PRM system benefit patients should be asked.

5.8 Essentiality of E-PRM

The results from the questionnaire show that 91% of the participants indicated the significance in using the E-PRM. According to the users it would support the PRM processes. Simpler, quicker and effective work optimisations are elements supported according to 88% of the participants. It is therefore vital to assess the importance and value addition brought about by the use of E-PRM in the healthcare. Questions such as what E-PRM system is used for PRM and how does this system make work simpler, quicker and effective should be asked.

5.9 Primary and Secondary Factors contributing to the use of E-PRM

Primary factors refer to factors that influence the use of E-PRM. Lists of these factors are:

- Patient aid: Findings show that 70% of the participants indicated that patients would benefit from the system. If the system does not benefit the patient as the customer to the healthcare, the system would be unnecessary.

- Resources: 91% of the participants indicated that financial resources must be allocated by the healthcare towards purchase and maintenance of ICT for PRM. It is important to assess the budgeting or availability of financial resources at the healthcare to determine affordability and up keep of the technologies.
- Hospital aligned electronic PRM: 91% of the participants indicated that the ICT utilised for PRM must suit the healthcare context. Healthcare PRM requirements assessment is necessary to determine if the system meets the local healthcare requirements.
- Infrastructure: Healthcare infrastructure to support the use of ICT for PRM should be assessed to determine if these technologies can be housed and supported.
- Staff capacity: Staff to use the ICT for PRM must be assessed to determine the support efforts from the IT administrators for the users of the system.
- Training: 96% of the participants agreed that training must be provided to the users of the PRM system. Assessment of training needs is necessary to ensure competency of the users of the system and help boost interest in using the system.

Secondary factors refer to factors that influence the use of E-PRM indirectly. A list of these factors is:

- ICT Culture: 54% of the participants indicated that there is a culture in using ICT for PRM. Assessing an existence of such a culture will help determine why it is that healthcare participants use or does not use the system and help emphasise the importance of a good working environment.
- ICT policies: 52% of the participants agreed that the healthcare encourage the use of ICT policies. Assessment of existing ICT policies are needed to determine the presence and extend to which these policies are applied. Additionally whether these policies are plain enough to be followed. The policies would be a control mechanism in determining the does and don'ts of using ICT for PRM.
- Employee motivation: 61% of the participants agreed that healthcare encouraged the use of ICT for PRM. Assessment of what mechanisms by the healthcare are in place to encourage the use of ICT for PRM.
- GRN investments: 52% of the participants agreed that the Namibian government is willing to invest in ICT for PRM. Assessing the need to motivate investment of ICT for PRM by reporting on effectiveness of the PRM system would be necessary.
- Internet infrastructure: 74% of the participants agreed that there is a satisfactory level of the internet infrastructure in the healthcare to support the use of ICT for PRM. Assessing the internet infrastructure in the healthcare for availability and effectiveness as it would be the backbone in connecting different departments in the healthcare for PRM would be vital.

6. BUSINESS BENEFITS

The R-PRIA Framework would aid the hospital executives and management at the MOHSS to make informed decisions regarding investments in ICTs. Evidence to support huge investments in the use of ICTs in the healthcare is lacking, resulting in uncertainty on whether these technologies actually aid work processes in contributing towards better service delivery. The framework could be used to guide the gauging of impact of existing ICTs, E-PRM in particular, to substantiate investment efforts. This in turn would ensure appropriate governance of GRN finances. As such management would no longer need to shoot in the dark as it relates to the continued maintenance of existing ICTs or the need to invest in newer technologies that would possibly produce an impact towards enhanced healthcare service delivery.

7. CONCLUSION

The GRN has invested significantly in HIS for purposes of quality healthcare. Such an investment can however not be substantiated, as it is not clear whether ICTs have an impact in the healthcare. Thus an E-PRM framework to assess the impact is lacking to aid in measuring impact. The developed framework is filling this research gap. The framework can be used as a tool that would aid in measuring the change brought about by ICTs on PRM. The application of the framework may yield results that could aid Namibian

healthcare decision makers on whether or not further investments on ICTs, specific to E-PRM, would be necessary based on the outcomes provided by the framework. The drive of the MOHSS is on ensuring that public healthcare which was designed for those who cannot afford private healthcare services is of the right quality. E-PRM automates the process of admission, discharges and billing to mention a few. A number of impact assessment frameworks ranging from Sustainable livelihoods (SL) approach, Sen's capability framework and Rural ICT comprehensive evaluation framework (RICT-CEF) were reviewed. The purpose of the reviewed frameworks was to identify criteria that could be used in the development of an E-PRM impact assessment framework specific to the Namibian public healthcare. The developed E-PRM framework (R-PRIA) was the result of this research paper.

For future research work this study recommends that aspects that were not covered as a result of the focus and scope of the study be considered by researchers. This could include investigating how the use of the E-PRM directly affects patients, as a key feature in the framework, by collecting data from them, as they are the recipients of service provided by the healthcare institutions. Another recommendation would be to investigate and include cost-related aspects in using the E-PRM, which would aid in establishing the impact of costs related to using the E-PRM. Additionally, the inclusion of change management aspects towards technology adoption by healthcare personnel would be beneficial in shedding light on the effectiveness of change agents in the healthcare, motivating the use of ICTs for work processes. Further recommendation would be to investigate the functionalities of the E-PRM system to determine if the final work outputs related to the PRM are equal to the full system potential. This would establish whether or not the system is being fully utilized. Finally, testing the framework in other healthcare settings to determine if the developed framework can be applied to any public or private healthcare facility is recommended.

REFERENCES

- Economist, N. 2016. *No Need for State Hospital Queues*. [online] Available at: <https://economist.com.na/17299/health/no-need-for-state-hospital-queues/> [Accessed 15 Sept. 2016].
- Fabbri, D., LeFevre, K. and Hanauer, D. A. 2011. Explaining Accesses to Electronic Health Records. *Accumulated Computer Machinery*, pp. 10-11.
- IFRC. 2005. *Impact Assessment*. [online] Available at: <http://www.ifrc.org/docs/evaluations/impact-handbook.pdf> [Accessed 27 Sept. 2016].
- Jafari, M., Safavi-Naini, R. and Sheppard, N. P. 2011. A Rights Management Approach to Protection of Privacy in a Cloud of Electronic Health Records. *ACM Digital Library*, pp. 23-24.
- Khan, T. and Edwards, D. 2012. Assessment of National Health Information Systems: Ministry of Health and Social Services (MOHSS), Republic of Namibia. *ProQuest*, pp. 3-7.
- Lee, B. C. 2013. Human Cognitive Performance in Healthcare Information System Environment and Its Application on Nursing Tasks. *Open Access Dissertations*, pp. 8-11.
- Mthoko, H. and Khene, C. 2015. Assessing Outcome and Impact: Towards a Comprehensive Evaluation Approach in ICT4D. *Accumulated Computer Machinery*, pp. 1-2.
- Nyella, E. 2009. Challenges in Health Information Systems Integration: Zanzibar Experience. In: *Information and Communication Technologies and Development (ICTD), 2009 International Conference*. Doha: IEEE, pp. 1-2.
- Poulymenopoulou, M., Malamateniou, F. and Vassilacopoulos, G. 2014. E-EPR: a Workflow-Based Electronic Emergency Patient Record. *Accumulated Computer Machinery*, pp. 91-92.
- Ruxwana, N., Herselman, M. and Conradie, P. 2010. ICT Applications as E-health Solutions in Rural Healthcare in the Eastern Cape Province of South Africa. *Health Information Management Journal*, pp. 24-25.
- Streatfield, D. and Markless, S. 2009. What is Impact Assessment and Why is it Important? *Emerald Insight*, pp. 134-141.
- Tossy, T. 2014. Major Challenges and Constraint of Integrating Health Information Systems in African Countries: A Namibian experience. *International Journal of Information and Communication Technology Research*, pp. 278-279.
- Walker, J. P. 2015. The Effect of Neoliberalism on Capabilities: Evaluating the Case of Mexico. *ProQuest*, pp. 34-35.
- Wilson, B. 2009. Economic Diversification and Prospects for Sustainable Rural Livelihoods in a Dryland Agrarian Village: A Case Study in Bijapur District Karnataka, India. *ProQuest*, pp. 5-35.

TEACHING MOBILE DEVELOPMENT WITH APP INVENTOR AND PAIR PROGRAMMING

Paolo Musmarra

University of Salerno, Department of Computer Science, Italy

ABSTRACT

The relationship of the teenagers with the new technologies is very complex and it is the center of attention today for the schools. Everyone is always connected: some searches show that the teenagers under 18 spend online about ten hours every day and are connected also for 24 hours with mobile devices and thanks to free wi-fi. Pair programming is a collaboration paradigm that has been strongly adopted in particular in the computer science education. In fact, Pair Programming has demonstrated benefits for many aspects in education, but unique concerns of mobile software design raise questions about the effectiveness of Pair Programming in this evolving field. This paper probes unique challenges for Pair Programming when used in mobile software design classes, focusing on four mobile design topics: dealing with interface and data management, recording and playing audio via mobile device and App Inventor components, using sensors and collecting GPS data from mobile device, using a TinyDB component of App Inventor for storage data. The paper highlights successes and challenges for Pair Programming and mobile applications, concluding with recommendations on building assignments.

KEYWORDS

Computer Science Education, Pair Programming, Mobile development, App Inventor

1. INTRODUCTION

Evidence suggests that pair programming (PP) can improve students' programming efficiency and ability. In fact, numerous studies have found that both code quality and efficiency of student pairs is much greater than solo students (Cockburn & Williams, 2001; Nagappan et al, 2013; Rahimian & Ramsin, 2008). Agile practices provide practices and techniques to improve software developers to develop software quickly and efficiently, valuing working software and face-to-face interactions over processes, events and documentation (Manifesto for Agile Software Development, 2001; Seyam and Galal-Edeen, 2011). One of the most important agile practices is *pair programming*, which is an approach that features two developers working on the same development task (Cockburn & Williams, 2001; Beck & Andres, 2004). One developer, called the *driver* initially writes code, but the other, called the *navigator*, watches the driver, provides advice, and seeks to grasp the overall picture of the task currently under development (Beck & Andres, 2004). These roles are not fixed, but two roles are exchanged at regular intervals, and the two developers together become the owners of the resulting product, and also become knowledgeable about the development topic. The next section provides background on research that has been done on applying PP in classroom settings. The paper proceeds with showing how we apply PP in our mobile development class and how we use app inventor to create an own app. We conclude the paper with a discussion on the findings in four PP lessons we performed, with a set of guidelines that would help teachers facilitate PP lessons in their own mobile development classes.

2. RELATED WORK

PP also gives students a major sense of responsibility, which leads to increased retention rates (Simon & Hanks, 2007), even for non-majors (Kuppuswami & Vivekanandan, 2004). In a study comparing individual students to paired students, individual students tended to put off critical thinking activities, while

student pairs spent the beginning phases of their programming task discussing their plan (Katriou & Tolias, 2009). Some students found PP to be time consuming, particularly if partners' opinions differ and/or tasks are not distributed evenly (Cockburn & Williams, 2001). Most of these prior studies focused on the final outcome of PP, but recent research has begun to investigate the process of pair programming. Individual students have been observed to jump right into programming and address design concerns during testing activities, while student pairs spent more than 30 minutes at the beginning of the activity discussing the requirements before starting to program (Katriou & Tolias, 2009).

3. APP INVENTOR: STEPS AND METHODS

App Inventor (AI) has been successfully used to teach introductory programming concepts in both secondary and higher education courses (Soares & Martin, 2014). O' Malley et other describe the mobile learning "Any sort of learning that happens when the learner is not at a fixed, predetermined location, or learning that happens when the learner takes advantage of the learning opportunities offered by mobile technologies" (O' Malley et al, 2003).

Below we describe, the main steps to create own mobile app with AI are described:

1. Decide what you want your app to do, chose the needed components and the non-visible components, and set the properties for each component using Component Designer.
2. Learn to work with the Blocks Editor, using blocks of codes as a visual code and connected like puzzle pieces, to make the components work just drag-and-drop it into the Blocks Editor.
3. Test the app with tablets, smartphone or cell phone, or an emulator. In this case, we prefer to use the cell phones, because it's more fun that way, and it's easier to test phone-specific functionality.

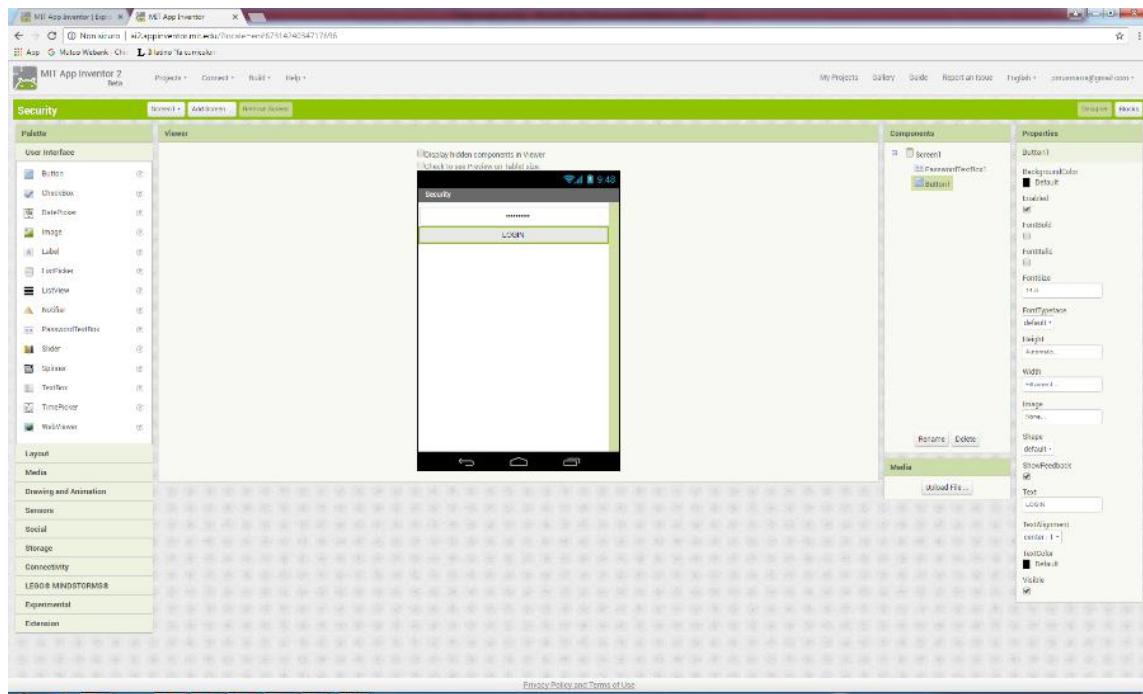


Figure 1. App Inventor UI

4. PAIR PROGRAMMING FOR MOBILE DEVELOPMENT CLASS

In this paper we are focused in PP for classroom settings, in particular in higher education who are enrolled in computer science (CS). One of the original studies of PP suggested that it is a promising approach to use as a pedagogical tool due to its capability of increasing learning capacity (Cockburn & Williams, 2001), with foundational studies demonstrating how PP as a pedagogical tool that can help students enhance productivity and learn technology design and programming (Nagappan et al, 2003; Simon & Hanks, 2007). A review of PP studies shows areas of success with evidence suggesting that PP could improve enjoyment, reduce workload, increase student's confidence level, enhance course completion rate, improve exam performance (Salleh, Mendes, & Grundy, 2011). Rahimian and Ramsin (2008) show that mobile development has certain issues that differentiates it from web development related to communication, portability, standards and protocols, mobility, storage, and user interfaces. Tigrek and Obadat (2012) proposed a pedagogy for teaching smartphone programming, which considered PP as a method to enhance guided teamwork. Most recent research paper highlighted teaching programming for smart watches as part of an undergraduate mobile development class (Esakia, Niu & McCrickard, 2015). This research shows the need to consider several extensions, as many kind of sensors or variable platform, while managing with mobile development. Williams et al. (2008) presented a study of using PP for teaching human computer interaction (hci) class. In this paper we explore how PP using can be applied to create mobile apps, and describe challenges associated with students developing mobile apps in pairs.

5. PAIR PROGRAMMING LESSONS

We conducted the lessons with expert reviews with two pairs of android developers specifically experienced with AI tool. Two active walkthrough PP sessions were conducted to get the participants expert reviews about the process. The expert reviews also provided some recommendations that the experts think would enhance students' performance, such as considering pairing preferences, handling quality requirements as basic requirements. In detail, we applied PP in class activities for a mobile development course. To achieve these goals, we facilitated five 90-minute PP sessions for the 48-student class, which covered four different aspects related to mobile development:

- 1- Dealing with user interface and data management
- 2- Recording and playing audio via mobile device and App Inventor components
- 3- Using accelerometer and GPS data from mobile device
- 4- Using a TinyDB App Inventor component for storage data

The driver and navigator exchanged roles every 15 minutes. The class was managed by a main facilitator for PP and two assistants to answer questions concerning AI and programming. Every lesson started with a 10-minute introduction that worked like a retrospective analysis for previous lessons and preparation for the new one. At the end of every lesson (except the first introductory one), students were required to fill-in an online questionnaire after they deliver their app. It's important to underline that was not mandatory for students to answer every question, because that questions might have different number of responses within the same lesson. Below was synthesized what is performed in every lesson, and the findings resulting from each lesson.

5.1 Lesson 1

An initial lesson featuring a 20-minute overview of PP and a 10-minute presentation about the activity, which was on creating a simple app with a user interface for a password like login screen and store the personal data. PP rounds were 15 minutes each, so each student worked twice as a driver and twice as a navigator. Students could select their partner, select activities, and create their own UI.

5.1.1 Findings

- 1- To help a cooperative learning, students got annoyed when asked to work on only one screen (i.e., no second laptop).
- 2- Students didn't like to switch while in the middle of coding certain code segments, and they ignored the time checks until the driver finishes the part in hand happened subconsciously while others argued that it makes them work faster.

5.1.2 Questionnaires

There was no questionnaire for this lesson. It was meant only to introduce students to PP methodologies.

5.1.3 Considerations

After beginning, most students immediately started coding, giving no time to consider the application design. Indeed, they did not design a draft interface until second or third rounds. At this lesson, many students struggled to adapt their work to fit the activity's minimal requirements. Because students did not pay attention to the time box, so the facilitator asks students to switch roles every 15 minutes.

5.2 Lesson 2

The activity exercised image processing tasks. The session started with instructor comments on student work during the previous PP session, mainly related to the roles of driver and navigator. Students were randomly paired based on a seating location rubric.

5.2.1 Findings

- 1- Students got more used to switching roles and stick to the assigned turns. However, they needed the facilitator to manage time and remind them when to switch
- 2- Students managed to give themselves more flexibility in switching time. i.e. to wait until driver finishes what s/he is working on before switching.
- 3- It was still hard to keep pairs working on the same screen, as the navigator sought to use another laptop screen for other supporting functions as online search, displaying class material.

5.2.2 Questionnaires

Out of the 36 responses to the open ended question: "How would you describe your experience of today's Pair Programming session?" 30 students provided positive feedback about the lesson, while 6 had minor complaints especially with problems with code, classroom environment, or about App Inventor features.

5.2.3 Considerations

During this session, the number of questions raised by students was noticeably fewer than in lesson 1. Navigators were allowed to hold the mobile devices to test code augmenting to the traditional PP tasks of watching and testing with the driver on coding. This is a special case for mobile application development as rapid parallel testing is done on another screen, rather than the development screen as in web/desktop apps.

5.3 Lesson 3

Students had a previous activity on GPS, and for this reason in lesson 3 we can combining GPS and gyroscope. Students were assigned to random pairs. To get students to feel more involved with the sessions, the facilitator displayed anonymous comments from the previous session's questionnaire and discussed them with students during the introduction.

5.3.1 Findings

- 1- Navigators and drivers want to be more active by looking things up while the driver is coding
- 2- Students, again, tried to use another laptop screen for tasks other than coding, not stopping until the facilitator asks them to go back to a single screen

3- PP seems to work better when students have prior knowledge about the activity, as this increases their ability to exchange knowledge and learn from each other.

5.3.2 Questionnaires

29 out of the 36 recorded responses indicated they were able to deliver more than 80% of the required functionality. This rate caused the level of satisfaction in students' answers to the open ended questions to be higher than the previous session. Many students indicated that they were able to understand how to use GPS thanks to discussions with their partners.

5.3.3 Considerations

Since students had a previous lab on GPS, they kept referring to prior material and finding complementary resources online. Although the activity introduction asked students to consider how users would use the application, only five groups included guidance for users to be able to use their apps. At the end of the session the facilitator discussed with them how a user will use their app.

5.4 Lesson 4

The objective of the activity was to familiarize students with TinyDB components. This last PP session of the semester added the statement “You are required to include any components that help users understand and use your app”. Based on previous session results, students were assigned to random pairs..

5.4.1 Findings

- 1- To address TinyDB component with AI need to be clearly requested rather than broadly referenced; otherwise students do not realize the need to dedicate time for it.
- 2- Students are able to practice PP on their own after training.
- 3- Online resources and course material are important during the session. Planning for PP sessions should incorporate such resources within the development process.

5.4.2 Questionnaires

29 out of 36 responses provided positive feedback regarding the lesson. Some students indicated that PP helped them get the required skills needed to work with the activity with no need to ask for instructor's help.

5.4.3 Considerations

Students were able to explore how to deal with the activities on their own through discussions and web search. About half of the groups missed the first switch and needed to be reminded by the facilitator, but only 3 pairs missed it on the second round. On the third and final switch, all pairs managed to switch without reminders.

6. RESULTS

Questionnaire results showed how PP affects student performance on in-class activities. For example, the percentage of functionality delivered per session changes with the topic being introduced in such session. In details, It is clear that the using accelerometer and GPS data from mobile device (lesson 3) and using TinyDB components (lesson 4) activity were those with the increase number of assignments completed. Therefore, student answers to open-ended questions provided more positive feedback (based on the terms they use to describe the session)—increasingly so as sessions go on. While this suggests PP was helping students deal with complex problems, it also shows that students work better with PP when they get sufficient practice time. Along with information from questionnaires, our observations provided different aspects on how students worked with PP.

7. SUGGESTIONS

In this section we propose some practices, based on our explanatory study, that can help instructors get the best of PP in tune with course learning objectives. Below is a list of our suggestions.

1- Decide on the strategy to be used to introduce UI/UX requirements based on the amount of time available for each session. Lesson 3 and 4 show students giving more attention to UI/UX issues much more than the previous sessions. In this situations the instructor should dedicate the introduction of the lesson to only work on UI/UX issues if students will not use PP themselves.

2- Allow students to choose partners on the first session, then move to random pairing. The first session lets students unfamiliar with PP focus on learning how PP works instead of figuring out how to deal with an unfamiliar work partner. Students tend to be more productive when they deal with partner they may not know as well. Two possible reasons for this are that students are more relaxed when they work with someone they know, and they gain more knowledge from someone whom they did not have interaction with before.

8. CONCLUSIONS

This paper highlights that mobile development requires further investigation on how PP will perform and how it might change. We present App Inventor, a visual programming tools that enables students to see and test your own app immediately. PP seems to provide a reasonable approach to handle such challenges. For PP to perform better in mobile development classes, instructors should consider how PP has traditionally been applied, as well as why some changes may positively influence student learning and performance. We expect that this work introduces more questions than it answers, and we look forward to learning about the answers through future PP efforts.

REFERENCES

- Beck, K. and Andres, C., 2004. *Extreme ProgrammingExplained: Embrace Change*. Addison-Wesley Professional
- Cockburn, A. and Williams, L., 2001. The Costs and Benefits of Pair Programming. In *Proceedings of Second International Conference on Extreme Programming and Flexible Processes in Software Engineering* (pp. 223-243).
- Esakia, A., Niu, S., and McCrickard, D.S., 2015. Augmenting Undergraduate Computer Science Education with Programmable Smartwatches. In *Proceedings of the Proceedings of the 46th ACM Technical Symposium on Computer Science Education* (Kansas City, Missouri, USA2015), ACM, 2677285, 66-71.
- Katriou, S. and Tolias, E., 2009. From twin training to pair programming. In *Proceedings of the 2nd India software engineering conference, ISEC '09* ACM, New York, USA.
- Kuppuswami, S. and Vivekanandan, K., 2004. The effects of pair programming on learning efficiency in short programming assignments. *Journal of Information Education* 3, 2, 251-266
- Manifesto for Agile Software Development (2001). <http://agilemanifesto.org> Accessed: June 15, 2015
- McDowell, C., Werner, L., Bullock, H., and Fernald, J., 2006. Pair programming improves student retention, confidence, and program quality. *Communications of ACM* 49, 8, 90-95.
- Nagappan, N., Williams, L., Ferzli, M., Wiebe, E., Yang, K., Miller, C., and Balik, S., 2003. Improving the CS1 experience with pair programming. In *Proceedings of the 34th Technical Symposium on Computer Science Education (SIGCSE '03)*.
- O'Malley, C., Vavoula, G., Glew, JpP., Taylor , J., Sharples, M. & Lefrere, P. (2003). *Guidelines for learning/Teaching/Tutoring in a Mobile Environment. MOBIlearn project report*, D4.
- Rahimian, V. and Ramsin, R., 2008. Designing an agile methodology for mobile software development: A hybrid method engineering approach. In *Research Challenges in Information Science, 2008. RCIS 2008. Second International Conference on*, 337-342.
- Salleh, N., Mendes, E., and Grundy, J., 2011. Empirical Studies of Pair Programming for CS/SE Teaching in Higher Education: A Systematic Literature Review. *Software Engineering, IEEE Transactions on* 37, 4, 509-525..
- Seyam, M. and Galal-Edeen, G.H., 2011. Traditional versus Agile: The Tragile Framework for Information Systems development. *International Journal of Software Engineering (IJSE)* 4, 1, 63-93.

- Simon, B. and Hanks, B., 2007. First Year Students' Impressions of Pair Programming in CS1. In *Proceedings of the Third International Computing Education Research Workshop*, 73-86.
- Soares, A. and Martin, N. L., 2014. Teaching Non-Beginner Programmers with App Inventor: Survey Results and Implications. In *Proceedings of the Information Systems Educators Conference* ISSN, Vol. 2167. Citeseer, 1435.
- Tigrek, S. and Obadat, M., 2012. Teaching smartphones programming using (Android Java): Pedagogy and innovation. In *Information Technology Based Higher Education and Training (ITHET), 2012 International Conference on*, 1-7.
- Williams, L., McCrickard, S., Layman, L., and Hussein, K., 2008. Eleven Guidelines for Implementing Pair Programming in the Classroom. In *Agile, 2008. AGILE '08. Conference*, 445-452.

INCLUSIVE JOBS: A PROJECT TO HELP PEOPLE WITH IMPAIRMENT FIND OPPORTUNITIES – HEURISTIC EVALUATION TO DIAGNOSE USABILITY PROBLEMS

Felipe Araujo and Adriano Bernardo Renzi
Serviço Nacional de Aprendizagem Comercial, Brazil

ABSTRACT

Inclusive Jobs is a web project to help empower impaired people who struggle to reach professional demands of the market. It is an experimental project, part of the interactive design post-graduate program at SENAC. The whole project has 4 stages of research and development: mapping users' expectations, structuring the information, developing prototype and usability testing. This paper presents the project's high fidelity prototype evaluating phase, aspiring to validate the development directions and find usability problems for further development and discussion. The usability analysis is based on Nielsen and Molich's Heuristic Evaluation method.

KEYWORDS

Information Systems, Usability, Heuristics, Employment

1. JOB OPPORTUNITIES FOR IMPAIRED PEOPLE

The collected data from unsystematic interviews with impaired people and human resources professionals – based on Muccielli's directives (1979) – surfaces some integration difficulties of impaired workers with the market demands of Brazil. The country has been through an economic crisis in the last 3 years that has been affecting deeply its economic structure, resulting in lack of job opportunities and depressive expectations from society.

Although the crisis is considered recent, the diminishing of job opportunities seems to have been affecting impaired workers for a much longer time, independent of the country's crisis. According to collected data from the interviews, impaired participants demonstrate to have difficulties to find jobs directly related to their professional expertise. Instead, participants have shown a tendency to grab any available job, which most of the time are linked to professional areas not related to their educational investment.

From the human resources professionals' point of view, they signalize that it is hard to find qualified impaired workers, fit for the demands of specific job positions. In which, without the proper qualifications, they end up not hiring them for the job openings.

In Brazil, it is mandatory by law that companies offer 2-5% of their job opportunities to impaired people. It is called "*lei de cotas N° 8.213/91*", and according to the country's annual report regarding social development, there should be 828,000 job opportunities throughout national companies. However, only 403,200 of the positions are filled, less than 50% of total opportunities, leaving still 424,800 opportunities waiting to be filled by impaired professionals.

Most of the interviewed impaired participants search for jobs using well-known Brazilian job hunting websites, such as Catho.com and Vagas.com, or search engines (Google, Bing etc). In either choice, participants tend to start looking for a job using their desktop computers or tablet, for their bigger display to gather information and interact. Although inserted in their day-to-day chores, smartphones are not often used as an option for job search.

The understanding of users' necessities and expectations, the mapping of market opportunities and case studies were part of earlier stages of this research and helped the first drafts for the development of the Inclusive Jobs project. Sequential stages enveloped information architecture studies, wireframe proposals, personas and low-resolution prototyping. The 3 proposed personas represented users with impairment and users who act as recruiters for companies:

- Agatha is 32 years old. She is server management specialist with a bachelor degree in computer science. She has movement impairment and although having 5 year experienced as a project manager, she has been having difficulties in finding a job after her accident and movement impairment.
- Bartholomeu is 45 years old. He is an administration assistant and has a high school diploma. He has hearing and movement impairment. He has always operated basic functions and worked in jobs that required low qualification. Due to his lack of qualifications, he has difficulties in finding better opportunities.
- Erica is 31 years old. She is a Human Resources Recruiter and has a bachelor degree and post-graduate specialization. Erica is responsible for selecting candidates for hiring, but has difficulties in finding qualified professionals among candidates with impairment.

Due to noticeable users preferences and mapped cultural conventions, the Inclusive Jobs web project has been developed to be an adaptive system web portal, instead of an App. Inspired by cases such as Kerjbilitas, the project intends to bring social impact and bring opportunities for life change. The major proposal of the Inclusive Jobs project is to provide information on job opportunities directed for impaired people, to present professional growth opportunities and to create a bridge between employers and employees. The difference between Inclusive jobs and other employment websites is its proposal to be exclusive for people with impairment and help companies fill the 2-5% of their opportunities.

This prototype focus on users with mobility impairment needs as a first step for analysis and discussion. A project with a broader reach, involving intellectual and visual impairment users will need further research and support from web accessibility tools.

As a tool to concentrate relevant information about impaired users and help employers find the best candidates to fill the 2-5%, obliged by law, the project has to present candidates' best assets up front and help impaired users to prepare and post their curriculum with cognitive actions. In order to make the process easier, the impaired user can sign up in the system and update his professional profile, filling an online step-by-step curriculum form. The system provides flexibility and users can also upload resumes in PDF or DOC formats. If the user already has a LinkedIn account, it is possible to transfer his/her information directly. The Inclusive Jobs system reformat the information to better correlate users' profile information mapping with opportunities compatible with their expertise and possibilities of integration with the inclusive law "*lei de cota N° 8.213/91*". The disclosed information about professionals have to be objective also to recruiters and human resources professionals, as they are the complementary users of the system and the project intends to help explicit candidates' skills that are relevant to employers. For users, despite any impairment, that lack skills expected to certain opportunities, the project intends to provide courses, programs and workshops for users' skill improvement, so they can better reach job requirements and fill the 424,8 thousand gap in Brazil. These actions need to go beyond the portal itself and embrace knowledge from pervasive user experience (Renzi 2017) theories, integrating user experience journey with service design. Based on interview results and preliminary research on benchmark services and their interactive flow structure proposal, users pass through 4 steps of experience: search > filter > position opportunities > application.

The Inclusive Jobs project intends to shorten the phases on the impaired users' side to 2 steps: position opportunities > application. After users follow the system steps to build their skill profile, the system will bring to each user the best suitable job options and courses for improvement. For the recruiters, the system presents match options for the job selection and takes into account if candidates' impairment could be impeding to specific positions. The development of the prototype tried to simulate the steps, in which first time users would follow to be part of the program (figure 1).

The prototype for experimentation has all major interaction functionalities, to enable specialists explore throughout all expected steps to prepare the skills profile and apply for a job. Since the availability of users with impairment was restricted during this phase, the use of heuristic evaluation was chosen as a starting method to pinpoint usability problems in the prototype, before testing with final users. The preference befalls on Nielsen and Molich heuristic evaluation, due to users' preference on using desktop to access website and search for jobs. Heuristic evaluations for smartphones (Inostrozza and Rusu 2013) or tabletop (Apted *et al.* 2014) did not fit properly the research. The prior interview results, mapping of users' journey and personas were presented to the specialists before the evaluation.



Figure 1. Inclusive Jobs project homepage prototype

2. HEURISTIC EVALUATION

Heuristic evaluation is a method created by Jakob Nielsen and Rolf Molich in 1990, during the second wave of computing (Renzi 2017), to help determine usability problems in systems with interfaces. The evaluation is conducted with the help of usability experts, who analyze and test all interface pages and mark usability problems and their respective gravity. According to Nielsen, it is difficult to apply the technique with just one expert, as one individual can't map all problems at once. The author has experimented with the method in 1992 with 19 experts to better understand the relationship between number of experts and number of problems found. The system used to conduct this experiment had 16 usability problems in total (Nielsen 1995). Using 19 experts, all 16 problems were found in a variety of ways: experts who pointed out many problems not necessarily found the difficult ones, and experts who found much less problems, got hold of the hard ones to find.

Although the 19 experts were able to find all 16 problems, Nielsen denotes that only 3 to 5 experts are necessary to conduct the technique. After observing heuristic evaluation in more than 6 projects, it was noted that each expert individually could locate 35% of usability problems. It is possible to reach better results when aggregating observations from different experts and Nielsen proposes that using 5 experts can surface 75% of the problems. But adding more than 5 experts would not bring a substantial increase in the percentage value. For this research, 9 experts were invited to conduct the evaluation, but only 4 agreed to be part of the research.

The selection of experts for the heuristic evaluation must consider their professional experience in digital projects, their knowledge regarding usability and human-computer interaction, familiarity with the heuristics and previous experience with the method. All 4 experts, who agreed to participate, have at least 5 years of experience as interaction designer, and one of them has a post-graduate degree specialization in interaction design. After selecting the experts, the heuristic evaluation follows 3 phases in sequential order:

- (1) Each expert receives the prototype to be analyzed and the evaluation guide. The document has the description of the project purpose, all 10 heuristics and a form to include the several usability problems and their respective heuristics. Each expert insert the usability problems in the guiding form without direct contact with the other evaluators. It is important to note that the usual 10 heuristics can be modified and more heuristics can be increased, as fit to the necessities of the project and context (Renzi and Freitas 2013);
- (2) The researcher collects all noted problems (and respective heuristics) and compiles them in one single document. The form is sent back to all experts and they include a gravity degree to each problem (from 0 to 4);

- (3) All forms are collected once more and the researcher lists all problems in one final document. The gravity of each problem is a weighted average from all experts' opinions. If a pointed problem brings up gravity values far apart from different experts, a meeting should take place to reach common sense.

The 10 heuristics used in this research were the original ones proposed by Nielsen and Molich:

- (1) Visibility of system status: The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
- (2) Match between system and the real world: The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
- (3) User control and freedom: Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
- (4) Consistency and standards: Users should not have to wonder whether different words, situations, or actions mean the same thing.
- (5) Error prevention: Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
- (6) Recognition rather than recall: Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
- (7) Flexibility and efficiency of use: Accelerators — unseen by the novice user — may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
- (8) Aesthetic and minimalist design: Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
- (9) Help users recognize, diagnose, and recover from errors: Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
- (10) Help and documentation: Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

The gravity of each noted problems varies from 0 to 4 and have to be included after all usability problems are compiled: 0 (zero) – not considered a usability problem; 1 (one) – aesthetic problem; 2 (two) – minor usability problem; 3 (three) – major usability problem; 4 (four) – catastrophic usability problem.

3. RESULTS

The four experts found 80 usability problems. Following Nielsen's gravity chart, 3 of them were considered catastrophic, 25 were considered as major problems, 35 considered as minor problems. The rest was marked as aesthetic problem and not a usability problem at all. Figure 2 shows, comparatively, the 80 problems distributed throughout the gravity scale.

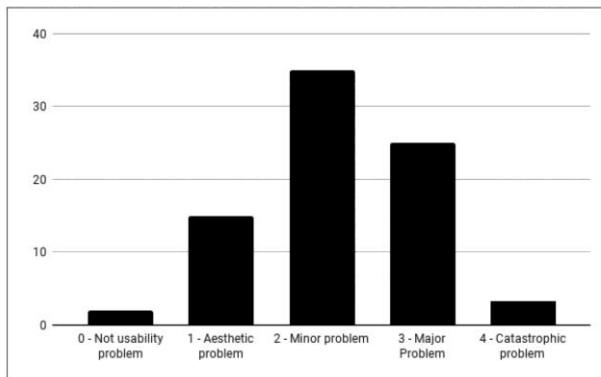


Figure 2. Usability problems categorized by degrees of severity

The most cited heuristics by the evaluators related to usability problems with the prototype were heuristics 4 and 8. Both linked to a total of 48 noted problems and related to visual design aspects: consistency and standards; aesthetic and minimalist design. Heuristic 3 (user control and freedom) was linked to 15 noted problems and heuristic 1 (visibility of system status) was linked to 11 problems (figure 3).

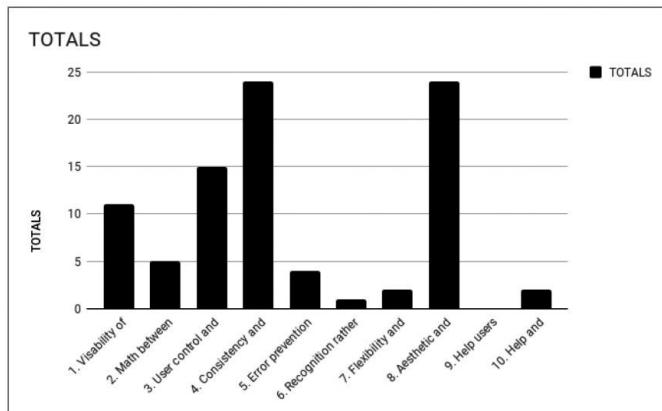


Figure 3. Usability problems categorized by heuristics

3.1 Catastrophic Problems

Figure 4. After education information is included, users can only edit, but never delete it

The problems with most urgency to be fixed in the prototype are mostly related to user profile and curriculum edition, both important items to help users get a job. When users start to fill the form with their skills and abilities, it is not possible to advance without including education information (heuristic 3), and this could exclude users without formal education. And after adding education information it is not possible to exclude it (heuristic 3) – figure 4. Navigation throughout the prototype also showed basic mistakes, unnoticed while developing, but surely pointed out as catastrophic problems by experts in their exploring evaluation: it is not always possible to return to the homepage.

3.2 Major Problems

Although not an emergency as in the catastrophic problems, major issues pointed by experts need attention as soon as possible to avoid users' interaction problems with the system. Among the 25 identified problems, this paper presents the ones considered most relevant for the project. As in the catastrophic problems, most of the major problems pointed out are related to filling the skills profile and curriculum information.

While inserting the users' data in the curriculum page, there is no indication of the document's completion (heuristic 1), nor possibility of including a profile photo (heuristic 3). There is no assistance to help users complete the steps of this important phase (heuristic 10) and no alert in case there is a mistake in the form (heuristic 9). There is no indication of which fields are required and which are optional (heuristic 6). There is an evolution of the process bar in yellow (figure 5), but no phase is clickable, leaving no possibility for users to jump back and forward within the process (heuristic 3).

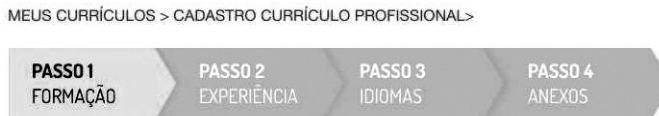


Figure 5. No possibility of jumping through phases

The use of pop-up (figure 6) window to complete the curriculum form was considered a major problem by the experts, because if users have browsers setup to inhibit pop-ups, it would be impossible to finish the procedure. Although the project is focused on desktop access, the system intends to be adaptive. But the pop-up would be impossible to access from smartphones (heuristic 5).

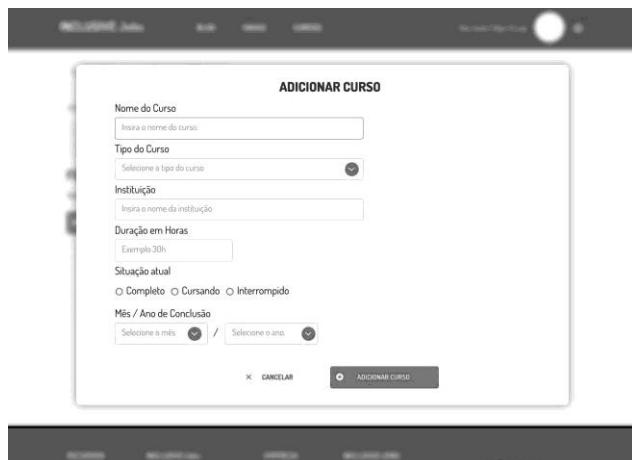


Figure 6. Pop-up window for curriculum completion

The nomenclature for signing in changes throughout the pages: sometimes is written as "login", sometimes is written as "enter" (heuristic 4). Informational banners were considered with low contrast, making legibility difficult (heuristic 8) even for users with normal sight. If considering color blinded users, the problem certainly reaches a higher usability degree.

3.3 Minor Problems

Most of the 35 pointed minor problems are related to heuristics 4, 6 and 8. The use of lack of legibility from small size typography and used in different styles is pointed in a few pages (heuristics 4 and 8). Problems noted with consistency and standards (heuristic 4), regarding the shifting position of the login button and the search engine throughout the pages. Some noted problems were focused on the blog section, regarding navigability to find specific contexts, derived from misunderstanding of icons' function, small size of texts and not enough contrast between figure and background (heuristic 6 and 8).

If users start putting together their curriculum from a job alert, there are many pages to go through and after finishing the curriculum set up, the application to the open position is not sequential. It is not cognitively perceived that users should go back to the alert in order to apply for the job (heuristic 6).

4. CONCLUSION

From the collected results, it became clear that many adjustments must be done for the development of the project with proper usability. The pages, related to profile and curriculum, are the most critical from usability heuristics aspect, and also from users' experience journey expectations. These contain the core information to help impaired people show their skills, their working experience and find proper jobs related to their expertise. The system is expected to bring easy cognitive actions to facilitate users build their profile, but the prototype still doesn't shorten the actions necessary to fulfill the task. An interface with more drag-and-drop interactions could bring a more obvious set of actions for further investigations. Experts suggest the possibility of inserting more than just textual information, such as images for a graphic portfolio.

Since the project is focused on people with impairments, it is important to ensure cognition of information as easily as possible. Some graphic items showed lack of legibility even for the experts, which certainly can be a major issue for people with any minor sight perception. The noted contrast issue brings important discussion for the project regarding its color palette and color perception limitations by color-blind users.

If the project intends to reach a broader audience and include people with vision impairment, the website must follow accessibility interactions set of directives for blind users and consider Paddison and Englefield's heuristics (2003) for accessibility interaction: (1) Provide meaningful and relevant alternatives to non-text elements, (2) Support consistent and correctly tagged navigation, (3) Allow complete and efficient keyboard usage, (4) Respect users' browser settings, (5) Ensure appropriate use of standard and proprietary controls, (6) Do not rely on color alone to code and distinguish, (7) Allow users control of potential distractions, (8) Allow users to understand and control time restraints, and (9) Make certain the Web site is content compatible with assistive technologies.

The use of heuristic evaluation for this phase of the research brought up discussion regarding the method itself and the selection of usability experts. It was observed that designers with a broader experience with the method were more assertive in conducting the analysis. Although 3 participants are experienced designers in interactive projects, it was noticeable the difference between their analysis performance and expert n°2's observations. Expert n° 2 is the only one with post graduation education and with larger experience in conducting digital systems evaluation. Her heuristic report has largely outnumbered the rest of experts and her notes were richer in specifications. But even with such different numbers of problems, the results completed each other, such as Nielsen exposed in his experiment with 19 experts.

After collecting the various noted problems from the experts and prioritizing them by severity, all necessary adjustments are being prepared for a new prototype. Next steps for the project's development seems to orientate this research for a further analysis using Think-aloud Protocol or cooperative analysis methods in order to observe final users' actions and interaction. Observing impaired users and recruiting users interacting with the prototype system will bring data to help compare the project's proposed user experience journey with the users' real time interactive action.

As the construction of an experience journey brings UX design closer to service design (Renzi 2016), it is necessary to go beyond the website and its digital interactions. To be successful and an employment project for social impact, the Inclusive Jobs project has to reach out companies and present the variety of skills and experience from professionals with special needs. To build the bridge, the project must leap forward from digital and take interaction to actors and spaces.

REFERENCES

- Apted, Trent; Collins, Anthony; Kay, Judy. 2009. *Heuristics to Support Design of New Software for Interaction at Tabletops*. Computer-Human Interactions, Boston, USA.
- Inostroza, Rodolfo; Rusu, Cristian; Roncagliolo, Rosana; Rusu, Virginica. 2013. *Usability heuristics for touchscreen based mobile devices*. 2013. 9th International Conference on Information Technology – IEEE. Computer Society. Temuco, Chile.
- Mucchielli, Roger. 1979. *O questionário na pesquisa psicosocial*. Ed. Martins Fontes. São Paulo, Brazil.
- Nielsen, Jakob. 1995. *10 usability heuristics for interface design*. <www.nngroup.com/articles/ten-usability-heuristics>
- Nielsen, Jakob. 1995. *Characteristics of Usability Problems Found by Heuristic Evaluation*. <https://www.nngroup.com/articles/usability-problems-found-by-heuristic-evaluation/>
- Nielsen, Jakob. 1995. *How to Conduct a Heuristic Evaluation*. <https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>
- Paddison, Claire; Englefield, Paul. 2003. Applying heuristics to perform a rigorous accessibility inspection in a comercial contexto. In ACM SIGCAPH Computers abd the Physically Handicapped.
- Portal do governo do Brasil. 2016. Cresce o número de pessoas com deficiência no mercado de trabalho formal. <<http://www.brasil.gov.br/economia-e-emprego/2016/09/cresce-numero-de-pessoas-com-deficiencia-no-mercado-de-trabalho-formal>>
- Portal Inclusivo. 2010. *Cartilha censuo 2010 pessoas com deficiencia*. Impaired people census data: <<http://www.portalinclusivo.ce.gov.br/phocadownload/cartilhasdeficiente/cartilha-censo-2010-pessoas-com-deficiencia.pdf>>
- Portal do Planalto. 1991. *Lei de Cota*. <http://www.planalto.gov.br/ccivil_03/leis/L8213cons.htm>
- Renzi, Adriano Bernardo. 2017. Experiência do usuário: construção da jornada pervasiva em um ecossistema. SPGD2017. Rio de Janeiro, Brazil
- Renzi, Adriano Bernardo; Freitas, Sydney. 2013. *Aplicação de avaliação heurística para diagnosticar problemas de usabilidade na procura e compra de livros em livrarias online*. Revista Ergodesign & HCI. N1. vol.1. year 1. Rio de Janeiro, Brazil. p. 10-21.
- Renzi, Adriano Bernardo, Sande, Axel. Schnaider, Silvia. 2017. Experience, usability and the sense of things. In: Design, User Experience, and Usability: Designing Pleasurable Experiences. HCI International 2017 Proceedings, part II. p.77-86. Vancouver, BC.
- Renzi, Adriano Bernardo. Experiência do usuário: a jornada de Designers nos processos de gestão de suas empresas de pequeno porte utilizando sistema fantasiado em ecossistema de interação cross-channel. Thesis. 239p. Escola Superior de Desenho Industrial. Rio de Janeiro, Brazil, 2016

PLAYER-CENTRIC ADAPTATION OF A CAR DRIVING VIDEO GAME

Boyan Bontchev¹, Dessislava Vassileva² and Dobrin Ivanov²

¹Dep. of Software Technologies, Fac. of Mathematics and Informatics, Sofia University St. Kl. Ohridski,
5 J. Bourchier Blvd., Sofia 1164, Bulgaria

²Scientific Research Center, Sofia University St. Kliment Ohridski,, 8 Dragan Tsankov Blvd., Sofia 1164, Bulgaria

ABSTRACT

Dynamic player-centric adaptation continues being a challenge for software applications and, in particular, in video games. The paper presents a dynamic adjustment of difficulty and visual effects in a car driving 3D video game based on changes of both the player performance (outcomes) shown within the game session and, on the other hand, the player arousal (excitement) recognized by a special, free game component. The component is applied for estimation of arousal inferred by the electrodermal activity of an individual player measured by a cheap custom device. The paper proposes a simple model for keeping the player in flow while adjusting dynamically both task difficulty and game visual effects, providing a better immersion, challenge, and tension. The presented initial results are statistically significant and answer positively to the main research question of the study: does the adaptive version of the game provide a better gaming experience than the non-adaptive one?

KEYWORDS

Video Game, Player-Centric, Adaptation, Arousal, RAGE

1. INTRODUCTION

In recent years, the game industry has been one of the fastest growing in the area of entertainment. Video games are engaging, amusing and exciting (Egenfeldt-Nielsen et al., 2015). They are everywhere around us and their popularity grows every day. Video game applications go beyond entertainment and they are used successfully also in the areas of psychology, education, neuroscience, medicine, etc. (Murphy et al., 2013).

Due to the great competition among game developers, they are increasingly challenged to make their games attractive, engaging, and useful, by improving the user experience of players. According to Csikszentmihalyi (1997) games are successful if they invoke positive emotions and immersion. The audience of a game should be as big as possible and it has to cover different type of players. Usually, keeping the attention of different type of players is achieved by adjusting/changing of the level of difficulty, game mechanics, game story or speed depending on shown player's results (Bontchev, 2016) or player's biofeedback (Bersak et al., 2001). Adapting a game depending on physiological data of players is a relatively new approach used in affective video games. Affective video games use physiological signals in order to recognize emotional excitement of players and to adapt the gameplay based on it (Fairclough, 2009; Kivikangas et al., 2011). The adaptation mechanisms are based either on player's performance, or on player's physiological signals, but not on both of them. Therefore, new adaptive models combining player's performance with emotional status recognized by means of player's physiological data may provide a better gaming experience including flow, immersion, tension, challenges, and competence and, hence, an enhanced overall playability.

The paper presents a study of a dynamic player-centric adaptation of both task difficulty and visual effects in a car driving video game. In contrast to other affective games (Bontchev, 2016), here we apply a new way of dynamic adjustment of difficulty and visual effects, which is based on changes of both player outcomes shown within the game session and player arousal. The last is inferred by a special, free game component named Real-Time Excitement Recognizer and developed in the scope of the RAGE (Realising an Applied Gaming Eco-System) H2020 research project. The component is applied for estimation of arousal

(excitement) inferred by the electrodermal activity of an individual player measured by a cheap custom device. The paper proposes a simple model for keeping the player in flow (Csikszentmihalyi, 1997) while adjusting dynamically both task difficulty and game visual effects, on the base of player performance and achieved general arousal. It explains how the model is applied to the car driving affective video game and, as well, the experimental setup and the procedure of playing the game and filling a post-game questionnaire about the gaming experience. Next, it presents some initial results answering positively to the main research question of the study: *does the adaptive version of the game provide a better gaming experience than the non-adaptive one?* The final section of the paper provides some conclusions about the experimental results and directions for future works.

2. BACKGROUND

The section briefly outlines the methods of using player-centric adaptation in affective video games and presents the previous works of the authors applying a software component for real-time excitement recognition.

2.1 Player-centric Adaptation in Affective Video Games

Physiological computing systems capture various types of responses of the central and peripheral human nervous system and apply them for dynamic adaptation of the software application (Fairclough, 2009). The term *affective feedback* was coined by Bersak et al. (2001) by imposing the idea for “*an active intelligent participant in the biofeedback loop*”. The affective (or biocybernetic) loop applies physiological adaptation mechanisms in order to “*augment human-computer interaction by interpreting human behaviour via physiological responses*” (Muñoz et al., 2017). The loop provides an adaptation of software systems or hardware devices by means of a negative or positive feedback control. When applied in affective video games, it keeps player’s arousal or emotions near the optimal, safe level of performance effectiveness and, thus, leads to a stable gameplay (Fairclough, 2009). On the other hand, positive physiological feedback control enlarges the difference between the measured psychophysiological input and the desired player performance, i.e. the desired anxiety level (Rani et al., 2005) or the level of player’s emotion (Lindley & Sennersten, 2006) is steadily increased for reaching greater performance and pleasure. Physiological responses were applied to adapt dynamically the game for maintaining the player in an optimal status of enjoyment (Tognetti et al., 2010).

Affective video games apply measurements of physiological or neural signals for recognizing player’s emotional states and, hence, for adapting the gameplay in of more exciting ways (Kivikangas et al., 2011). They apply the affective loop for implementation of three main types of gameplay adaptation, as summarized by Novak et al. (2012):

- (1) adaptive automation of tasks with possible provision of automated assistance – according to Murphy et al. (2013), game tasks can be: explicit - e.g., objectives, goals, and missions posed to the player; implicit - not explicitly stated but expected to be fulfilled; player-driven tasks - created by the player;
- (2) adjustment of game difficulty level (referred as dynamic difficulty adjustment, or DDA) for appropriate balancing challenge with player skills in three ways (Chang, 2013): by means of automatic level generation, modification of artificial intelligence, or adjusting level content;
- (3) adaptation of audio-visual effects such as ambient light (Bontchev & Vassileva, 2017).

In last two decades, various dynamic models of player behavior were proposed using tracking and monitoring of player interactions either for recognition of player’s emotions or for inducing desired emotions during gameplay. A review on adaptation models, mechanisms and practices used in affective video games including ways of their application, types of measured signals and their features, details about biofeedback devices, emotions/arousal recognized through classification or estimation, experimental validation using questionnaires or interviews, and assessment of outcomes and impact of the affect-based adaptation in 14 video games can be found in (Bontchev, 2016).

2.2 The RAGE Real-Time Excitement Recognizer

Within the scope of the RAGE project, the authors designed and developed a software component for real-time arousal (i.e., excitement) detection based on measuring electro-dermal activity (EDA), also known as skin conductance response (SCR), and skin conductance level (SCL) (Kivikangas et al., 2011). EDA is related to the activity of the sweat glands, which are regulated by the sympathetic nervous system. It is measured by a particular player in order to produce some real-time signal features. The level of arousal may be useful for emotion detection and for adaptation purposes.

The software component receives a filtered and coded signal from an advanced, custom built, compact biofeedback device, developed in the scope of the RAGE project and named Galvanic Skin Response Meter. It is powered by a risk-based microcontroller system and allows measurements with sampling rate set by control system from 0.4Hz up to 800Hz. The applied output voltage is less than 3.9V according to human electrical safety (Dorr, D., 2009). The range of measurements is from 0.488 μ S up to 35 μ S. Its precision class is 1%. For client convenience, the Galvanic Skin Meter is powered by the same data connection through the USB port, and the consumption power is less than 500mW. The device has a very compact size (90x50x15mm) and a weight only 60g (figure 1). Human galvanic measurements are carried out with two electrodes placed on two adjacent fingers.



Figure 1. A view of the Galvanic Skin Meter

Recording, filtering and feature extraction might be executed on a computer (server) different from the game machine, in order to speed up all the required processing. The results are communicated from the server-side to the client component in order to be used for game adaptation. A calibration period is used for measuring the SCL baseline, which will be subtracted from the SCL level for normalizing it. The baseline determined during the calibration period implying total rest of the user sets the zero tonic arousal level. In the same way, the zero phasic arousal level is set by the SCR features measured during the calibration period. After the calibration period (set by the developer), the component produces several metrics featuring user arousal based on the EDA signal, namely:

- Current level of phasic user arousal measured in N levels, i.e. from 0 to N-1, where N is user-defined and has a default value (the level is indicated by the area under the curve of the filtered EDA signal within the time window using a sampling rate chosen by the user of the component). The user arousal level is determined after the first time window is expired and is based on statistical analysis of human EDA signals, thus, the accuracy is based on statistically significant data. With collecting more data from given user, the component starts providing more accurate data;
- Current level of tonic arousal measured in N levels, i.e. from 0 to N-1, where N is user-defined and has a default value (the level is indicated by the mean amplitude of the tonic component of the signal within the time window and using a sampling rate chosen by the component user);
- Current level of general arousal measured in N levels, together with real-time features of EDA signal measured from particular player such as: phasic activity represented by mean and maximum amplitude of skin conductance response (all in micro-siemens), rate of phasic activity (response peaks/sec), SCR rise time, SCR 1/2 recovery time, mean tonic activity level, and slope of tonic activity (in micro-siemens/sec).

The Real-Time Excitement Recognizer game component has been used in the scope of the ADAPTIVES project. The component was applied for estimation of excitement (arousal) inferred by measuring electrodermal activity of an individual player and used together with emotions recognized through facial expression analysis, for a dynamic player-centric adaptation of both game difficulty and some visual effects (Bontchev & Vassileva, 2016) and for recognition of playing styles (Bontchev & Georgieva, 2018), within an action-adventure video game. The conducted experiments found statistically significant correlations between emotional arousal states from one side and from another side game engagement, current player emotions, and gameplay metrics. As well, affectively adaptive gameplay was proven to bring better distribution normality and to increase task's efficiency and difficulty with a statistically significant effect and, on the other hand, to reduce game session time and relative time for performing a single task (Bontchev & Vassileva, 2017).

3. EXPERIMENTAL SETUP

This section explains the experimental setup including a car driving video game applying the RAGE real-time excitement recognizer, the adaptation methods based on player excitement and outcomes, and the experimental procedure.

3.1 The Car Driving Video Game

Tognetti et al. (2010) applied physiological responses to adapt dynamically a car racing game for maintaining the player in an optimal status of enjoyment. In contrast to that research, we applied the RAGE real-time excitement recognizer for adapting a cost-free car racing video game for analyzing and testing benefits of player-centric adaptation based on both player arousal and game outcomes. The goal of this game is from one hand to be used as a car simulator and from another hand to be used for an improvement of the driving skills of players in a more effective, easier and faster way. Therefore, both the car and the environment of our game have to be as possible as realistic and to represent an adequate representation of real-world scenarios (figure 1). Moreover, it has to be able to detect progress and regression of driving skills of players and to take appropriate action in order to help them improve their achievements (related to learning/improving driving skills).

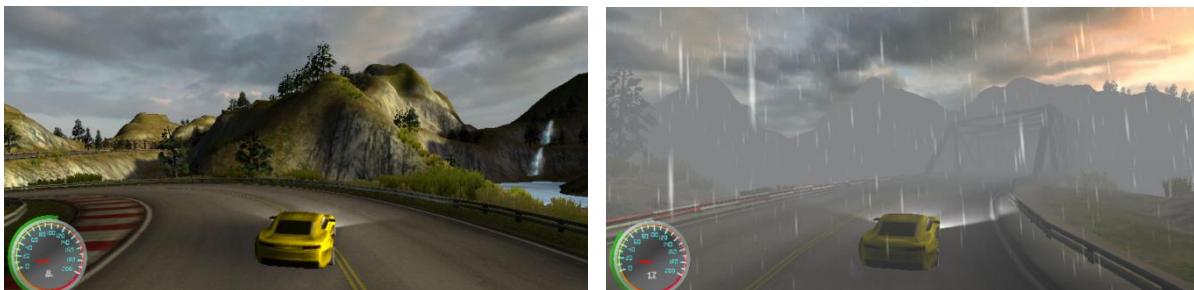


Figure 2. Views of the game prototype – driving in good weather (left) and in a foggy, rainy day (right)

For these reasons, the game was developed by the authors of the paper on the base of an existing video game of car simulator available at GitHub¹. The existing game was re-designed by adding to it necessary gauging tools, gadgets, and as well as, availability of different weather conditions, and implementation of adaptability based on player performance and arousal. Hereby, it provides a fascinating first-person car driving simulation with maneuvering on a road with various types of curves, through attracting and realistic terrains. It increases player motivation and fun by visualization of extra tools like a speedometer and trip counter, by playing amazing sounds (of the motor, of each collision, etc.), and by driving under different weather conditions such as fog, rain, and nightfall.

The learning goal applied in the game is players to develop best skills for driving the car achieving at the same time a maximal velocity and a minimum number of collisions at different road scenarios and weather

¹ <https://github.com/rostrage/CarSimulator>

conditions. Under collision, it is understood every exit from the road and hitting on the fence guard or on another object. Each collision is characterized by an impact measured on a basis of the current velocity and the angle between the car and the hit object. The higher speed and greater angle lead to greater impact that simulates the degree of the damage in a real-life situation.

The game registers a metric for the overall player performance (OPP) and logs its value at every second. It evaluates this metric depending on the current impact of collisions and achieving average speed, according to the following formula:

$$(1) \ OPP = V_{avg} / (I_{normalized} + 1),$$

where V_{avg} is the average velocity for a time window of 10 seconds and $I_{normalized}$ is the average normalized impact of collisions calculated following the formula:

$$(2) \ I_{normalized} = \text{SUM}(V * |\sin(\alpha)|) / N_{coll} / I_{avg},$$

where V is the car velocity at the moment of collision, α is the angle between the car direction and the hit body, N_{coll} is the number of collisions for the time window, and I_{avg} is the average impact value for the game.

Together with the player's excitement (arousal) inferred by the RAGE component, OPP is applied for controlling the adaptation process.

3.2 Adaptation Methods based on both the Player Excitement and Outcome

With respect of three possible value levels of individual outcome (i.e., OPP) and excitement (i.e., arousal) – low (from 0% up to a threshold of 33% of the maximal value), medium (with thresholds of 34% - 67%), and high (with thresholds of 68% - 100%) – we have six possible cases:

1. Medium or high outcome and high excitement – at such a moment, the player does the task very well, therefore, nothing should be changed;
2. High outcome and low/medium excitement – a situation demonstrating the gameplay has become boring to the player – he/she is fulfilling very well the task but without fun and engagement. Therefore, the task difficulty should increase;
3. Low outcome and medium/high excitement – shows that the player cannot meet the game requirements and, most probably, is being under stress. The arousal provides no emotional valence (Bontchev, 2016), however, the low outcome proves the impossible way to meet the game challenge, thus, the task difficulty should decrease;
4. Low outcome and low excitement – this could happen when the player is at the beginning of the task and yet will start playing seriously, or after a moment of collision. Therefore, task difficulty should be kept low;
5. Medium outcome and low excitement – reveals the player's performance goes higher than the excitement; therefore, the task difficulty should increase with Δd up to D_{max} ;
6. Medium outcome and medium excitement – the situation is typical for players gathering more skills in task execution. Nothing should be changed.

Table 1. Dynamic adaptation control of task difficulty according to player outcome and arousal

Outcome	Arousal	Difficulty
Medium/High	High	Not changed
High	Low/Medium	Increased with Δd up to D_{max}
Low	Medium/High	Decreased with Δd up to D_{min}
Low	Low	Decreased with Δd up to D_{min}
Medium	Low	Increased with Δd up to D_{max}
Medium	Low/Medium	Not changed

Table 1 represents the dynamic dependency of the task difficulty on both the player outcome and arousal. The game difficulty has a maximal value denoted D_{max} and a minimum given as D_{min} . Each adaptable game level starts with task difficulty equal to D_{min} for the level. The task difficulty cannot be greater than D_{max} and depends on three adjustable game mechanics:

1. fog density – at driving in fog;
2. road slipperiness – at driving in rain;
3. light intensity – at driving at nightfall.

In this way, current task difficulty can be changed by a constant amount of $\pm\Delta d$ until reaching D_{\max} or D_{\min} . The change $\pm\Delta d$ is implemented by a linear increment/decrement of the current values of the three game mechanics given over.

During the initial experiments, we found the task difficulty may fluctuate (i.e., may change several times up and down) when the conditions alternatively change. In order to avoid such fluctuations, we applied a hysteresis for the threshold-based control applying $\pm10\%$ of the thresholds of 33% (low-middle) and 67% (middle-high).

3.3 Experimental Procedure

For conducting the practical experiment, the game was built for the Windows platform as a desktop executable connected to the RAGE component through a simple socket interface and including all the drivers and libraries needed for the EDA measuring device. The game included three driving tours on the same road (through a generated terrain):

- Level 1: a non-adaptive tour (i.e., without any player-centric adaptivity), with best driving conditions – applied for calibration of the RAGE component thanks to the fact the player starts with the game and might demonstrate various levels of arousal;
- Level 2: an adaptive tour through a foggy weather – the dynamic difficulty of driving is adjusted according to player outcome and arousal, by changing the fog density and the illumination at nightfall;
- Level 3: an adaptive tour through a foggy and rainy weather – the dynamic difficulty of driving is adjusted according to player outcome and arousal, by changing the fog density and the road friction in the rain.

For the practical experiment, we used a gamer laptop gamer laptop with 15" LCD using a Logitech wireless mouse and the Galvanic Skin Meter device. The electrodes were placed at the palmar and hypotenar eminences of the middle and ring fingers of the player's non-dominant hand. To avoid pressure artifacts, participants were instructed not to use these fingers while playing. The participation was entirely anonymous and voluntary. After an individual explanation about the experiment, short demonstration and a procedure of informed consent, each participant was asked to spend 2 minutes in relaxation, with the electrodes placed on his/her fingers, while listening to calm music and watching playing instructions. Next, he/she played one game session through the three levels of the game, unrestricted in time, with logging the playing time, outcome and arousal once per second. The calibration period for the EDA measuring device was finished at the end of the first (non-adaptable) level of the game. Each higher game level proposed a higher difficulty compared to the previous one. Only within the first level, the game difficulty was not adjustable and depended only on the curves of the road. The difficulty of the next levels was changing on outcome and arousal of the particular player. After the game session, the player was asked to fulfil the reduced Game Experience Questionnaire (GEQ) questionnaire (Nacke, 2009) asking 36 questions (translated in Bulgarian) about competence, immersion, flow, tension, challenge, and positive/negative affect felt by the player in both the non-adaptable and adaptable version of the car driving game.

4. EXPERIMENTAL RESULTS

The section outlines some of the initial results of a case study exploring the dynamic player-centric adaptation of both task difficulty and visual effects in a car driving video game, based on both player's performance (outcomes) and player's arousal inferred on the electrodermal activity of the skin. Nine volunteers (5 men and 4 women, at average aged 29.22 years) participated the experiment following the procedure described over.

Figure 3 presents charts of dynamic changes of player's OPP, player's arousal and game difficulty trend for the adaptive game version, i.e. for the second and the third driving tour. The positive difficulty trend means an adaptation control striving to increase the difficulty up to D_{\max} , while the negative difficulty trend stands for and decreasing the difficulty up to the minimum D_{\min} . After reaching the minimum/maximum value of difficulty, the adaptation control stops decreasing/increasing the difficulty according to the Table 1 in order to provide a better flow and immersion of the gameplay.

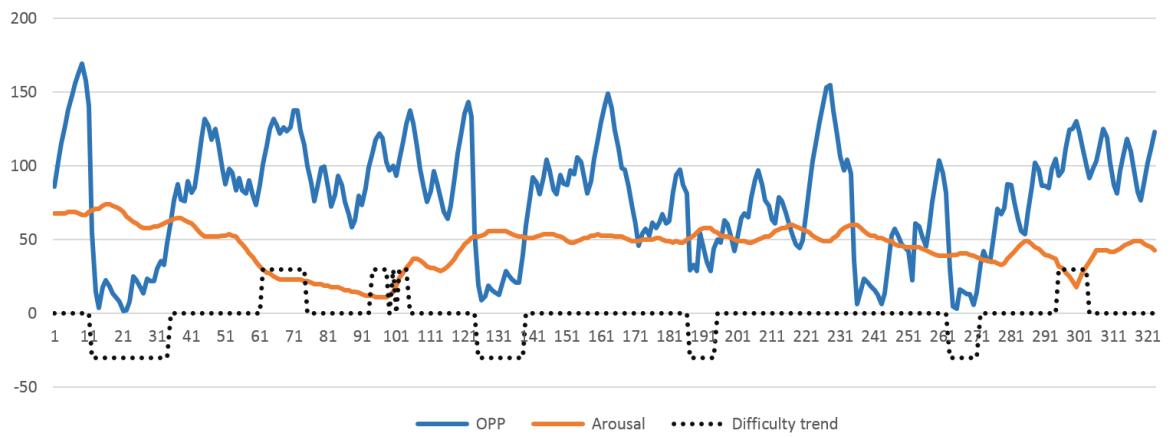


Figure 3. Charts of OPP, arousal and difficulty trend for the adaptive game version

After each game session, the participant was asked to answer the 36 questions the GEQ questionnaire translated into Bulgarian. Table 2 presents the results of the seven GEQ factors for the non-adaptive game and the adaptive game, namely mean (M), standard deviation (SD), and standard error (SE). The last column provides the statistical significance of the differences of the means of results for the non-adaptive and adaptive game obtained by a paired two-sided T-test. The results prove that even with a rather small number of participants ($N=9$) the differences of the means for immersion, flow, competence, tension, and challenge are statistically significant at values of $p<0.01$, and for the negative affect at $p<0.001$. The differences clearly show that the adaptive version of the game provides a better immersion, flow, competence, tension, and challenge than the non-adaptive one. At the same time, in the non-adaptive version of the game the player feels more competent, which can be explained by the lower challenge of that version. As well, it appears somehow boring for the player bringing a higher negative affect than the adaptive one.

Table 2. Results about the GEQ factors for the non-adaptive game and the adaptive game

GEQ factor	Non-adaptive game			Adaptive game			Significance
	M	SD	SE	M	SD	SE	
Immersion	3.963	0.450	0.159	4.481	0.426	0.151	0.003306*
Flow	3.089	0.880	0.311	3.978	0.769	0.272	0.002001*
Competence	4.156	0.759	0.268	3.089	1.178	0.416	0.007449*
Tension	1.489	0.390	0.138	2.689	0.971	0.343	0.005930*
Challenge	3.133	0.542	0.191	4.178	0.607	0.215	0.004397*
Positive affect	3.889	0.213	0.075	3.533	0.899	0.318	0.269346
Negative affect	1.844	0.409	0.144	1.333	0.313	0.111	0.000953**

* $p<0.01$
** $p<0.001$

5. CONCLUSION

The paper presented an experimental study of a dynamic player-centric adaptation of both task difficulty and visual effects in a car driving video game, where the dynamic adjustment of difficulty and visual effects is based on changes of both player performance shown within the game session and the player arousal. The initial results are statistically significant and answered positively to the main research question of the study - the adaptive version of the game provides a better gaming experience than the non-adaptive one.

The player arousal was inferred by a component named Real-Time Excitement Recognizer developed in the scope of the RAGE H2020 research project. The practical experiments proved that the effectiveness of the component, when applied for adaptive 3D video games, is high. Game developers can use arousal detection functionality in their games to support an adaptive gameplay and, thus, to enhance game playability. Thanks to its simple but effective API, the output from the component can be easily integrated into any digital game. On other hand, education specialists can track the learners' arousal in order to assess

and to optimize learning experiences in a game. Thus, the component was proven to be easily applied in more complex online 3D entertainment video games in order to achieve an efficient player-centric adaptation similar to the experiments done in the scope of the ADAPTIVES project (Bontchev & Vassileva, 2017).

Future experiments with the component will include a greater number of participants, different types of games (e.g., a box shooter game), usage of both tonic and phasic arousal of the individual player, and a bigger questionnaire extended with some questions asking about game playability issues. In such a way, we will be able to generalize the conclusions about the advantages and merits of player-centric affective video games.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the EC H2020 project RAGE (Realising an Applied Gaming Eco-System), Grant Agreement No. 644187, and from the APOGEE project, funded by the Bulgarian National Science Fund, Grant Agreement No. DN12/7/2017.

REFERENCES

- Bersak, D., McDarby, G., Augenblick, N., McDarby, P., McDonnell, D., McDonald, B., and Karkun, R., 2001. Intelligent biofeedback using an immersive competitive environment. *Designing Ubiquitous Computing Games Workshop*, UbiComp, Atlanta, USA.
- Bontchev, B., 2016. Adaptation in Affective Video Games: A Literature Review. *Cybernetics and Information Technologies*, Vol. 16, No 3, pp.3-34, DOI: 10.1515/cait-2016-0032.
- Bontchev, B., Vassileva, D., 2016. Assessing engagement in an emotionally-adaptive applied game. *Proc. of the 4th Int. Conf. on Technological Ecosystems for Enhancing Multiculturality (TEEM'16)*, ACM, November, pp.747-754.
- Bontchev, B., Vassileva, D., 2017. Affect-based adaptation of an applied video game for educational purposes. *Interactive Technology and Smart Education*, Vol. 14, Issue 1, January, pp.31-49, <https://doi.org/10.1108/ITSE-07-2016-0023>.
- Bontchev, B., Georgieva, O., 2018. Playing style recognition through an adaptive video game. *Computers in Human Behavior*, 82, pp. 136-147, <https://doi.org/10.1016/j.chb.2017.12.040>.
- Chang, D. M. J., 2013. Dynamic Difficulty Adjustment in Computer Games. *Proc. of 11th Ann. Interact. Multimedia Systems Conf.*, Univ. of Southampton, UK.
- Csikszentmihalyi, M., 1997. *Finding Flow: The Psychology of Engagement with Everyday Life*, New York, Basic Books.
- Dorr, D., 2009, July. Determining voltage levels of concern for human and animal response to ac current. *Power & Energy Society General Meeting (PES'09)*, IEEE, pp. 1-6.
- Egenfeldt-Nielsen, S., Smith, J.H., Tosca, and S.P., 2015. *Understanding video games: The essential introduction*, Routledge.
- Fairclough, S. H., 2009. Fundamentals of Physiological Computing. *Interacting with Computers*, (21)1, pp.133-145.
- Kivikangas, J. M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Järvelä, S., and Ravaja, N., 2011. A review of the use of psychophysiological methods in game research. *Journal of Gaming & Virtual Worlds*, 3(3), pp.181-199.
- Lindley, C. A., Sennersten, C. C., 2006. Game play schemas: from player analysis to adaptive game mechanics. *Proc. of the Int. Conf. on Game research and development*, pp.47-53.
- Muñoz, J.E., Rubio, E., Cameirao, M., and Bermúdez, S., 2017. The Biocybernetic Loop Engine: an Integrated Tool for Creating Physiologically Adaptive Videogames. *4th Int. Conf. in Physiological Comp. Systems (PhyCS17)*. INSTICC.
- Murphy, C., Chertoff, D., Guerrero, M., and Moffitt, K., 2013. Design Better Games: Flow, Motivation, and Fun. *Design and Development of Training Games: Practical Guidelines from a Multidisciplinary Perspective*, p.1773.
- Nacke, L. 2009. *Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment*, Doctoral Dissertation, Blekinge Institute of Technology.
- Novak, D., Mihelj, M., and Munih, M., 2012. A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interacting With Computers*, 24, pp.153-172.
- Rani, P., Sarkar, N. and Liu, C., 2005, July. Maintaining optimal challenge in computer games through real-time physiological feedback. *Proc. of the 11th Int. Conf. on Human Computer Interaction*, Vol. 58, pp. 22-27.
- Tognetti, S., Garbarino, M., Bonarini, A. and Matteucci, M., 2010, August. Modeling enjoyment preference from physiological responses in a car racing game. *Computational Intelligence and Games (CIG)*, IEEE, pp. 321-328.

CLASSIFICATION OF DRIVING TRAITS USING FUZZY-LOGIC CONTROL

Dylan Seychell and Steven Farrugia

Saint Martin's Institute of Higher Education, Malta

ABSTRACT

Younger as well as less experienced drivers have largely contributed to the alarmingly high vehicle collision rate globally. Improving smartphone and onboard technology can help to profile drivers and subsequently detect abnormal driving traits that can be potentially hazardous while providing an opportunity to reward good driving. This paper explores the relationship between driving traits and collision rates among different age groups and driving experience. This paper presents a solution that uses a smartphone application to classify drivers by using sensor data and feeding it into a fuzzy inference system, mapping acceleration and speed to the corresponding driving classification output through defined fuzzy logic rules. A survey with 806 respondents provided insights about drivers' attitude towards such profiling with an encouraging 65% stating they are willing to provide their data in return for driving reports. The fuzzy inference system was then evaluated with 20 drivers of different age and experience. They were asked to follow a designated route while having the mobile application recording the journey. The system successfully profiled the drivers in relation to their driving record with the drivers classified as most aggressive by the system being the same drivers with most collisions in their driving history.

KEYWORDS

Driver classification, Driving Behavior, Fuzzy logic inference

1. INTRODUCTION

During a three-year period, starting from 2012 to 2015, the number of vehicles present on the world's roads have seen an increase of 16%, with a staggering 67-million passenger vehicles coming into circulation during 2014 alone (WHO, 2015). This has directly impacted the number of road collisions per year. A number of strategies have been implemented to tackle this issue. In 2012, research carried out by the World Health Organization revealed road traffic crashes were the leading cause of death for people aged between 15 and 29. To address this global issue, the UN General Assembly announced the Decade of Action for Road Safety, with the intent to decrease the overall predicted amount of fatalities (WHO, 2015). Although multiple methods may be implemented (improving road safety standards, better education to the driver etc.) in order to address this worrying issue, *smartphone technology* can also be used to mitigate and prevent road traffic collisions.

While most in-vehicle data recorders (IVDR) have focused on the collision event, the data gathered can be used to tackle other important areas such as training and collision avoidance. Through smartphone hardware technology, inbuilt sensors may capture numerous quantitative values which may collectively provide a descriptive representation of the drivers' actions. Although the sensors available may not be as accurate or powerful as specialized apparatus, they are considered to be a viable option when carrying out a behavior analysis in a practical driving setting (Antoniou *et al.*, 2014).

This paper presents a profiling approach that uses fuzzy logic inference to classify drivers using a mobile device. The implementation of a mobile application that enables data collection and feedback together with a complementary server-side application is presented. The subsequent evaluation through a survey with 806 respondents and an experiment with 20 drivers is presented in this paper. The results show that drivers are willing to use such a system provided that they receive feedback and are able to improve their driving through the use of mobile technology.

2. BACKGROUND

This section provides background about techniques that allow for driver profiling. The first subsection provides information about profiling and classification techniques, focusing on those that make use of mobile technology. It is then followed with another subsection that outlines how fuzzy logic inference can enable such profiling.

2.1 Driver Profiling and Classification

The potential behind providing feedback to drivers on their driving style, specifically those who possess an aggressive driving style, is very promising. Informing such drivers would help them become aware of and also contribute towards a better driving behavior. Antoniou *et al* proposed a driving behavior classification scheme based on acceleration measurements gathered from smartphone sensors. The k-means clustering algorithm was applied to the data generated by two drivers using different vehicles. From a preliminary analysis of the gathered information, driving behaviors were evidently different on the longitudinal axis, representing the acceleration and deceleration of the vehicle. Although the exact same traffic situations were not replicated, the mapped clusters of data were able to provide a classification of two drivers having different driving behaviors (Antoniou *et al*, 2014).

Numerous applications have obtained valuable information from smartphone sensors in order to assess ones driving behavior as well as provide feedback to the driver. Information gathered from the multiple sensors embedded in the smartphone could act as a foundation in providing safety features to drivers which have been identified as aggressive or high-risk. Smartphone sensors can be grouped into three categories; motion, environmental and position sensors (Magrin *et al*, 2015). A study carried out by Johnson and Trivedi classified aggressive driving behavior through the development of a smartphone application which did not require any external processing or components. The drivers' style, based on characteristics, was classified into two groups; aggressive and non-aggressive. Once the smartphone sensors detected aggressive vehicle maneuvers, including excessive speed, hard breaking and abrupt turns, audible feedback informs the driver while also initiating video recording, tracking the drivers' location, storing the vehicle speed and documenting any recent maneuvers (Johnson & Trivedi, 2011). From all the data gathered, the results proved that smartphones sensors were able to detect movements at a quality level similar to that obtained by a vehicle Controller-Area-Network (CAN) bus. In 2012, an application was developed, using the smartphones hardware sensors and internal IMU, to calculate and compare critical driving events experienced by a driver, with data generated by the vehicles installed Inertial Measurement Unit (IMU) (Paefgen *et al*, 2012). This app utilizes the gyroscope, accelerometer and GPS sensors. Data collected from the sensors serve as an input to two functional elements; calibration and journey recording. The calibration process defined the devices three-dimensional position in the vehicle and once complete, the parameters are stored. Following calibration, journey recording could commence at the drivers' discretion. Throughout the trip, three different data types were collected; time at which the journey commenced and finished, the trip duration time, GPS readings at regular intervals at which velocity and covered distance were measured, and acceleration and gyroscope readings. An event was triggered once sensor readings exceeded threshold acceleration values, which were defined for maneuvers normally executed during a driving trip. These included braking and turns. Once an event was triggered, the device's GPS location along with a timestamp were recorded, allowing users to view where driving violations occurred after their journey. On analyzing the data gathered, a large quantity of smartphone generated events correlated to those produced by the vehicle IMU, indicating the smartphone is to a certain extent a means for reliable event detection.

2.2 Fuzzy Logic Inference

Dynamic real-world unconstructed environments generate a number of ambiguous sources (Hagras, 2007). These uncertainties have a direct impact on the membership function. There are a number of different membership functions, namely; the triangular, trapezoidal, Gaussian, bell-shaped, sigmoidal and S-Curve waveforms. Each function has a suggested use. In situations where a system will carry out large dynamic discrepancies in a small-time frame, one should make use of the triangular or trapezoidal waveform. Gaussian or S-Curve waveforms should be utilized when a high level of control accuracy is required

(Bai & Wang, 2006). An approach put forward in 2010, to further improve the recognition of specific driving maneuvers for Advanced Driver Assistance Systems (ADAS), combined the use of a fuzzy rule base with that of a probabilistic finite-state machine (Hulnhagen *et al*, 2010).

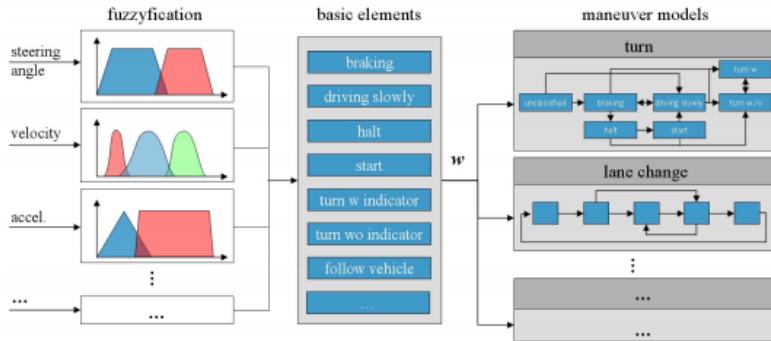


Figure 1. Sequence of basic elements forming a maneuver (Hulnhagen *et al*, 2010)

The approach described, breaks down a maneuver into a sequence of basic elements which directly correlate to a set of input values obtained from the test vehicle CAN bus and installed GPS. Using membership functions, input variables (such as acceleration) were mapped to a normalized output set [0, 1] with a number of linguistic terms. The relation to a specific linguistic variable was determined by a defined membership function. Combining these variables logically using the MIN and MAX operators to represent any ‘and’ and ‘or’ logical operators, each fundamental element of a maneuver was expressed by a fuzzy rule (implemented using Fuzzy Control Language (FCL)).

A similar approach was put forward in 2012 (Aljaafreh *et al*, 2012). Using a 2-axis accelerometer, any g-force experienced in both longitudinal and horizontal acceleration was measured. A driver executed a ‘start, drive and stop’ driving procedure in a normal and aggressive manner, which was then represented graphically. The resultant value combined with the vehicles velocity, allowed for the classification of 4 distinct driving styles; below normal, normal, aggressive, and very aggressive. The fuzzy inference system used the average of the Euclidean Norm of both the longitudinal and lateral acceleration, combined with the vehicle velocity. All membership functions, defining the relationship between the input and output value, were derived from expert driver knowledge. Fuzzy values for the accelerative inputs, were specified as Low, Medium and High, whereas Very Low, Low, Medium, High and Very high corresponded to the speed fuzzy values. The overall output represented the driving behavior, simplified into 4 fuzzy values; below normal, normal, aggressive and very aggressive. Based on feedback provided by three expert drivers and analysis of real test data, manual adjustments were applied to the membership functions as well as the fuzzy rules.

3. METHODOLOGY

3.1 Smartphone Application

The proof of concept application was developed on Android. Sensor data values are displayed once the application is opened, grouped by the respective sensor type. These include; latitude and longitude values, speed in km/hr (calculated using location values), acceleration in the x, y and z-axis, gyroscope values, timestamp once a new location is received and the difference in seconds between the current and last location update. Assuming location services were switched on, the current location is updated on the Google Maps interface. Any changes in location are depicted using a blue line, indicating the user’s path. Using fine location updates, produced relatively accurate results, which actively contributed towards a +/- 2km/hr discrepancy when visually compared to the car’s speedometer value.

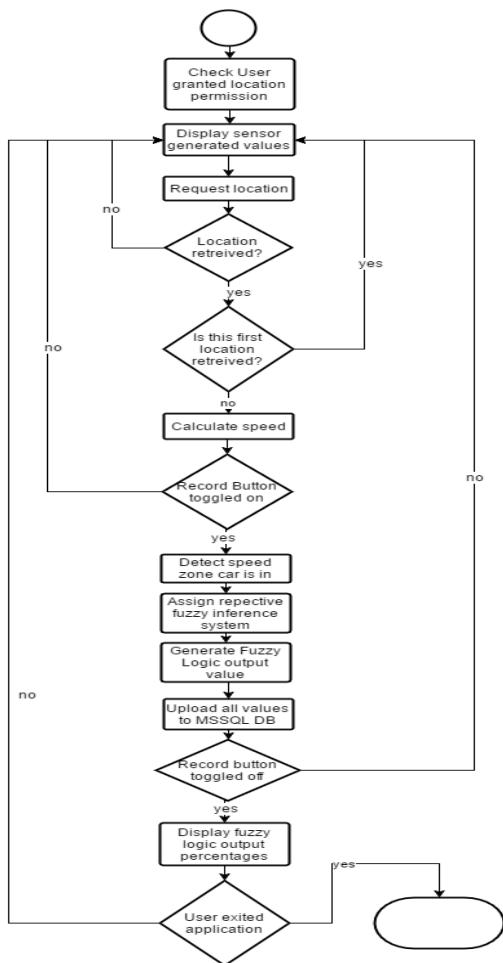


Figure 2. Flow Diagram of application

3.2 Server-side Profiler

At server-side, the fuzzy logic inference system was deployed and a secure web interface was developed to manage the data being collected and processed. In terms of GUI, the following inputs were collected: z-axis acceleration and speed terms along with their respective ranges, minimum and maximum values. Input term values were based on sensor values generated while executing different driving styles (i.e. driving in aggressive, calm and normal behavior) which I had conducted along the defined route. Therefore, input ranges were defined based on my judgment of what is deemed to be aggressive, calm etc. For speed, three terms were defined; Fast, Normal and Slow. With respect to acceleration, five terms were specified; Hard Deceleration, Deceleration, Neutral, Acceleration and Hard Acceleration.

Once the input variables had been defined, three different driving classification outputs were defined. These were ‘Calm’, ‘Aggressive’ and ‘Normal’. Having 3 terms defined for speed input variable and 5 for acceleration required 15 fuzzy logic rules to be specified so as to produce the corresponding fuzzy logic output value. Multiple changes were made in order to refine the Fuzzy Inference System (FIS). Two fuzzy logic rules were specified, corresponding to the 50 and 70 speed limit zones present along the predefined route, as may be seen in tables 2-4 below.

Table 1. Speed input variable ranges and corresponding output terms (speed limit 50/70 km/hr.)

Speed (Km/hr.) 50 Zone	Speed (Km/hr.) 70 Zone	Output Term
0 - 35	0 - 50	Slow
30 - 55	45 - 75	Normal
50 - 100	70 - 100	Fast

Table 2. Acceleration input variable ranges and corresponding output terms

Acceleration/Deceleration Z-axis (50/70 km/hr. zone)	Output Term for corresponding acceleration range
1.5 – 4.5	Hard Deceleration
0.5 – 2.5	Deceleration
-0.5 – 0.75	Neutral
-2 – (-0.25)	Acceleration
-4.5 – (-1.5)	Hard Acceleration

Table 3. Fuzzy logic rules output

Speed	Acceleration/Deceleration	Fuzzy Logic output
Slow	Hard Deceleration	Aggressive
Slow	Deceleration	Calm
Slow	Neutral	Calm
Slow	Acceleration	Calm
Slow	Hard Acceleration	Aggressive
Normal	Hard Deceleration	Aggressive
Normal	Deceleration	Normal
Normal	Neutral	Calm
Normal	Acceleration	Normal
Normal	Hard Acceleration	Aggressive
Fast	Hard Deceleration	Aggressive
Fast	Deceleration	Aggressive
Fast	Neutral	Aggressive
Fast	Acceleration	Aggressive
Fast	Hard Acceleration	Aggressive

4. EVALUATION

This solution was evaluated through a survey that enabled us to get a better understanding of drivers' attitude towards such an approach. It was followed with an experiment where 20 volunteers used the application while driving on a designated route. The results for each evaluation technique are presented in this section.

4.1 Survey

A total of 806 candidates responded the online form over a period of 7 days. The questionnaire was split into 6 sections, as follows;

- 1.General information regarding driver - Demographic data
- 2.Collision responsibility - This section was primarily targeted at filtering out drivers who were involved in a collision but were not responsible for the accident.
- 3.Collision involvement - Section 3 requests information on the number of collisions the questionnaire respondents were involved in, severity of the worst crash and time elapsed from obtaining his/her driving license and first collision.
- 4.Driving behavior analysis – Included a few questions present in the Aggressive Driving Behavior Questionnaire (ADBQ), which sought to identify aggression amongst different age groups.
- 5.General overview of application development, feedback and reports - A general description was provided on the application being developed. This section seeks to analyze whether drivers would consider an application which could analyze and provide live feedback based on their driving. Participants were also

asked if they would be interested in receiving driving reports, informing them of their overall driving performance and the frequency at which they would like to receive such reports.

6. Participation in research project - Section 6 asks whether participants would like to be involved in the driving behavior experiment with a short description of what it would entail.

All questions in sections 1,2,4,5 and 6 required a candidate response. Section 3 was based on the candidates' response from question 5.

4.1.1 Survey Results

An online questionnaire, created on Google Forms, was shared among numerous Facebook groups and friends in order to gather information related to driver aggression, collision rates and driving feedback. A total of 806 people responded to the questionnaire titled 'Motor Vehicle Collision & Driver Aggression Questionnaire', with the majority of respondents being female.

Driving experience varied in the younger age groups. Although in the majority of drivers 24 months had passed since having received their driving license, the older age groups (26-44, 45-60 and 1 over 60-year-old) included a few drivers who had only recently received their license (i.e. Less than 2 years). 626 of all respondents had been involved in a collision throughout their lifetime, with 259 of them being held mostly responsible for the collision they were involved in with another participant accountable for half the accident.

Comparing the number of times drivers were responsible for collisions, 54% had been involved in 1 collision especially among the 19-21 age group. More than a fourth of all drivers had caused two collisions, 12% for 3 and 7% were accountable for over 3 collisions. Assessing the duration after which one was to blame for a collision after having obtained their driving license, 53% had a collision within 2 years with 10% responsible for a collision after less than 3 months. This in turn highlighted the strong relationship present between a driver's experience and collision involvement.

Particularly addressing the questions present in section 5, overall 34% of respondents would make use of such an application while 46% and 20% answered 'maybe' and 'no' respectively. More than half preferred audible feedback such as 'slow down', 'good driving' if they were to have any sort of feedback which in turn would provide less distractions when compared to the other options provided (i.e. vibrations or visual feedback through a change in interface colours which would require the drivers' visual attention). A respondent suggested making use of a head up display (HUD) to provide feedback to the driver, preventing him/her from taking their eyes off the road. 65% of drivers were interested in receiving driving reports, with the majority preferring reports generated over longer driving durations.

4.2 Driving Experiment

Altogether, 20 participants took part in the driving experiment over a period of 2 weeks, all following the same pre-defined route of approximately 8.8km. The participants, 9 females and 11 males, were selected knowing they had a different driving record, level of driving experience and formed part of a different defined age group. During this experiment, the participants' cars were used for each subject and the application was installed on a device that was used for all subjects, it was safely secured in the car and transmitted data to a server. They were then asked to drive along the coast road to the roundabout present in Salina, Malta and back, with an approximate total distance of 8.8km.

Participants were asked for the total number of collisions they were responsible for and to drive as they would normally. Once the route had been completed, a break down percentage of the participants' driving behavior (Three classification percentages generated by the Fuzzy Logic outputs – Calm/Normal/Aggressive Styles) was told to the driver.

The stated route was chosen for three specific reasons;

1. The road was recently re-constructed (end of 2015) and therefore no holes/bumps would affect the smartphones sensor values.
2. The defined route is free of traffic congestion and would therefore produce an uninterrupted recording of driving behavior.
3. Traffic lights, roundabouts (depicted as round circles in figure 5) and different speed limits (red and blue lines in figure 5 referring to 50km/hr and 70km/hr respectively) would allow for different driving characteristics to be analyzed.

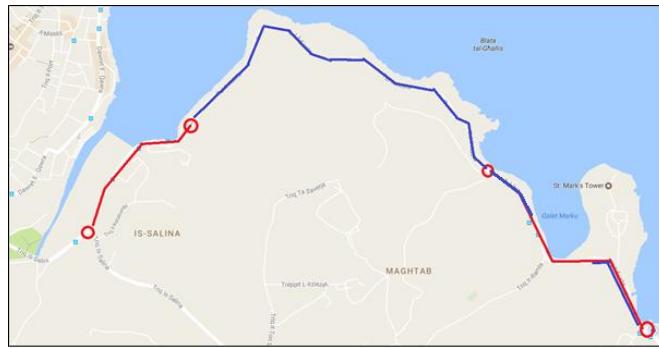


Figure 3. Illustration of the route used in the experiment

Six age groups were defined, with the majority of drivers lying in the 19 to 21 age group. One male was found for each respective 18 and over 60 age group. Younger drivers were split into smaller categories to possibly identify whether younger drivers possess a more aggressive driving style which may contribute to a higher collision rate.

To assess the number of collisions drivers were responsible for and compare their driving records, drivers were classified into 4 groups (0, 1, 2 and over 3). Collision involvement varied, with 2 participants accountable for over 3 collisions. Nearly half, 9 of the drivers, were never involved. The 18-year-old was responsible for one collision, having recently attained his driving license.

When evaluating the percentage of generated fuzzy output classification of driving behavior values according to their age group, a direct relationship was present as may be seen from the trend lines in Figure 4.

When comparing maximum and average speeds according to driver age groups, younger age groups exhibited a marginally higher average speed. Two aggressive drivers greatly contributed towards the higher average speed and average max speed in the 45-60 age group. The average maximum acceleration and deceleration values were also noticeably higher in younger age groups when compared to older age groups, implying that the drivers accelerated harder and decelerated quicker along the route.

Aggressive driving behavior was invariably found amongst drivers with a higher amount of collisions. Participant 1 displayed the highest aggressive driving behavior even though he was never involved in a driving collision. Opposing results were present for calm behavior, whereby drivers with no collisions had a high calm fuzzy output percentage. Drivers who obtained a high 'Normal' fuzzy output percentage, were accelerating or decelerating at an 'above normal' rate (determined by myself during preliminary tests and subsequently applied to fuzzy logic rules).

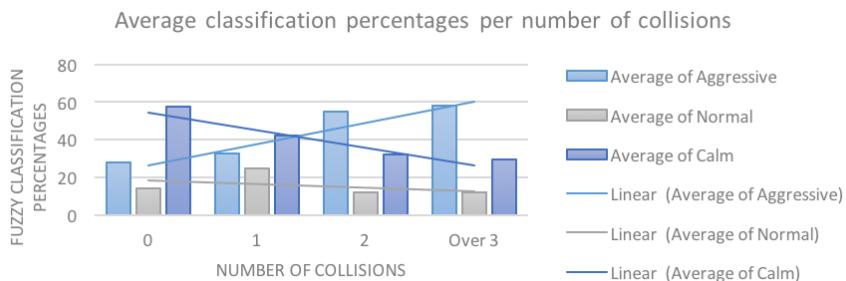


Figure 4. Number of collisions vs. average fuzzy output percentages

Average route time also exhibited relatively associated results, with drivers previously involved in zero collisions taking the longest to complete the route and drivers with over 3 completing the pre-determined route in approximately 1 minute 30 seconds less. Based on all the data generated from the driving experiment, one may conclude that different driving behaviors were found to be directly related to the drivers' age group as well as the number of collisions the driver was responsible for. Noticeably younger age groups and/or drivers responsible for more driving collisions (except one driver who had no collisions), as expected displayed higher levels of driving aggression.

Once the route had been completed, drivers were informed on the percentages of fuzzy output values generated (i.e. Driver's 'Aggressive', 'Normal' and 'Calm' values produced by the Fuzzy Inference System). Drivers generally asked how they fared with other drivers, with the majority falling into the same 'category' based on their age and/or number of responsible collisions. Aggressive drivers did not question the high aggressive percentage they produced, as they were aware of their driving style.

5. CONCLUSION

Taking into account the quantitative analysis that has been carried out on both the questionnaire and the driving experiment, results indicate age groups and the number of collisions a driver have are related to different driving styles and levels of aggression. The driving experiment accompanied by the survey produced insightful results, nevertheless having more drivers with a vast driving history, level of experience and age would have contributed to stronger results. Fuzzy Logic Inference and smartphone technology were combined throughout the project. Noticeably, results pointed out a correlation between driving behavior and aggression which not only was related to age but also to the number of collisions a driver had been involved in. The automatic profiling demonstrated in this paper provides various opportunities for research and potential applications. This profiling can be used for personalized educational campaigns, assisting road designers with general driving traits in specific locations or even categorizing drivers for risk assessment situations such as insurances. The architecture of this solution provides an opportunity for scalability since the software at client side can also be integrated in onboard computers running operating systems such as Android Auto. This all presents an opportunity for applications that can render our roads safer for the benefit of everyone.

REFERENCES

- Aljaafreh, A., Ashabatat, N., & Najim Al-Din, M. S. (2012). Driving Style Recognition Using Fuzzy Logic. International Conference on Vehicular Electronics and Safety (pp. 460-463). Istanbul, Turkey: IEEE.
- Antoniou, C., Papathanasopoulou, V., Gikas, V., Danezis, C., & Perakis, H. (2014). *Classification of driving characteristics using smartphone sensor data*. 3rd Symposium of the European Association for Research in Transportation. Greece.
- Bai, Y., & Wang, D. (2006). Fundamentals of Fuzzy Logic Control — Fuzzy Sets, Fuzzy Rules and Defuzzifications. In D. W. Ying Bai, Advanced Fuzzy Logic Technology in Industrial Applications (pp. 17-36). Springer London.
- Johnson, D. A., & Trivedi, M. M. (2011). *Driving Style Recognition Using a Smartphone as a Sensor Platform*. Intelligent Transportation Systems (ITSC) IEEE, San Diego.
- Hagras, H. (2007). Type-2 FLCs: A New Generation of Fuzzy Controllers. IEEE Computational Intelligence Magazine, pp. 30-43.
- Hulnhagen, T., Dengler, I., Tamke, A., Dang, T., & Breuel, G. (2010). Maneuver recognition using Probabilistic Finite-State Machines and Fuzzy Logic. 2010 IEEE Intelligent Vehicles Symposium (pp. 65-70). San Diego, CA, USA
- Magrin, E., Seychell, D., & Briffa, D. (2015). *Evaluating the use of mobile sensors in improving the user model in mobile recommender systems*. Proc. of the 8th IADIS International Conference on Information Systems. Madeira, Portugal.
- Paefgen, J., Kehr, F., Zhai, Y., & Michahelles, F. (2012). *Driving behavior analysis with smartphones: insights from a controlled field study*. MUM '12 Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia.
- World Health Organization (2015). *Global status report on road safety 2015*. Geneva, Switzerland

TOWARDS A MULTI-MODAL IT

Thomas Keller, Thierry Baumgartner and Elke Brucker-Kley

Zurich University of Applied Sciences, School of Management and Law

St.-Georgen-Platz 2, 8401 Winterthur, Switzerland

ABSTRACT

With the introduction of Gartner's term "bi-modal IT" in 2014, the problem of the need for a much faster response time in service delivery due to increasing digitalization has been highlighted. On the one hand, IT departments function as highly efficient and effective service organizations; on the other hand, IT is seen as a key factor for innovation. However, these two roles have completely different quality attributes and must therefore be treated adequately by management. Every company tries to solve this challenge independently and within the individual entrepreneurial context. This paper deals with this problem and tries to derive heuristics for multimodal IT structures from practice.

KEYWORDS

Bi-Modal IT, Multimodal IT, Innovation, Strategy

1. INTRODUCTION

The development of highly competitive markets and the rapidly increasing technological change is a new challenge for the IT of established companies. The development of new business models and innovative services is largely linked to the adaptation and innovation of new technologies (Bharadwaj et al., 2013:472). In this context, the importance of IT is becoming increasingly important for the further development of companies and the survival of the company as a critical success factor. The need to respond as quickly as possible to new market requirements is increasing without neglecting quality. In addition, existing systems and services must be able to continue to operate without interference and must meet both legal and safety requirements.

Gartner's bi-modal IT concept (Gartner Inc, 2014) describes two distinct modes within the architecture. Mode 1 focuses on the traditional sequential service delivery and ensures availability, continuity, security and reliability for the core business. Evolutionary and historically evolved legacy systems often form the core of existing business services. In contrast, Mode 2 as a counterpart manifests an explorative, non-linear approach with a focus on speed and agility as the basis for innovative developments through the use of the latest technologies (Gartner Inc., 2016).

Bi-modal entrepreneurship is the combination of a foreseeable, evolutionary development of existing IT services as well as the innovative and highly accurate development of new products and services (Bossert, Ip, & Laartz, 2014). This poses a particular challenge for long-established companies, which have to adapt their IT systems to digitalization and agile service delivery. Risk-hungry start-ups and small businesses without dependencies on legacy systems are less affected and have the advantage of being able to build up agile IT from scratch.

2. STATE-OF-THE-ART

The definition of a bi-modal IT by Gartner is already being reinterpreted, since it can also be a multimodal IT (Dion, 2016). The number of modes is not necessarily limited and can affect different areas of the IT architecture. Therefore, there are various theoretical implementation variants, which are strongly dependent on the business environment and its existing architecture. Possible approaches in the literature are described below.

As an example, the concept of bi-modal IT can be applied to the Archimate Framework (The Open Group, 2016). For example, it is possible to restrict mode 2 to the application layer. This allows the design and development of fast-paced and customer-oriented front-end applications (Mode 2), while in the background, operational services are used as a technological basis for operation on the basis of existing legacy systems (Mode 1). The difficulty is the interface between these two architectural layers. It must be possible to ensure compatibility on a permanent basis through the use of possibly very different technologies.

A further approach is based on the areas of responsibility when defining the modes. While the Chief Information Officer and the Chief Operation Officer control the Mode 1 domain, the Chief Digital Officer and the Chief Marketing Officer assume responsibility for Mode 2. Gartner's (Gartner Inc, 2016) approach does not provide any further detail on architecture, technology or innovation management, but demonstrates the importance of business leadership in the context of dynamic IT development. In particular, transformation is highly dependent on the existing corporate culture and, according to McKinsey, is an important component of the implementation success (Bossert, Ip, & Laartz, 2014).

In the recent past, there have also been critical voices pointing out the weaknesses of bi-modal IT. Being aware of these obstacles is essential for a successful implementation of the concept. Forrester Research criticizes bi-modal IT in that the incremental change and formation of a two-class IT system leads to an unnecessary increase in complexity and destroys the corporate culture (McCarty & Leaver, 2016). This is a short-term solution and not a meaningful investment in the long term, since the incentive to further develop existing legacy systems is lost and at the same time customer orientation suffers as a result (McCarthy, 2016). They recommend a comprehensive approach that is not based on different silos, places the customer at the centre of technology management rather than technology, and is not limited by the organization of the CIO. However, it is not explained how exactly the technology and innovation management is to be developed and where the potential for it comes from. The question remains whether or not innovations can be derived from customer requirements (Ulwick, 2005:105).

3. PROBLEM STATEMENT

There is a high diversity of implementation variants of multimodal IT structures in the literature. However, only limited information is available on its practical implementation and the underlying evaluation criteria. The procedures are at best theoretical approaches and have not been checked in practice or assessed qualitatively. Despite this, Gartner's CIO survey shows that two years after the need for a bi-modal delivery platform, 40% of the CIOs surveyed are already in a bi-modal IT environment (Gartner Inc., 2015). The Gartner CIO Survey 2016 was based on the feedback from 2,944 CIOs in 84 countries with a total IT budget of over 250 billion. After the preliminary literature analysis and the consultation of worldwide CIO survey surveys, it is clear that extensive knowledge regarding the implementation and operation of a multimodal IT architecture is already available in practice.

There is currently no formalization of a concrete procedural model or reference architecture that can be consulted as best practice approaches. The implementation of a multimodal IT must be built up successively without being able to assess its quality, effectiveness and achievement of objectives. The influence of the changes made on the development of new services, corporate culture, architecture and organization can only be determined retrospectively. As a result, the implementation of multimodal IT is strongly associated with unpredictable impacts and contradicts a goal-oriented approach.

Based on initial interviews, it became clear that the companies were developing multimodal IT structures using the "trial and error" method. The interest for alternative possibilities, structures and procedures is just as great as a qualitative and cross-sectoral IT and innovation benchmarking. Since neither an adequate reference architecture nor the corresponding measurement criteria have been defined, it is not possible to make an assessment with regard to one's own level of development.

4. OBJECTIVES

The aim of this paper is to derive heuristics for multimodal IT structures from practice based on the analysis of existing implementations of multimodal IT structures in reference companies. In addition, qualitative

evaluation factors are determined on the basis of renowned quality attributes from software engineering (Bass, Clements, & Kazman, 2013). These serve to evaluate and further develop an existing architecture. The inductively developed reference model is to be considered as a best practice approach for the initiation and further development of a multimodal IT architecture. This work is of scientific relevance, as there is currently no visualized representation of the reality of a coherent multimodal IT architecture.

5. RESEARCH DESIGN

The formulation of open research questions makes it possible to pursue a qualitative research approach. When deciding on the research design, it was taken into account that the main component of existing knowledge can be derived from practical systems. Based on this information, the following construction of a reference model is carried out. Such a system can only contribute to quality improvement if it is as up-to-date as possible and tailored to the specific application context (Becker, Delfmann, & Rieke, 2007:103). It is essential that the artifact reflects the actual state of knowledge. Consequently, a constructivistic approach is pursued which has its origin in Design Science (Hevner et al., 2004). As a research method, surveys in the form of interviews as well as inductive reference modelling are used. The underlying research paradigm follows the principles of Action Research. This makes it possible to transfer knowledge directly from practice to a reference model (action), operationalize it and critically assess and optimize it using qualitative attributes (research).

The reference model is designed and constructed step by step. Following each individual survey, the information collected is integrated into the reference model. The iterative process integrates the individual implementation variants into an overall model and forms the design cycle of the ISR framework (Hevner et al., 2004). After the integration of an additional architecture variant, the model is validated and optimized anew. This shall facilitate the process of standardization and allow misinterpretations to be detected at an early stage. In case of inconsistencies between individual experts, these are examined for their context and presented to the experts for discussion again. The modeling process is completed as soon as all individually collected implementation variants of the practice partners are part of the model.

Following table lists the involved companies in the interviews.

Table 1. Associated companies

Enterprise	Industry
Swiss Re	Insurance
Staatskanzlei Kanton Zürich Stabsstelle E-Government	Public Administration
SIX Group LTD	Finance
Zürcher Kantonalbank (ZKB)	Finance
Post CH LTD	Logistics

In a subsequent step the model has been checked for its applicability and pragmatism. For the following companies a case study has been elaborated which will be published separately.

Table 2. Validation companies

Enterprise	Industry
Swisslog Ltd	Logistics
Belimed Ltd	Production
SIX Group Ltd	Finance
Post CH Ltd	Logistics

As can be seen from the above two tables SIX Group Ltd as well Post CH Ltd have been involved in both phases, the interviews as well as the case studies. This poses a shortcoming. However, both companies are very large and have multiple examples of multimodal structures.

6. RESEARCH FINDINGS

The reference model is based on the results of the expert interviews and on the subsequently elaborated case studies. The model is valid for all experts and does not conflict with individual statements. Each dimension describes a specific topic that is relevant for multimodal IT and justifies as a dimension when it shows varying manifestations depending on the mode. The corresponding design is described on the basis of characteristics that are assigned to a single dimension. The shown characteristics are often in the focus area of one of the three modes. However, they can overlap the boundaries of modes. Figure 1 illustrates the graphical representation and describes the elements of the artifact.

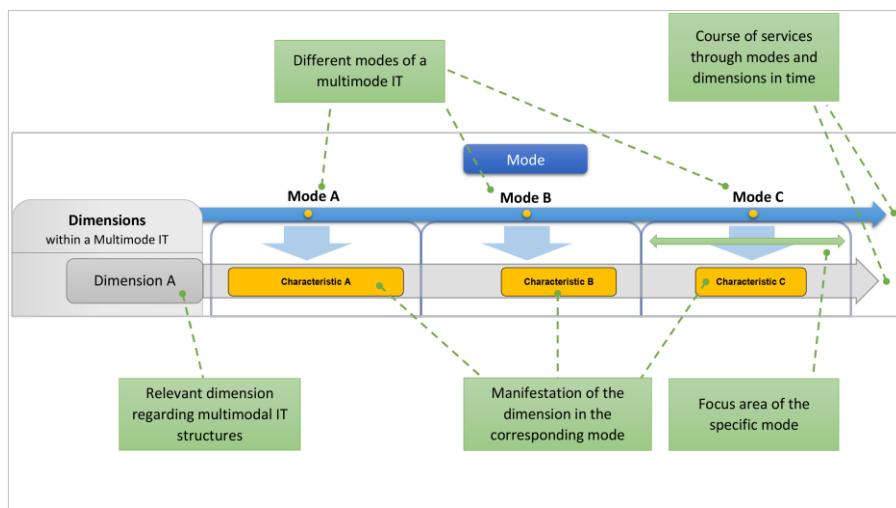


Figure 1. Explanation of elements of the framework

The three modes, which follow a different philosophy and speed, consist of the agile mode of the "Pioneers", the traditional mode of the "Town Planners" and the transformation mode of the "Settlers".

The external influencing factors that are subject to the reference model were derived on the basis of the St. Gallen Management Model (Rüegg-Stürm & Grand, 2015) and the information provided by the surveyed experts. The individual statements can be seen in the expert interviews on the basis of the question "Environmental". These factors can only be influenced slightly or not at all. However, their overall influence on the design of multimodal IT is very high. Figure 2 shows the external influencing factors of the reference model and the final reference model. The relevant dimensions are described within the reference model by characteristics. This defines the necessary competencies with which an IT system can be operated in several modes. In the following, differentiated action guidelines are formulated for each dimension. These guidelines can be consulted in two ways. On the one hand, to review and optimize an existing application of multimodal IT structures. On the other hand, as a strategic transformation planning instrument for the transition from traditional to multimodal IT. All involved practice partners have an extensive operational IT architecture. This forms the basic mode of "town planners" through continuous growth. The goal of the strategic transformation planning is to create the two more agile modes of the "settlers" and "pioneers".

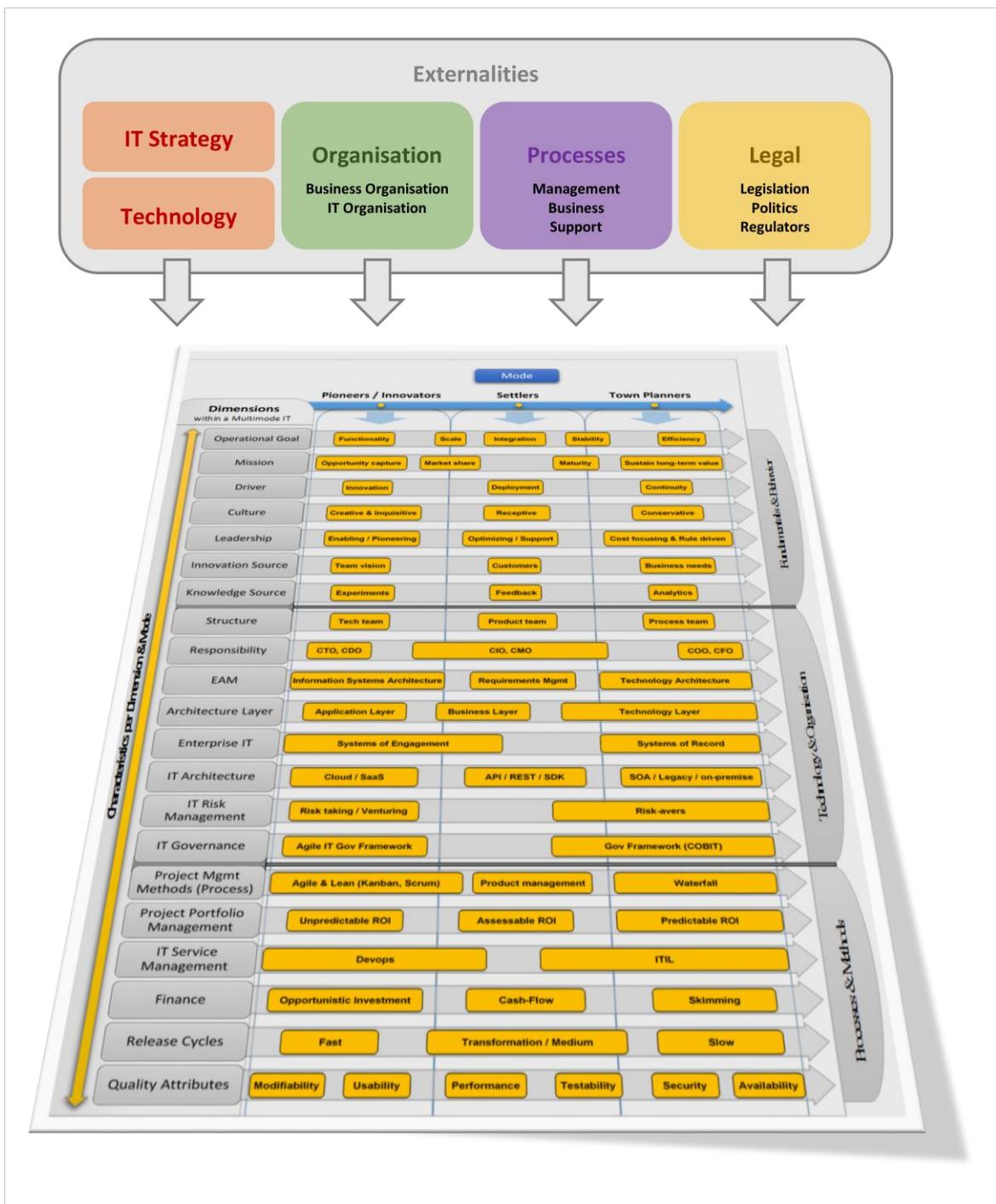


Figure 2. Complete framework

The guidelines for action provide a more detailed insight to enable the necessary competencies and possible transformation. It should be noted that the reference model includes all dimensions which were considered relevant by the practice partners. These are all large corporations with many years of business activities and core competencies, which now have to be coupled with flexible service provision and innovation. In the changed corporate context, these are relevant in different ways. As an example, the governance dimension has only limitations on the agile side of the reference model if there is clear and enforced IT governance at present. If this is only true to a limited extend, this dimension does not have to be

taken into account for the time being. Accordingly, when using the reference model the dimensions most relevant to your own company can be derived (whether they are used for optimization or for initial transformation is irrelevant). However, no dimensions should be omitted that are subjectively regarded as less important. In Gartner's view, leadership and culture are crucial for the success of transformation (Rivera & Van der Meulen, 2014).

According to the defined delimitation of this work, known frameworks (such as ITIL or COBIT) as well as known philosophies (DevOps) and project management methods (SCRUM, Lean, Kanban) are not discussed further. There are sufficient sources of information in the literature to introduce such rules and regulations. The practical experience with the aforementioned concepts is already well established and can be easily consulted for use in the corporate context. In this context, the reference model merely describes which concepts are relevant and why they need to be implemented.

For optimal alignment and planning of multimodal IT, the output required after application of the guidelines for action is summarized below. A more comprehensive description is available at the authors.

Table 3. Output after application of the action guidelines

Dimension	Output after application of the action guidelines	
Operational Goal	Independent operational target definition per mode.	Fundamentals & Behavior Technology & Organisation
Mission	Define an individual IT mission for each of the three modes.	
Driver	Establishing central drivers of all modes in the company.	
Culture	Establishment of a cultural basis for all three modes, taking into account their respective characteristics. Organizational reallocation or acquisition of employees with the required skills.	
Leadership	Development and expansion of the management culture to promote agile modes, innovation, agility and transformation. Organizational role assignment within the modes according to the dimension "Responsibility".	
Innovation Source	Establishment of the innovation culture and establishment of associated instruments (boards, meetings and registration portals).	
Knowledge Source	Definition and development of knowledge sources on innovative technologies and projects.	
Structure	Organizational definition of different teams per mode as well as their design and integration into the company.	
Responsibility	Define clear responsibility structures for the three modes.	
Enterprise Architecture Management (EAM)	Coordination of Enterprise Architecture Management through different iteration cycles of the architecture adaptation. Integration of the modes into the architecture development method of the framework TOGAF (The Open Group, 2011).	
Architecture Layer	Organizational assignment of the modes to the three layers Application, Business and Technology in correspondence with the Archimate Framework (The Open Group, 2016).	
Enterprise IT	Development of the ability to implement "Systems of Engagement" applications in short development cycles.	
IT Architecture	Creation of a service-oriented service architecture with a focus on cloud technologies. Use of application interfaces for orchestration and networking of innovative services with existing structures.	
IT Risk Management	Adaptation of risk management for agile projects by creating an increased acceptance of unknown risks.	
IT Governance	Establishment of a simplified governance framework for agile mode (pioneers) by selecting and weakening existing governance	

	rules in traditional mode (town planners).	
Project Management Methods (Process)	Implementation of agile project management methods.	Processes & Methods
Project Portfolio Management	Creation of an additional tool for qualitative project evaluation in agile projects (substituted business cases as a tool for the pioneers). Creating acceptance of project execution without a quantifiable return on investment.	
IT Service Management	Integration of DevOps as a philosophy for agile development methods. Ensuring alignment with existing IT Service Management frameworks such as the IT Infrastructure Library.	
Finance	Adjustment of capital expenditure with alignment of the new instruments of project portfolio management. Clear definition of financial processes to support innovative projects (pioneers).	
Release Cycles	Adaptation of fast release cycles (2-4 weeks for agile modes). Coordination of the different development cycles across the modes.	
Quality Attributes	Definition of the strategic direction of the "Settlers" and "Pioneers". Ongoing target verification through benchmarking.	

7. CONCLUSION

The reference model includes the dimensions considered relevant by the experts involved. It also shows the necessary measures to establish agile modes. However, since the large companies involved are still in the process of being developed, the practical experience gained cannot be regarded as an irrefutable definition. Due to the fact that the multimodal structures in these companies have been developed iteratively, there is currently no possibility to define an adequate prioritization as order of application. The resources required to transform individual dimensions are very demanding. Due to usually limited financial resources, it would make sense to provide a cost-benefit assessment.

By delimiting this work, the multimodality was strongly concentrated on IT. According to some experts, however, such an IT can only function in practice if the business is aligned to this. This approach has to be taken into account in particular by the high criticality of business/IT alignment. The connection to the business requirements takes place in the "settlers" mode. However, it has not been followed up in more detail as part of this work. It is therefore advisable to simplify and extend the reference model to the business organization.

The collection of the information used to create the reference model and guidelines is based on data and statements of only five companies. This is partly due to the specific context in which such large companies are currently located. The thin concentration of the consulted experts was necessary in order to narrow down the application to the existing problems.

REFERENCES

- Bass, L., Clements, P., & Kazman, R. (2013). *Software architecture in practice - 3rd edition*. Reading Massachusetts: Addison-Wesley.
- Becker, J., Delfmann, P., & Rieke, T. (2007). *Effiziente Softwareentwicklung mit Referenzmodellen*. Heidelberg: Physica-Verlag.
- Bharadwaj, A., El Sawy, O., Pavlou, P., & Venkatraman, N. (June 2013). Digital Business Strategy: Toward a next generation of insights. *MIS Quarterly*, 37(2), S. 471-482.
- Bossert, O., Ip, C., & Laartz, J. (Dezember 2014). A two-speed IT architecture for the digital enterprise. *McKinsey on Business Technology*. Digital McKinsey.

- Dion, H. (1. Mai 2016). *IT leaders inundated with bimodal IT meme*. Von ZDNET: <http://www.zdnet.com/article/it-leaders-inundated-with-bimodal-it-meme> abgerufen
- Gartner Inc. (15. Juli 2014). *Bimodal IT: How to Be Digitally Agile Without Making a Mess*. Gartner Digital Business.
- Gartner Inc. (5. Oktober 2015). Gartner Survey of More Than 2,900 CIOs Reveals Shift Toward Platform Thinking. Orlando, Florida, USA. Abgerufen am 15. November 2016 von <http://www.gartner.com/newsroom/id/3143018>
- Gartner Inc. (18. Februar 2016). *Deliver on the Promise of Bimodal*. Gartner Digital Business.
- Hevner, A., Ram, S., March, S., & Park, J. (März 2004). Design Science in Information Systems Research. *MIS Quarterly*, 8(1).
- McCarthy, J. (2016). *Avoiding The Bimodal IT Traps And Pitfalls*. Forrester Research.
- Rivera, J., & Van der Meulen, R. (2014). *Gartner Says CIOs Need Bimodal IT to Succeed in Digital Business*. Barcelona, Spain: Gartner Inc.
- Rüegg-Stürm, J., & Grand, S. (2015). *Das St. Gallen Management-Modell (2. vollständig überarbeitete und grundlegend weiterentwickelte Auflage)*. Bern: Haupt-Verlag.
- The Open Group. (2011). *TOGAF Version 9.1*. Zaltbommel: Van Haren Publishing.
- The Open Group. (2016). *ArchiMate® 3.0 Specification*. Zaltbommel: Van Haren Publishing.
- Ulwick, A. (2005). *What Customers Want: Using Outcome-driven Innovation to Create Breakthrough Products and Services*. Lewiston, NY, U.S.A: McGraw-Hill.

DNA-BASED USER AUTHENTICATION SCHEMES FOR WIRELESS BODY AREA NETWORK

Esma Ergüner Özkoç

*Department of Management Information Systems Başkent University
Ankara-Turkey*

ABSTRACT

Today, the rapid spread of wearable/implantable wireless technology, especially in the field of health, has brought some security weaknesses. In this study, an authentication scheme is proposed with the aim of preventing unauthorized persons from accessing the private data of individuals while communicating over Wireless Body Area Network (WBAN). The proposed scheme is separated from previously generated schemes by using a randomly generated DNA sequence that increases the level of security but does not add any additional cost to the system. Security analysis of proposed scheme is given.

KEYWORDS

Pseudo DNA Cryptography, DNA Computing, DNA-OTP Wireless Body Area Network-WBAN, Authentication

1. INTRODUCTION

WBAN can be defined as a technology based on the automatic collection of data with the help of tiny sensors attached to the human body. It is essentially a wireless technology with a very short range. The main reason for the recent increase in interest in WBAN can be explained as the diversity in the areas of use, such as telemedicine, education and entertainment, sports, military, medical monitoring, disease early detection and treatment. Today, it is used most effectively in healthcare systems. WBAN sensors can be directly attached to human body as well as placed under the skin in order to gather some health data (Electro Cardio Gram - ECG, Electro Encephalo Gram-EEG, temperature, blood pressure, blood glucose, etc.) One of the major security requirements of the secure WBAN communication is the mutual authentication which refers to two sides authenticating each other at the same time. Previously developed authentication protocols for WBAN can be categorized as cryptographic based methods (Li, X. et al., 2017-a) and non-cryptographic based methods (Li X. et al., 2017-b) Public key infrastructure (PKI), identity-based cryptosystem (IBC) and elliptic curve cryptography (ECC) are included in the cryptographic method class while physiological signal based, channel-based and the proximity-based schemes are included in non-cryptographic based methods. Beside these DNA-based authentication method has not been used previously for WBAN. In proposed scheme, randomly generated DNA sequence is act as an one time pad (OTP) which is in principle unbreakable.

Adelman's use of DNA molecules for the solution of the Hamiltonian path problem (Adleman, L.M., 1994), which is an NP problem in 1994, is regarded as the beginning of "DNA Computing" studies. In 1995, DNA-based cryptography studies can be categorized as encryption algorithms(Cui, G. et al., 2008, Geng, X., Pan, L. and Xu, J., 2008, Prabhu, D. and Adimoolam, M., 2011, Sadeg, S. et al., 2010, Wang, X., Zhang, Q. and Wei, X.P., 2010, Yunpeng, Z. et al., 2011) -image encryption algorithms (Maniyath, S.R. and Supriya, M., 2011, Zhang, Q. et al., 2009, Zhou, C. et al., 2010, Zhou, S., Zhang, Q. and Wei, X., 2010) -OTP (One Time Pad) (Borda, M. and Tornea, O., 2010, Gehani, A., LaBean, T. and Reif, J., 2003, Shreyas, C., 2013.), data hiding, steganography (Huang, Y.H., Chang, C.C. and Wu, C.Y., 2014, Jiao, S. and Goutte, R., 2008, Liu, H., Lin, D. and Kadir, A., 2013, Tuncer, T. and Avci, E., 2016 , Chang, C.C. et al., 2007), and authentication (Lakshmi, P.V. and Susan, V.S., 2010, Raju, P.V.S.N. and Parwekar, P., 2015, Vijayakumar P., Vijayalakshmi V. and Zayaraz G. 2014) and watermark (Heider, D. and Barnekow, A., 2007, Lee, S.H., 2014). DNA studies are performed either by biochemically testing tubes, or by simulating DNA functions

(Pseudo or Virtual DNA cryptography). The use of DNA, for cryptographic purposes requires a very high laboratory specifications. This has led researchers to develop methods for cryptographic processes that simulate real-world DNA processes or inspired rather than use real DNA. This kind of cryptography is called Pseudo DNA cryptography (Saeed Al-Wattar A.H. et al., 2015).

Some characteristic features of DNA make it attractive for use in solving complex problems. These features are important for computer and encryption such as parallelism, energy efficiency, randomness of endless DNA sequences, easy access to DNA sequences, similarity to binary system, data storage capacity (Saeed Al-Wattar A.H. et al., 2015). Although there are many studies on DNA cryptography, the process has not yet reached maturity and security features must be provided at a higher level than existing.

The rest of the study is organized as follows. In section 2, the DNA structure and calculation principles, in section 3, the proposed scheme is explained in detail, and in the last section, the security analysis of the proposed method is included.

2. DNA STRUCTURE

DNA (Deoxyribo Nucleic Acid) is a nucleic acid that carries the genetic instructions necessary for the vital functions and biological evolution of all organisms and some viruses. A DNA's single-stranded (ssDNA) is form of a sequence of molecules named bases. There are four type of DNA bases: adenin (A), thymine (T), guanine (G), and cytosine (C). These bases constitute the DNA alphabet and grouped into complementary pairs (A-T, G-C) (Figure1). One of the most basic features of the DNA's strand sequence is that its order: such as AGCA is not equal to ACGA. Just as the words in the alphabet are organized to form words, bases that compose the DNA are also a specific order which provide all the information necessary to make a new copy of the entire molecule. This specific order of bases might compose a "genetic code".

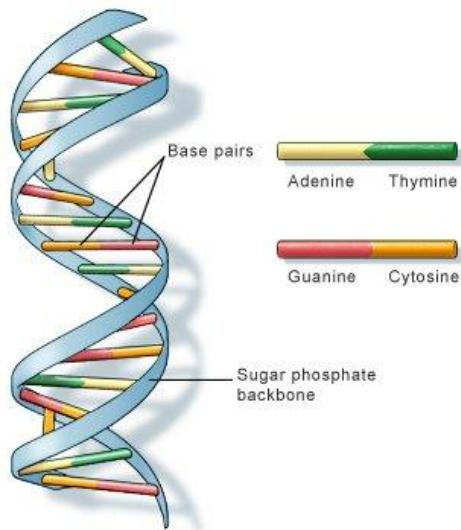


Figure 1. DNA structure –Watson-Crick Double Helix (U.S. National Library of Medicine, 2018)

As shown in Figure 1, DNA strands exist as paired reverse complementary words or strands. All the logical operations can be performed over DNA bases like XOR, AND, OR, XNOR, NOT and binary coding for the DNA bases could be changed. In Table 1, eight kind of binary code conversion are given for DNA bases. If first conversion {A↔ 00, T↔ 11, G↔01, C↔10} are chosen from Table 1, corresponding results for XOR operation are given in Table 2.

Table 1. Kinds of schemes encoding map rule of DNA sequence

DNA Base	Binary Code							
A	00	00	01	01	10	10	11	11
T	11	11	10	10	01	01	00	00
G	01	10	00	11	00	11	01	10
C	10	01	11	00	11	00	10	01

Table 2. XOR operation for DNA sequence for the case: A- 00 T- 11 G-01 C-10

XOR	A	T	G	C
A	A	T	G	C
T	T	A	C	G
G	G	C	A	T
C	C	G	T	A

Table 3. Complemented table for A,T,G,C

DNA Base	Complimented Base
A	T
T	A
G	C
C	G

3. PROPOSED SCHEME

In this section, proposed WBAN authentication scheme based on pseudo DNA cryptography will be explained. Proposed authentication scheme contains three entities: (i) WBAN client who request a service from application provider, (ii) Network Manager that manages the whole WBAN and plays the role of private key generator, (iii) Application provider which provides a services in the network for authenticated WBAN client. Proposed scheme consist of three phases; the registration phase, the verification phase and the password changing phase. Firstly, the notation, which we used in the rest of the paper, is presented at the Table 4.

Table 4. Notation

Notations	Description
U_i	User i (WBAN client)
NM	Network Manager
AP	Application Provider
ID_i	U_i identity
DNA_i	Random DNA strand sequence which is generated by NM for U_i
\overline{DNA}_i	Compliment of DNA_i strand sequence
T, T', T''	Timestamp
$h(\cdot)$	Cryptographic one-way hash function
$h_k(\cdot)$	Cryptographic keyed one-way hash function
\parallel	Concatenation of bits
\oplus	XOR operation

3.1 Registration Phase

This phase occurs only once for each user when the WBAN client wants to register the system over network. The steps of this phase are given as follows (Figure 1):

1. WBAN client (U_i) submits her/his ID_i to the NM over network.
2. The NM generates a random DNA sequence (DNA_i) for U_i . Then calculates $X = h(S_{NM} || ID_i)$ and $Y = DNA_i \oplus X$ by using NM's Secret value (S_{NM}). NM calculates also $Z = h_{DNA_i}(\overline{DNA}_i || T)$ value by using randomly generated DNA sequence's complements (\overline{DNA}_i) (see Table 3) and Timestamp. Calculated Z, Y and Timestamp transmitted to WBAN client.
3. When WBAN client takes the NM's message, first he/she checks Timestamp, is it in valid time interval or not. If it is valid, U_i , Calculates X' , and obtain DNA'_i which is used to calculate Z' . If calculated Z' and received Z are equal, WBAN client verifies NM and compose $PW_i = h(S_u || X)$. WBAN client sends $PW_i \oplus DNA'_i$, $h(PW_i)$ to the verified NM.
4. NM checks calculated $h(PW_i)$ is equal to received one, if it is, then NM verifies WBAN client too.

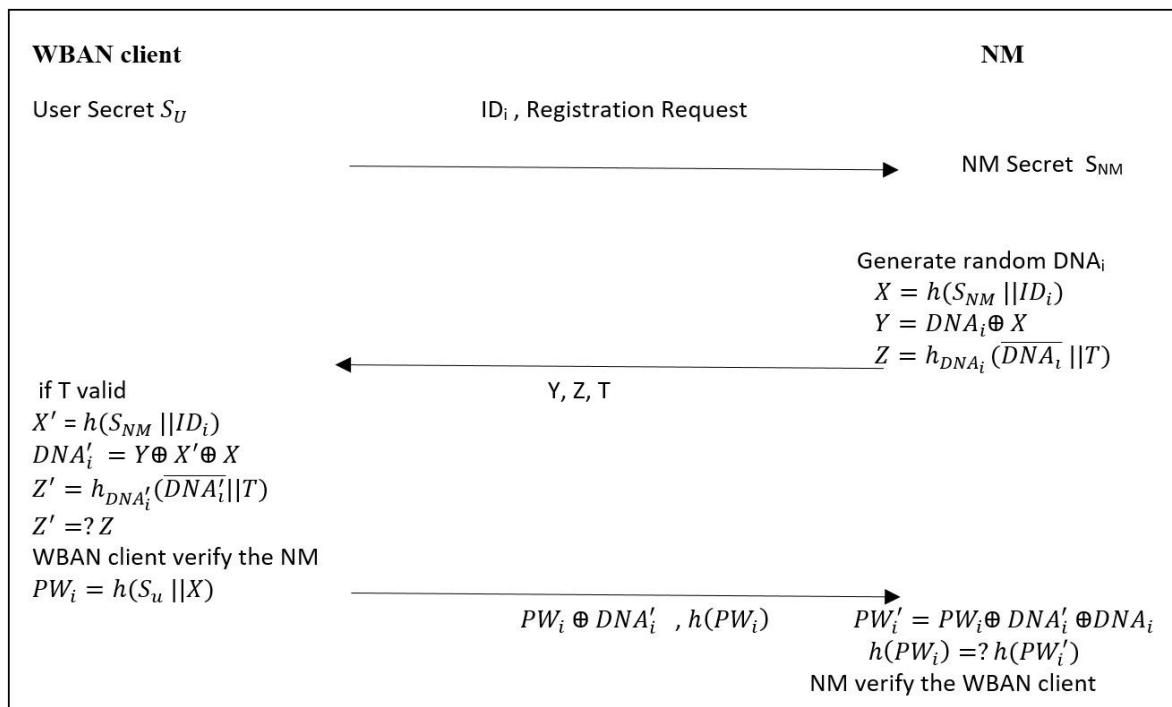


Figure 2. Registration phase of the proposed scheme

3.2 Login and Authorization Phase

This phase is implemented between 3 sides: WBAN client, NM and AP. The phase details are presented in the Figure 3 Shortly;

1. WBAN client (U_i) submits her/his ID_i and login request to the NM over network.
2. The NM generates a random DNA sequence (DNA_t) for U_i . Then calculates X and Y by using NM's Secret value (S_{NM}). NM calculates also $Z = h_{DNA_t}(\overline{DNA}_t || T)$ value .Calculated Z, Y and Timestamp T send to WBAN client.
3. When WBAN client receive the NM's message, first he/she checks Timestamp, is it in valid time interval or not. If it is valid, U_i , calculates X' , and obtain DNA'_t which is used to calculate Z' . If calculated Z' and received Z are equal, WBAN client verifies NM and compose $V_i = PW_i \oplus h(PW_i || T' || DNA'_t)$ WBAN client sends V_i, T' to the verified NM.

4. NM calculates PW'_i and checks its consistency with PW_i . If it is equal, NM is authorize WBAN client and calculate S_i . Calculated S_i and current timestamp T'' send to AP.
5. The AP, calculates S'_i by using its secret value S_n , and responds WBAN client's query if received S_i and calculated S'_i are equal.

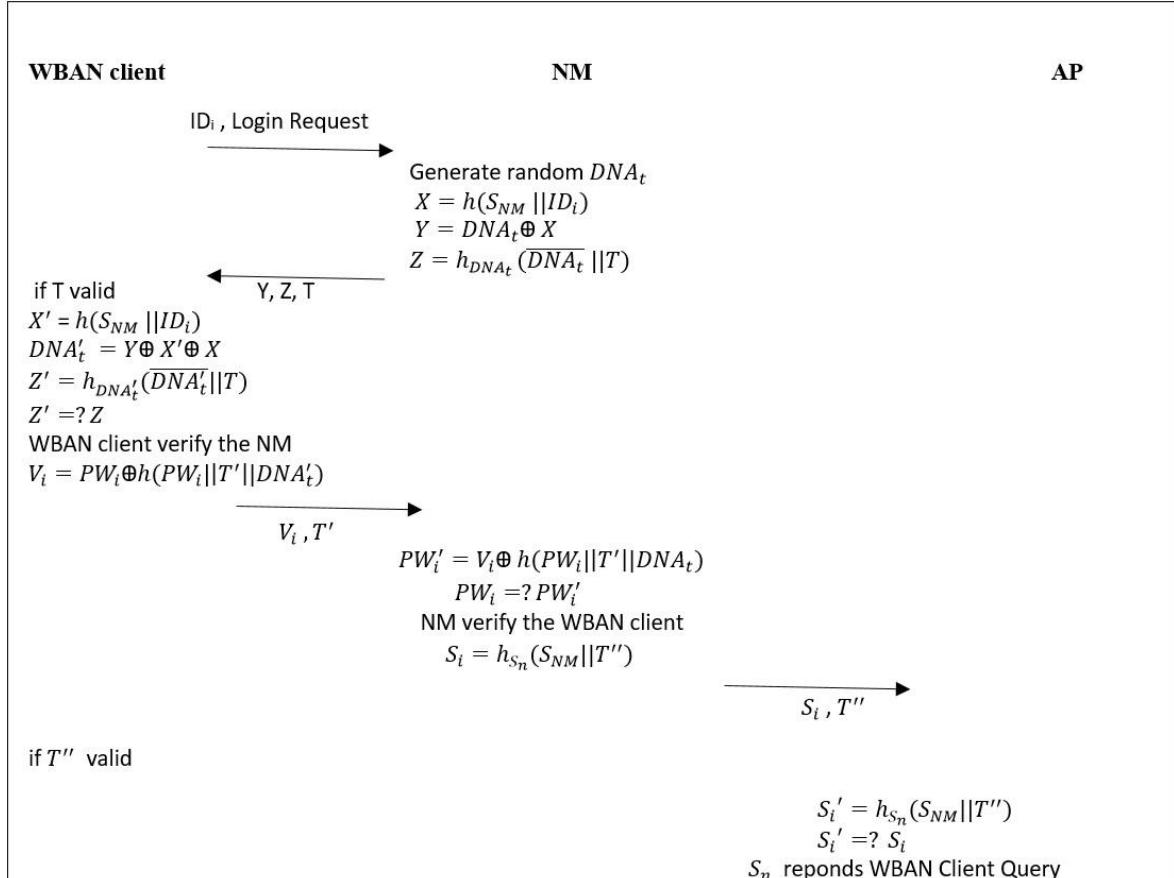


Figure 3. Authentication phase of the proposed scheme

3.3 Password Change Phase

Proposed scheme let WBAN client to change password after registration phase. Before changing password NM, controls existing password (PW_i) validity (Figure 4).

1. WBAN client selects new S'_U , and calculates new PW'_i . WBAN client compose a message which consist of $PW'_i \oplus S_{NM}$, $h(PW_i || PW'_i)$ and send it to the NM
2. NM calculates PW''_i from received message and checks $h(PW_i || PW'_i) =? h(PW_i || PW''_i)$ equality. If two hash value are equal then NM change the stored password with new one.

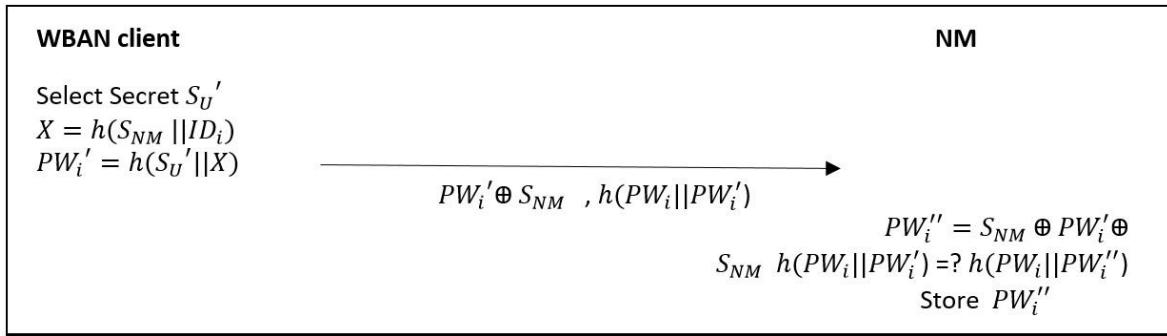


Figure 4. Password change phase of the proposed scheme

4. SECURITY ANALYSIS

Server Spoofing Attack: A server spoofing is an attack in which malicious person or program masquerading as a NM in order to access confidential information about the user. Proposed scheme can prevent this type of attack by using Secret of NM (S_{NM}) which is also stored in WBAN client's smart card in registration phase.

Stolen Smart card Attack: Even if the attacker obtains a smart card in which the (S_{NM}) is embedded, it cannot calculate the $PW_i = h(S_u || X)$ because attacker does not have the user secret $\{S_u\}$.

Stolen-Verifiers Attack: Even if the attacker obtains the PW_i value from NM. Since $h(.)$ is a secure one-way hash function, it is not possible for an attacker to calculate secret of the user $\{S_u\}$.

Pre-play Attack: it is not possible for an attacker to predict the next DNA sequence and change the timestamp value in a reasonable amount of time.

Active Attack and Revelation of Message Content: Messages sent/received during login and authentication phase are not in plain text format. All the transmitted messages are hashed with secure one-way (or Keyed) hash function.

Replay Attack: If an attacker wants to replay the authentication message $\{Y, Z, T\}$ in login and authentication phase. First he/she must choose a valid Timestamp $\{T\}$, then must calculate the corresponding $\{Z\}$ value that is depend on T and randomly generated DNA strand. It is fairly difficult task for proposed scheme thus attacker cannot validate himself as a legal NM.

5. CONCLUSION

WBAN security is still in its infancy even though there have been many studies on it. Balancing features such as security, efficiency and practicality in the design of a WBAN authentication scheme is extremely important. In this study, the authentication scheme for WBAN is proposed which security is enhanced by using pseudo-DNA sequence.

It provides mutual authentication between three actor of the scheme; WBAN client, network manager and application provider. The cost of the computation of the proposed scheme is low and it is based on keyed hash, hash function and XOR operations. To guarantee the WBAN's running time, the energy consumption of the wearable / installable sensors must be very low, which is why the low cost scheme is preferred. Finally security analysis of proposed scheme and it withstands the known attacks are given.

REFERENCES

- Adleman, L.M., 1994. *Molecular computation of solutions to combinatorial problems*. Nature, 369, p.40.
 Borda, M. and Tornea, O., 2010, June. *DNA secret writing Techniques*. In Communications (COMM), 2010 8th International Conference on (pp. 451-456). IEEE.

- Chang, C.C., Lu, T.C., Chang, Y.F. and Lee, R.C.T., 2007. *Reversible data hiding schemes for deoxyribonucleic acid (DNA) medium*. International Journal of Innovative Computing, Information and Control, 3(5), pp.1145-1160.
- Cui, G. et al., 2008, September. *An encryption scheme using DNA technology*. In *Bio-Inspired Computing: Theories and Applications*, 2008. BICTA 2008. 3rd International Conference on (pp. 37-42). IEEE.
- Gehani, A., LaBean, T. and Reif, J., 2003. *DNA-based cryptography*. Lecture notes in computer science, 2950, pp.167-188.
- Geng, X., Pan, L. and Xu, J., 2008. *A DNA sticker algorithm for bit-substitution in a block cipher*. Journal of Parallel and Distributed Computing, 68(9), pp.1201-1206.
- Heider, D. and Barnekow, A., 2007. *DNA-based watermarks using the DNA-Crypt algorithm*. BMC bioinformatics, 8(1), p.176.
- Huang, Y.H., Chang, C.C. and Wu, C.Y., 2014. *A DNA-based data hiding technique with low modification rates*. Multimedia tools and applications, 70(3), pp.1439-1451.
- Jiao, S. and Goutte, R., 2008, October. *Code for encryption hiding data into genomic DNA of living organisms*. In *Signal Processing, 2008. ICSP 2008. 9th International Conference on*(pp. 2166-2169). IEEE.
- Lakshmi, P.V. and Susan, V.S., 2010. *Biometric Authentication Using ElGamal Cryptosystem And DNA Sequence*. International Journal of Engineering Science and Technology, 2, pp.1993-1996.
- Lee, S.H., 2014. *DWT based coding DNA watermarking for DNA copyright protection*. Information Sciences, 273, pp.263-286.
- Li, X. et al., 2017-a. *An enhanced 1-round authentication protocol for wireless body area networks with user anonymity*. Computers & Electrical Engineering.
- Li, X. et al., 2017-b. *Anonymous mutual authentication and key agreement scheme for wearable sensors in wireless body area networks*. Computer Networks.
- Liu, H., Lin, D. and Kadir, A., 2013. *A novel data hiding method based on deoxyribonucleic acid coding*. Computers & Electrical Engineering, 39(4), pp.1164-1173.
- Maniyath, S.R. and Supriya, M., 2011. *An uncompressed image encryption algorithm based on DNA sequences*. Computer Science and Information Technology, 2, pp.258-270.
- Prabhu, D. and Adimoolam, M., 2011. *Bi-serial DNA Encryption Algorithm (BDEA)*. arXiv preprint arXiv:1101.2577.
- Raju, P.V.S.N. and Parwekar, P., 2015. *DNA Encryption Based Dual Server Password Authentication*. In *Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014* (pp. 29-37). Springer, Cham.
- Sadeg, S. et al., 2010, October. *An encryption algorithm inspired from DNA*. In *Machine and Web Intelligence (ICMWI), 2010 International Conference on* (pp. 344-349). IEEE.
- Saeed Al-Wattar A.H. et al., 2015. *Review of DNA and Pseudo DNA cryptography*, International Journal of Computer Science and Engineering (IJCSE) , 4(4), pp. 65-76
- Shreyas, C., 2013. *DNA Cryptography Based on DNA Hybridization and One Time pad scheme*. International Journal of Engineering Research & Technology (IJERT), 2(10).
- Tuncer, T. and Avci, E., 2016. *A reversible data hiding algorithm based on probabilistic DNA-XOR secret sharing scheme for color images*. Displays, 41, pp.1-8.
- U.S.National Library of Medicine, <https://ghr.nlm.nih.gov/primer/basics/dna> (Last access:01Jan2018)
- Vijayakumar P., Vijayalakshmi V. and Zayaraz G. 2014. *DNA based Password Authentication Scheme using Hyperelliptic Curve Cryptography for Smart Card* Proc. of Int. Conf. on Recent Trends in Information, Telecommunication and Computing, pp. 517-523.
- Wang, X., Zhang, Q. and Wei, X.P., 2010. *A new encryption method based on Rijndael algorithm and DNA computing*. In *Applied Mechanics and Materials* (Vol. 20, pp. 1241-1246). Trans Tech Publications.
- Yunpeng, Z. et al., 2011, October. *Index-based symmetric DNA encryption algorithm*. In *Image and Signal Processing (CISP), 2011 4th International Congress on* (Vol. 5, pp. 2290-2294). IEEE.
- Zhang, Q., Guo, L., Xue, X. and Wei, X., 2009, October. *An image encryption algorithm based on DNA sequence addition operation*. In *Bio-Inspired Computing, 2009. BIC-TA'09. Fourth International Conference on* (pp. 1-5). IEEE.
- Zhou, C., Wei, X., Zhang, Q. and Liu, R., 2010. *DNA sequence splicing with chaotic maps for image encryption*. Journal of Computational and Theoretical Nanoscience, 7(10), pp.1904-1910.
- Zhou, S., Zhang, Q. and Wei, X., 2010, November. *Image encryption algorithm based on DNA sequences for the big image*. In *Multimedia Information Networking and Security (MINES), 2010 International Conference on* (pp. 884-888). IEEE.

PERSONAL DATA PROTECTION MODELS: ASPECTS OF OWNERSHIP

George Bouchagiar¹ and Maria Bottis²

¹Attorney-at-law, Tutor/Research Fellow on Data Protection/IP, Ionian University, Kalokeretou 14, Corfu, Greece, 49100

²Associate Professor on Information Law, Ionian University, Ioannou Theotoki 72, Corfu, Greece, 49100

ABSTRACT

The potential application of a property rights model to personal data is an old subject of legal discussion. As some scholars have argued, such treatment could strengthen individuals' control over their information or could put pressure on firms to better manage people's data. Given that in the modern era personal data does have some economic value and is treated as a digital asset, this article mentions some examples to detect uncertainties that have, under current norms, occurred. Furthermore, it studies the status of personal data after a person's death and points out some common features that the concept of such data possession and the concept of ownership share. Finally, several shortcomings of the proprietary-based model are discussed to draw particular attention to the sensitive character of our information and its relation with the fundamental principles of human value and dignity.

KEYWORDS

Personal Data, Ownership, Deceased, Property Right

1. INTRODUCTION

Let us assume that George and Stella are friends. They live in different countries and meet once a year. They are about to meet on the 28th of January 2019. George sends an e-mail announcing the date of his arrival and Stella "scrutinizes" their previous e-mails to find places, where they have been before, and pick the best one to go for a drink with George. While scrutinizing their digital conversations, Stella finds a five-years-old e-mail concerning an old bickering. Five years ago, George behaved in an inappropriate manner, but Stella forgave him. But now that she remembers George's previous behavior, she changes her mind. Old digital conversations do not let her forgive George, nor do they let her forget her friend's previous behavior (Mayer-Schönberger 2009).

In present-day societies, digital remembering is the default. However, we need to remember to forget. We may need to think beyond laws and not turn into an unforgiving society. The fact that digital data did not let Stella forget and forgive might not be that dangerous. However, in other occasions, things could be worse.

In 2013, Jetpac collected the content of 150 million images uploaded on Instagram in order to create a digital collection/list of businesses¹. Photographs, taken in e.g. a restaurant, which contained faces with lipstick, were classified by Jetpac's algorithm and, thus, the place was characterized as "dressy". In cases where people in the photograph were only men, the algorithm characterized the places as "gay bars". Many Instagram images had also geo-location information and, hence, by having collected, analyzed and processed all above data, Jetpac could have created a list of all e.g. gay bars of Tehran.

Publishing such a collection would be a service or a disservice?

It would certainly be welcomed by a homosexual citizen or visitor of Tehran, who would not risk asking anyone, and perhaps the wrong person, to find the above gay bars. However, consequences would probably be devastating for Tehran's homosexual community, if this collection fell into the hands of the mullahs.

¹ See <https://techcrunch.com/2013/12/05/social-travel-app-jetpac-ditches-facebook-pivots-to-instagram-based-city-guides-for-at-a-glance-recommendations/>.

Some authors argue that a property rights model could apply to personal data (under the General Data Protection Regulation (GDPR), persona data is any information relating to an identified or identifiable natural person) to strengthen control of the data subject over her private information (Lessig 2006; Liebenau 2016). This model could also enable individuals to share profits from data's exploitation and force firms to make better decisions on the collection and use of such data (Samuelson 2000). Could this model apply to personal data, under current laws and norms?

2. CURRENT EU – US DATA PROTECTION REGIME

Under the European regime, the right to the protection of personal data is protected as a fundamental human right (Article 8(1-2) of the Charter of Fundamental Rights of the European Union, Article 16(1) of the TFEU). It is an established position of jurisprudence that the Convention for the Protection of Human Rights and Fundamental Freedoms does not protect a right to obtain remuneration for the waiver or sacrifice of a fundamental human right; European courts hold that human rights reflect personal integrity and liberty and, hence, one could conclude that there is no space for a property approach (E.C.H.R., *Mellacher v. Austria*, 1989, 12 E.H.R.R. 391; Nadezhda 2009). This means that, in accordance with the above positions, privacy, as a human right inseparable from personhood, cannot be waived or transferred (Prins 2006 at pp. 234-237). However, there have been cases, in which the European Court of Human Rights held that an individual may consent to waiving a fundamental right, albeit has to do so in an explicit manner (*Deweerd/Belgium*, ECHR, 27 February 1980, A35 par. 48-54; *De Wilde, Ooms, Versyp/Belgium*, ECHR, 18 June 1971, A12 par. 65; Lawson 1997).

Contrary to the European regime, the American legislator has adopted a "piecemeal" (Bottis 2012) approach that can be found in several Acts, such as the Fair Credit Reporting Act (1970, 15 U.S.C. § 1681), the Electronic Communications Privacy Act (1986, 18 U.S.C. § 2510), the Cable Communications Policy Act (1984, 47 U.S.C. § 551) or the Video Privacy Act (1988, 18 U.S.C. § 2710). In the United States, privacy (Warren, Brandeis 1890; Cavoukian, Tapscott 1997) is protected under a torts model (Restatement (Second) of Torts §§ 652A-652E, 1977). As some authors have argued, a common law tort could be used not only to enforce accountability on data traders but also to provide remedies for individuals, who have suffered harm to their privacy interests of choice and control (Ludington 2006). Other academics argue that people should own their personal information and be entitled to control what is done with it and that the tort solution might be preferable to a property rights approach, which however would likely offer only modest protection (Litman 2000).

Setting aside EU and US approaches, some economists and scholars argue that personal data could emerge as a new asset class in itself (World Economic Forum 2011; Manovich 2011) and speak of data's economic value, which is indeed measurable (Malgieri, Custers 2017; OECD 2013; Chirita 2016; European Data Protection Supervisor 2014). Such approaches are in accordance not only with some old academic opinions, that view commercialization of personal data as a potential way to strengthen individuals' control over data (Samuelson 2000), but also with some international institutions' views, which have recently recognized that personal data may be provided or used as money in exchange for the supply of digital content² and digital services³.

Given that in present-day society "*tout s' achète*" (Beigbeder 2000; Förster, Weish 2017), in this paper we study the status of personal data after a natural person's death and we highlight some common features that "traditional" ownership (Becker, 1980) and "personal data possession" share. We also discuss several shortcomings of the proprietary-based model to draw particular attention to the sensitive character of personal data and its relation with the fundamental principles of human value and dignity.

² See Article 3(1) of Proposal of the European Commission for a Directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content [(COM(2015)0634 – C8-0394/2015 – 2015/0287(COD)].

³ See Report of the European Parliament on the proposal for a directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content. Available at <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2F%2FEP%2FTEXT%2FREPORT%2BA8-2017-0375%2B0%2BDOC%2BXML%2BV0%2F%2FEN&language=EN>.

3. PERSONAL DATA AFTER DEATH

As we shift from atoms to bits (Mayer-Schönberger, Cukier 2014), digital assets, with economic or personal value, have emerged. For example, a domain name certainly has some financial value, while an e-mail may have emotional value. But let us focus on personal data and examine its status, as a digital asset, after an individual's death.

The European Data Protection Directive⁴ does not mention anything about the data of the deceased and, hence, some European Member States decided to protect it. For instance, under Bulgarian law, in case a natural person dies, her rights of access may be exercised by heirs, who are enabled, amongst others, to request, at any time and free of charge, from the personal data controller a confirmation as to whether or not data relating to the deceased are being processed, information as to the purposes of such processing, the categories of data concerned, and the recipients or categories of recipients to whom the data are disclosed (Article 28(3) of Bulgarian Law For Protection Of Personal Data).

Moreover, under Estonian law (Chapter 2 of the Estonian Personal Data Protection Act), the consent of a data subject shall remain valid during the lifetime of the data subject and for thirty years after the death of the data subject, unless the individual has decided otherwise. Interestingly, in Estonia, after the death of a natural person, processing of personal data relating to the deceased is permitted only with the written consent of the successor, spouse, descendant or ascendant, brother or sister of the data subject.

On the other hand, the General Data Protection Regulation focuses on living persons and does not apply to personal data of the deceased (Member States, however, may provide for rules regarding the processing of personal data of deceased persons, recital 27 of GDPR). By ignoring the deceased, one could argue that the GDPR leaves no space for a proprietary model. However, it should be noted that the GDPR's title refers not only to the protection of natural persons with regard to the processing of personal data, but also to the free movement of such data.

Could essential features of ownership be detected in the European concept of personal data possession?

4. DATA AS PROPERTY: IN WHOSE HANDS?

An essential element of the notion of ownership is the right to possess. This means that a person has the right to exclusive control of the thing (Becker, 1980). In case of intangible items, possession may be understood metaphorically. With regard to personal data, one of the fundamental principles of data protection law is respect for personal autonomy (Bottis 2014). Legal provisions, concerning such data, safeguard rights to informational self-determination (Kang et al. 2012). Thus, it has been consistently supported by authors that the right to the protection of personal data refers to control by the subject over the processing of her data (Oostveen, Irion 2016; Rengel 2014).

Another feature of ownership is the right to use, meaning the right to personal enjoyment of the benefits of the thing, other than those of management and income (Becker 1980). One could claim that a user has the right to use and to enjoy the benefits deriving from her personal data, e.g. e-mail or her IP address or "cookies" (Article 29 Data Protection Working Party, Opinion 4/2007 on the concept of personal data, Jun. 20, 2007; Opinion 1/2008 on data protection issues related to search engines, April 4, 2008). Besides, by using "cookies", companies offer "the best experience" and, thus, consumers enjoy the "benefits of the thing".

Furthermore, the right to manage, to decide how and by whom a thing should be used (Becker 1980), can also be detected in the concept of possession of personal data. Namely, the key tool for a legal control of personal data is the subject's consent (Solove 2013; Article 29 Data Protection Working Party, Opinion 15/2011 on the definition of consent, July 13, 2011) to its processing. Consent of the data subject is defined as any freely given, specific, informed and unambiguous indication of the individual's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her (Article 4(11) of GDPR). Thus, it would be fair to argue that the data subject has the right to manage her information and decide how and by whom her information should be used.

⁴ See Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, *infra* referred to as DPD (for Data Protection Directive).

Another important element of ownership is the right to the income –that is to the benefits deriving from personal use of an item and allowing others to use it (Becker 1980). At this point, things get complex. It is companies, rather than data subjects, that process data, and, thus, have the “right to the income” from this processing. However, some authors argue that consumers should have the right to know the value of their data (Malgieri, Custers 2017). Hence, if such rights were established, users might then be able to participate in profits. In any case, after having provided her data, a consumer does enjoy some kind of benefits, such as the innumerable “free” digital services.

Moreover, while obligations to delete personal data are also provided under Data Protection Directive (e.g. Articles 6(1)(e) and 12(b) of DPD), the GDPR establishes the right to erasure (‘right to be forgotten’, Flaherty 1998; Blanchette 2002; Dodge, Kitchin 2007; Mayer-Schönberger 2009), which is defined as the right of an individual to obtain from the controller the erasure of personal data, concerning him or her without undue delay (Article 17(1) of GDPR). This right resembles the right of the owner to destroy a thing (Becker 1980; Koops 2011).

Another crucial feature of ownership is the right to modify the thing, to effect changes less extensive than annihilation. Indeed, an individual has the right to modify her e-mail address or her name.

In addition, the right to alienate or to abandon ownership reminds us of cases *Deweerd/Belgium* (ECHR, 27 February 1980) and *De Wilde, Ooms, Versyp/Belgium* (ECHR, 18 June 1971), where the European Court of Human Rights held that an individual may consent to waiving a fundamental right.

The need for free flow of personal data illustrates some kind of utilitarian approach and data portability has been strongly supported by the European legislator. Thus, an individual has the right, amongst others, to transmit her personal data to another controller without hindrance from the controller to which these data have been provided (Article 20(1) of GDPR). This resembles the right to transmit, to devise or bequeath the thing (Honoré 1961).

One of the main principles relating to processing of personal data is that the latter shall be processed in a manner that ensures appropriate security of the personal data, including the protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures (Article 5(1)(f) of GDPR). This seems to resemble two features of ownership, i.e. the right to security -that is to immunity from expropriation- and the prohibition of harmful use -one’s duty to forebear from using the item in ways harmful to oneself or others (Becker 1980).

Ownership lasts, in theory, forever (Becker 1980); but the protection of data, not. However, if we examine personal data in a *sui generis* right database, then we could argue that protection of personal data may indeed last for an eternity (Davison 2006; Bottis 2004; Boyle 2003), in favor, however, not of the data subject but, of the database maker.

Another feature of ownership is liability to execution, to having the thing taken away as payment for a debt (Becker 1980). Here, again, it is firms, rather than individuals, which may have data taken away as payment for a debt: in fact, in recent years, some companies had no other choice than to sell their clients’ personal data as a way of silencing creditors (Enos 2001; Prins 2006, at p. 228), while in other occasions, databases that contained personal data changed hands and ownership during strategic company movements (Gauthronet 2001).

Finally, residuary rules govern the reversion to another, if any, of ownership rights, which have expired or been abandoned (Becker 1980). Such rules are, for example, those that determine the disposition of property left by intestate deaths. Such rules could resemble the provisions of the above Bulgarian and Estonian personal data protection laws.

To sum up, one could argue that the notion of ownership and the concept of personal data possession do share common features. However, with regard to certain of these features, personal data, as *quasi* property, has perhaps fallen into the wrong hands.

5. DISCUSSION

The right to the protection of personal data was a “wonderful idea” of an era, when information was rarely shared, societies were local and communications were expensive. In current era, when a pandemonium of information is circulated online, digital societies have no frontiers and communications are inexpensive, we may need to rethink models for protecting such fundamental human rights. A proprietary model might not resolve all the problems, albeit the treatment of the right to personal data as a *quasi* property right could remind people that an individual is the owner of –and the one entitled to control– her data and that she has the right and the human need to forget.

As a possible disadvantage of our approach here, one could note that a personal and a property model derive from dissimilar sources and concepts, which perhaps should not blend. A *quasi* property right in our case should not signal that a human being is equal to an object, since anything that is personal to us should normally not be subject to ownership. If such particular distinction is not as clear as it can be in law, then human dignity, the constitutional and fundamental principle upon which the right to the protection of personal data is indeed based, could be threatened. Moreover, the unambiguous identifier of our dignity, our self-determination, should always play a key role in these potential approaches of privacy, given that it is the feature that enables the individual to freely make decisions about her life.

The main GDPR’s objective, as we see in its title, is to strike a fair balance between the protection of individuals with regard to the processing of their personal data and the free movement of such data (e.g. recital (10) of GDPR). We do observe, on the other hand, the European legislator’s wish to treat individuals’ information as a highly protected moral good, incapable of being subject to ownership or any kind of measurement. This illustrates the direction, towards which the EU regime is heading, which is, of course, different than a pure ownership approach that would regard data as a mere object. However, we stress the GDPR’s provisions on data portability and free flow of data in general. These provisions strongly support, to us, a *quasi* property right, if the appropriate attention is also drawn to human dignity and self-determination in the interpretation of all GDPR provisions.

REFERENCES

- Becker C. Lawrence, 1980. The Moral Basis of Property Rights, in *Nomos XXII: Property*. J. Roland Pennock and John W. Chapman (eds). New York, New York University Press, Hollins Digital Commons, Web, pp. 187-220, at pp. 190-191. Available at <https://digitalcommons.hollins.edu/cgi/viewcontent.cgi?article=1020&context=philfac>.
- Beigbeder Frédéric, 2000. *99 Francs*, Grasset & Fasquelle, Paris, pp. 1-282, at p. 15.
- Blanchette J. F., Johnson D.G., 2002. Data Retention and the Panoptic Society: The Social benefits of Forgetfulness, *The Information Society*, 18:33, Taylor & Francis, pp. 33-45. Available at <https://pdfs.semanticscholar.org/f17f/fed4d7cd491796e7384f66204a11010ab0b0.pdf>.
- Bottis Maria, 2014. Law and Information: a “Love-Hate” Relationship, in Maria Bottis, ed., *The History of Information, from Papyrus to the Electronic Document*, Nomiki Vivliothiki, Greece, at p. 148.
- Bottis Maria, 2012. Surveillance, data protection and libraries in Europe and the US-notes on an empirical data case study on surveillance and Greek academic libraries, in Maria Bottis (ed.), *Privacy and Surveillance, Current Aspects and Future Perspectives*. Nomiki Bibliothiki, Greece, pp. 272-282, at pp. 275-278.
- Bottis Maria, 2004. *The Law of Information*, Nomiki Bibliothiki, Greece, p. 69 (In Greek).
- Boyle James, 2003. Foreword: The Opposite of Property? *Law and Contemporary Problems*, Volume 66 (Winter 2003), pp. 1-32, at p. 25.
- Cavoukian A., Tapscott D., 1997. *Who Knows: Safeguarding Your Privacy in a Networked World*, McGraw-Hill, New York.
- Chirita Anca D., 2016. *The Rise of Big Data and the Loss of Privacy, Personal Data in Competition, Consumer Protection and IP Law - Towards a Holistic Approach?* Bakhoun, M., Gallego Conde, B., Mackenordt, M.-O. & Surblyte, G. Berlin Heidelberg (eds), Springer, 2018 (forthcoming), Durham Law School Research Paper, p. 11. Available at <https://ssrn.com/abstract=2795992>.
- Davison Mark, 2006. Database Protection: The Commodification of Information, in *The Future of the Public Domain: Identifying the Commons in Information Law*, Lucie Guibault and P. Bernt Hugenholtz (eds), Kluwer Law International, The Netherlands. pp. 169, 187.

- Dodge M., Kitchin R., 2007. Outlines of a World Coming into Existence: Pervasive Computing and the Ethics of Forgetting, *Environment and Planning B: Planning and Design*, vol. 34, pp. 431-445. Available at http://personalpages.manchester.ac.uk/staff/m.dodge/cv_files/epb_ethics_of_forgetting.pdf.
- Enos L., 2001. Deal Afoot to Destroy Toysmart Database, *E-commerce Times*, January 10, 2001. Available at <https://www.ecommercetimes.com/story/6607.html>.
- European Data Protection Supervisor, 2014. Preliminary Opinion, *Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy*, March 2014, p. 9.
- Flaherty DH, 1998. Controlling Surveillance: Can Privacy Protection Be Made Effective? in PE Agre and M Rotenberg (eds), *Technology and Privacy: The New Landscape*, Cambridge, MA/London: The MIT Press, 1998, pp. 167-192, at p. 172.
- Förster Kati, Weish Ulli, 2017. Advertising Critique: Themes, Actors and Challenges in a Digital Age, in Gabriele Siegert, M. Bjørn von Rimscha, Stephanie Grubenmann (eds), *Commercial Communication In The Digital Age, Information or Disinformation?*, Walter de Gruyter GmbH, 2017, pp. 15-35, at p. 23. Available at <http://www.oapen.org/search?identifier=628402>.
- Gauthronet S., 2001. The Future of Personal Data in the Framework of Company Reorganisations, *23rd International Conference of Data Protection Commissioners*, Paris, September 2001.
- Honoré A. M., 1961. Ownership, in A. G. Guest (ed), *Oxford essays in jurisprudence, a collaborative work*, London: Oxford University Press, pp. 107-147.
- Kang Jerry et al., 2012. Self-Surveillance Privacy, *Iowa Law Review*, vol. 97:809, p. 820.
- Koops Bert-Jaap, 2011. *Forgetting Footprints, Shunning Shadows: A Critical Analysis of the 'Right to Be Forgotten' in Big Data Practice*. SCRIPTed, Vol. 8, No. 3, Tilburg Law School Research Paper No. 08/2012, pp. 229-256, at p. 247. Available at SSRN: <https://ssrn.com/abstract=1986719> or <http://dx.doi.org/10.2139/ssrn.1986719>.
- Lawson R.A., 1997. H.G. Schermers (eds), *Leading Cases of the European Court of Human Rights*, Nijmegen, Ars Aequi Libri, pp. 637-638.
- Lessig Lawrence, 2006. *Code, Version 2.0*. Basic Books, pp. 276, 228-230.
- Liebenau Diana, 2016. What Intellectual Property Can Learn From Informational Privacy, and vice versa. *Harvard Journal of Law & Technology*, Volume 30, Number 1 Fall 2016, pp. 285-307.
- Litman Jessica, 2000. Information Privacy/Information Property, *Stanford Law Review*, Vol. 52, 2000, pp. 1283-1313.
- Ludington Sarah, 2006. Reining in the Data Traders: A Tort for the Misuse of Personal Information, *Maryland Law Review*, Vol. 66, 2006, pp. 140-193.
- Malgieri Gianclaudio, Custers Bart, 2017. Pricing Privacy – The Right To Know the Value of Your Personal Data, *Computer Law & Security Review: The International Journal of Technology Law and Practice*, doi: 10.1016/j.clsr.2017.08.006.
- Manovich, L., 2011. *Trending: The Promises and the Challenges of Big Social Data, Debates in the Digital Humanities*, M. K. Gold (ed), The University of Minnesota Press.
- Mayer-Schönberger Viktor, Cukier Kenneth, 2014. *Big Data: A Revolution that Will Transform How We Live, Work, and Think*, Mariner Books, Houghton Mifflin Harcourt, Boston, New York. p. 78.
- Mayer-Schönberger Viktor, 2009. *Delete: The Virtue of Forgetting in the Digital Age*, Princeton: Princeton University Press.
- Nadezhda Purtova, 2009. Property in Personal Data: A European Perspective on the Instrumentalist Theory of Propertisation. *European Journal of Legal Studies*, Vol. 2, No. 3, 2010, pp. 193-208, at p. 203. Available at https://ejls.eui.eu/wp-content/uploads/sites/32/pdfs/The_Future_of_Autumn_Winter_2010/Full_The_Future_of_Autumn_Winter_2010.pdf
- OECD, 2013. *Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value*, OECD Digital Economy Papers, No. 220, OECD Publishing, Paris.
- Oostveen Manon, Irion Kristina, 2016. *The Golden Age of Personal Data: How to Regulate an Enabling Fundamental Right?* University of Amsterdam, Institute for Information Law, Paper No. 2016-06.
- Prins Corien, 2006. Property and Privacy: European Perspectives and the Commodification of our Identity, in L. Guibault and P.B. Hugenholtz (eds), *The Future of the Public Domain, Identifying the Commons in Information Law*, Kluwer Law International, The Netherlands, pp. 223-258.

- Rengel Alexandra, 2014. Privacy as an International Human Right and the Right to Obscurity in Cyberspace, *Groningen Journal of International Law*, vol. 2(2), Privacy in International Law, pp. 33-54.
- Samuelson Pamela, 2000. Privacy as Intellectual Property? *Stanford Law Review*, vol. 52, 2000, pp. 1125-1173, at pp. 1127-1129. Available at <https://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=3137&context=facpubs>.
- Solove J. Daniel, 2013. Introduction: Privacy Self-Management and the Consent Dilemma, *Harvard Law Review*, vol. 126:1880, p. 1894.
- Warren S. & Brandeis L., 1890. The Right to Privacy, *Harvard Law Review*, vol. 4, ed. 5, pp. 193-220. Available at <http://www.cs.cornell.edu/~shmat/courses/cs5436/warren-brandeis.pdf>.
- World Economic Forum 2011. *Personal Data: The Emergence of a New Asset Class*, p. 5. Available at http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf.

VISUALISATION – IT IS NOT THE DATA, IT IS WHAT YOU DO WITH IT

Chris Roast¹, David Patterson² and Victoria Hardman²

¹*Communication and Computing Research Centre, Sheffield Hallam University, Sheffield, S1 1WB, United Kingdom*

²*Learning Light Ltd., Sheffield Technology Park, Cooper Building, Arundel Street, Sheffield S1 2NS, United Kingdom*

ABSTRACT

The research described is focused upon the supporting the widespread adoption of data analysis skills as a means of empowering citizen engagement with the ever growing amount of open data. This has the potential to enhance the quality of engagement and effective use of data by helping minimize wrongly computed conclusions from given data. The work is applied in two technology areas where unskilled engagement with data is common. The first is that of the spreadsheet as a tool by which many analyze, plan and develop basic models for a number of purposes ranging from domestic budgeting to business planning. The second is more specialized language by which large data resources are commonly managed - the standard query language (SQL) for databases. The paper justifies the selection of these two areas and goes on to describe the approach taken to helping reducing potential user errors when working with both. This involves supporting users through visualizing the structures by which data is processed and characteristics of the steps involved in processing. This can be contrasted with the more common approaches to data visualisation that focus more up the outcome of data once processed. The two visualisation approaches are illustrated and initial results from the evaluation are described. Experience with using and evaluating these approaches has informed our understanding of the approach to visualization design.

KEYWORDS

Spreadsheets, SQL, Visualization, Empowerment

1. INTRODUCTION

This research is focused upon the supporting the widespread adoption of data analysis skills as a means of empowering citizen engagement with the ever growing amount of data. The objective is to help enhance the effectiveness of engagement with data by helping minimize wrongly derived conclusions. Although large, and growing, amounts of data are now available for scrutiny and analysis, the outcome of any analysis is dependent upon processing performed on the data. So for a member of the public to, say, assess air pollution data within their town, a level of numeracy needs to be applied in the processing of such data in order reach defendable conclusions. Similarly, for a town councilor to assess the impact of, say, reducing the budget for school bus services, they either defer the computation to a department or estimate it themselves with a numeric model. We believe the common trend with democratization of data and data transparency (and access) is that the public, government officials and business managers are more reliant upon their own skills to analyze and comprehend data than official units established for that purpose.

With this perspective in mind we have researched and developed tools to support non-specialist data analysis. We take non-specialists to be those whose job is specifically not data analysis but who enhance their work or activities through conducting analysis. The two tools focus that are dominant in many relevant sectors: spreadsheets and the standard query language (SQL) for database access.

Spreadsheets have been our initial interest simply because of their dominant role in informal modelling and basic numeric processing. Up until recently, spreadsheets have been part and parcel of school level ICT education, they are core tools in many vocational educational offerings and are a common general purpose tool for many numerate subjects (Chambers and Scalfidi 2010). In addition, spreadsheets are widely used in most businesses for unsophisticated model testing through to complex data analysis. Their ease of adoption means that in many cases their use in businesses persists to the point of being central to specific business

functions. Their responsiveness means sheets and their users quickly become entangled in 'solutions' that persist. Potentially reflecting this entanglement, Panko (2008) and Hendry and Green (1993) argue that up to 44% of spreadsheets contain errors, with specific areas of concern including the complexity of inter-cell referencing, the understanding of formulae.

SQL is the standard language for managing databases as well as accessing and processing data from them. As a language, it is not as accessible and commonplace as spreadsheets, though once accessing stored data it is equivalent in its prevalence. SQL forms an explicit component of over 80% higher education computing courses, with SQL expertise attracting some the highest pay rates of all IT professionals. Despite that, there is a recognised skills shortage in database related skills (Sector Insights, 2015), along with a global shortfall of Big Data expertise (McKinsey Global Institute, 2011). In addition, to the professional demand in the ICT sector, there is a growing interest in less skilled access to data for which SQL is core stepping stone to processing and analyzing large datasets.

So the problem we wish to address is that of the relative non-specialist using computationally powerful tools for data analysis. The risk in both cases, but to different degrees, is that the computation performed may not in fact be the one intended. Below we use two examples: in spreadsheet arithmetic "A1 + 2 * C1" may be written, and computed as: A1 + (2 * C1), when in fact the user thinks they are computing: (A1 + 2) * C1. Similarly, in SQL a user may query data using one a variety of variants of database JOIN and not understand which is the one they really mean - and in these circumstances checking correctness is relatively complex.

In the rest of this paper, we introduce our approach to visualization focused upon to address these challenges. We then, briefly describe the tools in development and their evaluation. We conclude with a reflection to applying our experience in other contexts where end user activities when working with data require support.

2. THE HUMAN IN THE LOOP AND VISUALISATION

Our approach to supporting users in working with data is one drawn from human factors and human computer interaction. Working with large data sets can be broadly characterized as "complex interaction" (Roast and Uruchurtu, 2013) since users are unlikely to fully appreciate the characteristics and features of the primary object they are interacting with (in this case a data set, or the language by which it is interrogated). So unlike common direct manipulation interfaces, users are faced with developing their understanding via broadly general purpose languages for probing a data set. In the two example technologies in this paper, the language of spreadsheet formula (and their re-use) is one such language and SQL is the other. Hence, interaction from this perspective is *indirect manipulation*, mediated by the language involved. Blackwell et al. (2001), have explored some of the characteristics of such interaction paradigms under the conceptual framework of Cognitive Dimensions of Notation. Moody (2009) has taken a slightly alternative approach structuring cognitive psychology findings in a framework to govern the effective design of visual notations. With both perspectives, the synthesis of a visual notation to support interaction is a non-trivial activity, which in our experience requires significant user feedback moderated by the rationale and considerations promoted by Moody (2009) and Blackwell et al. (2001), see Roast and Uruchurtu (2016).

The aims of a visual language is to employ geometric forms, colours, layout, etc. to take advantage of human perceptual ability to recognise patterns and associations - and support "visual thinking". Examples of how this might reveal itself include: learners recognising when a language element is not fit for its intended purpose and identifying what modifications are able to remedy that. For example, with a spreadsheet a user may be helped in: (i) recognising when a cell formula is computing an unexpected result (Burnett et al, 2002); (ii) identify where the mistake lies in the formula, and (iii) take some action to change the mistake. When considering data sets, this is intrinsically a cyclic process simply because (iii) may not correct the problem and (i) is dependent upon a clear expectation of the required result (something novice users may not have). The easiest area to address in terms of tool support in this simple account of interaction is that of (ii) - support in terms of *identifying where the mistake lies*.

The approach taken is captured by the following high-level principles for visualising languages:

- **Evidencing structure.** Within a given language or notation, the syntactic structure that defines its computed meaning is clearly evidenced in the visualisation.
- **Visual mapping.** Mapping between the original notation and the visualisation is well supported.

- **Evidencing categories.** Within a given expression when visualised it is easy to recognise the different categories of tokens and structures relevant to understanding the expression.
- **Evidencing abstractions.** There are various abstractions apparent in the way expressions are used (such as the re-use of a common sub-expression). Where possible, these should be evident in the visualisation.
- **Evidencing computation.** The base "meaning" of an expression can be interpreted as simply the result it generates when used. While this can belittle the importance or impact of a particular result, showing how an answer is generated helps with *identifying where the mistake lies* (when there is one).
- **Visual simplicity and scalability.** This general principle is to discourage wasted space in a visualisation approach and by contrast discourage dense unnecessarily complex visual structures. This is not necessarily straightforward when we consider large data sets and sophisticated analyses since there will be cases that are unavoidably complex.

It should be clear that these are just principles and thus subject to interpretation, and incapable of directing a visualisation designer to a single optimal solution. Experience of designing and evaluating visualisations indicates that the key features when applying the above principles are best established through users' engagement and feedback (Roast and Uruchurtu 2016). There are competing rationale for design alternatives, whether the user understands the same notions of "categories", "abstractions" and computation" as a visualisation pre-supposes, and whether they need that understanding to address the above three stages.

3. SPREADSHEET VISUALISATION

Research into addressing issues of spreadsheet quality has motivated many enhancements. This has included additional features to ensure they are more transparent and to encourage more discipline in their use (Burnett et al. 2002, Hendry and Green 1994, Hermans and Dig 2014, Panko and Sprague 1998, Sajaniemi 2000, Dix et al. 2016). Previous work has proposed ways of presenting and visualizing spreadsheets (Saariluoma and Sajaniemi 1991, Igarashi, et al. 1998, Ballinger, et al. 2003, Burnett, et al. 2001). However, these works only consider the wider structure of spreadsheets, and the dependencies between cells. None appear to have addressed the fact that the formulas language is computationally powerful but normally contracted onto a single line. It is this complexity of language presentation that can complicate its effective use, when a visual language could present a graph to make the relationship and sequence of formulae elements more evident. This should provide a visual 'scaffolding' that reduces problems of understanding what is being done with data, since it will no longer be something the user has to determine on their own.

The visualisation was developed extensively exploring various designs and allowing the authors, tutors and learners to provide feedback on characteristics and their perceived appropriateness. Initially, good visual design practice was informed by learning scenarios and educational uses of spreadsheets (Leitão and Roast, 2014, Roast, et al. 2016). The resulting tool is called EQUS, figure 1 provides two examples of the visualization. Key points arising from EQUS include: the importance of representing intermediate values; the visually distinct status of the final result; the colour coding of different arguments. One of the most challenging elements of the EQUS design was the visualisation for conditional expressions, such as: $\text{IF}(A1 < 10, B1/2, C1*2)$ (meaning that if A1 is less than ten the value is halved, otherwise it is doubled). In all cases only $B1/2$ is computed or $C1*2$ is computed (never both). Hence, following our principles, the visualisation would alter depending upon the specific data use. Thus, a user would have to check both conditions - to see the potential computation. Our approach in this case is illustrated in figure 2, both conditions are shown but only the valid one for the current data is shown as computed.

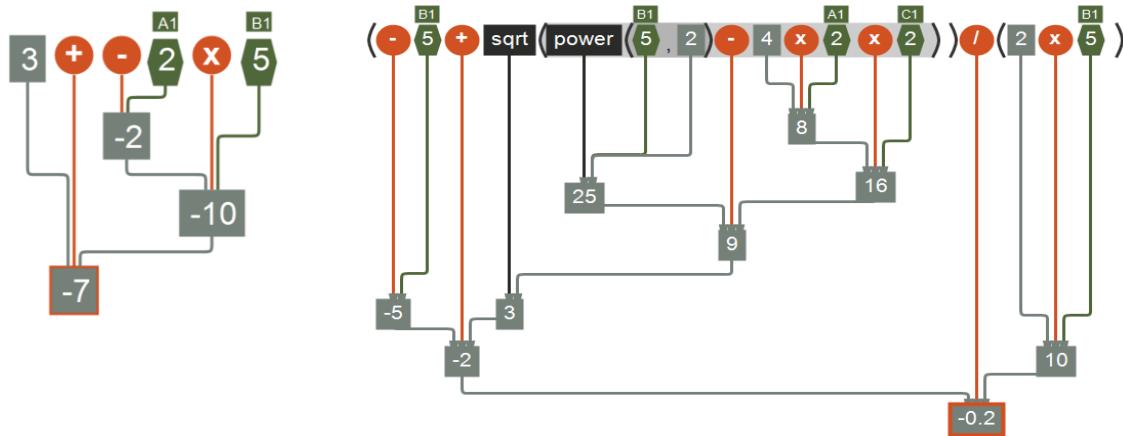


Figure 1. Two EQUS examples the visualisation output, for formula: $= 2 + -A1 * B1$ and $= (-B1+sqrt(power(B1,2)-4*A1*C1))/(2*B1)$.
In these cases: cell A1 is 2, B1 is 5 and C1 is 2

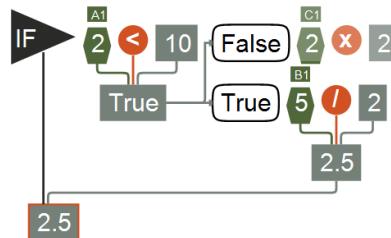


Figure 2. An EQUS example of a conditional statement

4. SQL VIEW

Historically our development of a visualisation for SQL has followed the design and evaluation of spreadsheets. Hence, the experience of some of the issues arising has shaped our approach with SQL. Specifically, since SQL commonly queries are applied to dynamic and large data sets, visualising a single representative result was a challenge. Another factor relevant to this research was that since SQL is a more specialist tool than spreadsheets, there is a mass of SQL specific expertise embodied in training and education resources, reference materials and on-line fora (such as: www.w3schools.com and www.stackoverflow.com). In addition, it can be argued that the technical capabilities and expectations of SQL exceed what is normal for spreadsheets. Despite this, Prior (2003) provides a simple account of how queries are developed that is close to what we have observed with spreadsheets:

"Professionals verify the results of this preliminary query once it has been executed online, and if it does not accurately return the required information, they refine the SQL query and re-execute it, repeating the verification and refinement steps until they are satisfied that the query is returning the desired results."

A premise for the visualisation was informed by our first design principle that, in effect, proposes that the relevant structure of the language used should be conveyed clearly to the user. In our initial research, we found that some presentation conventions were used in SQL query text, though with little consideration of the underlying syntax. In particular, few educational resources provided clear accounts of syntax and most assumed that it could be understood purely through examples. Redressing this concern our initial visualisation "structure" was chosen to use the parse-tree of SQL select statements. An early prototype of a select statement is shown in figure 2. This breakdown conveys some of the computational structure captured by a statement and thus supports the user in identifying the localise activity that they may want to examine.

In figure 3, the precedence structure of operators in the SQL where-condition is evident in the graphical composition.

Following a review of common training and teaching resources for SQL, and in consultation with high-education experts, it was concluded that the dominant pedagogic approach was one of using existing data sets to demonstrate SQL behaviour. Unfortunately, this provides a set of positive examples re-enforcing expectations and not exposing the full extent of what SQL will and won't do with arbitrary data sets. We combined this insight with a review of common or significant user errors within SQL based on a number of on-line technical support fora.

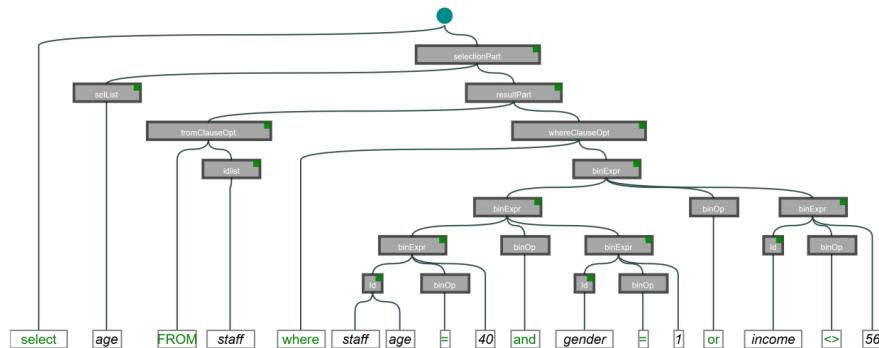


Figure 3. An illustration of a prototype parse tree for an SQL statement

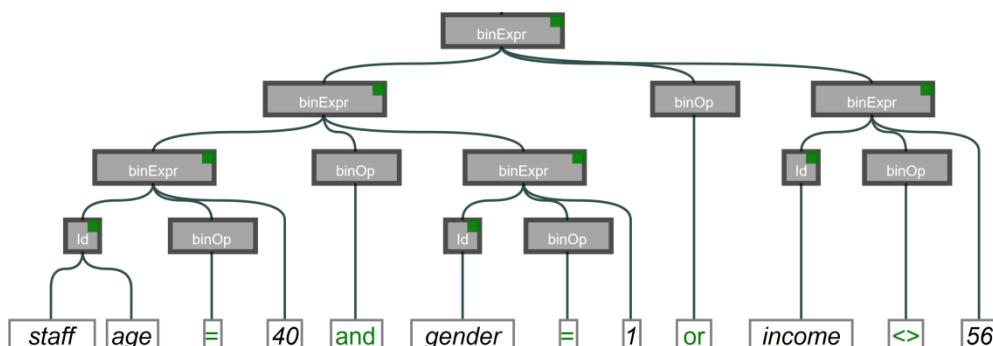


Figure 4. A close-up of the parse tree for an SQL expression that could be misinterpreted

Figure 5. Prototype BRIQs representing the generic result that can associate with specific parts of SQL queries: an equality, and two types of join

The visualisation approach developed in keeping with these findings led to the demand to adequately visualise the concept of the generic outcome of a SQL statement. SQL results are on-the-whole either specific values: 10, 20, "Jose", etc. (including a NULL the non-value) or a table of data. In contrast to existing materials, a generic presentation of an output was developed - termed a "BRIQ". The aim was to ensure the visual format for the BRIQ still conveyed the structure and features of a table and its features. Thus BRIQ's were designed to follow a grid layout with table and column names, and the rows capturing what can be inferred about table content. The BRIQ design was subsequently developed to:

- *Generalize results.* Hence, using a variable notation to represent arbitrary values.
- *Provide localised results.* Relevant BRIQs are interactively revealed on each relevant parse node (see those with a upper right "handle" in figures 3 and 4).
- *Ensure NULLs are explicit.*
- *Display rejected results.* This was principally motivated by the desire to demonstrate when NULL values were excluded or present in a computed result. (See the lower shaded sections of the BRIQs in figure 5.)

Developing a general BRIQ will always be a challenge because of the complexity of conditions and constraints the SQL is capable of computing. Early prototype BRIQs for specific example statements are shown in figure 5.

5. EVALUATION AND FEEDBACK

5.1 Spreadsheets

As described earlier we are treating the design of a visualisation for supporting end users as an iterative process governed by the core elements of the design principles introduced in section 2. As such, the process is one of discovering and reconciling the concepts held by the user with those essential to understanding the computation that is performed.

Historically more progress has been made with user based evaluation and refinement with the spreadsheet tool. Initial spreadsheet designs highlighted a tension between keeping the visualization succinct compared with displaying detail. Initial evaluation activities using pen and paper illustrations were used to establish the most effective approach. Following that, the spreadsheet visualization approach was formatively evaluated principally in educational contexts, gathering primarily qualitative feedback along with some performance measures.

Reliable performance data within an educational setting proved to be hard to obtain due to the difficulty of controlling experimental conditions and assessing impact longitudinally. Specifically, direct objective benefits were envisaged to cover specific hard cases "finding errors" or "improving comprehension" at a variety of educational levels. However, none of the performance data gathered (over 40 users at degree level) evidenced a negative influence.

Qualitative feedback was strongly positive. Over 100 learners at all levels including: apprentices, diploma students and undergraduates were engaged in trying out the tool - feedback included:

- "It would help me a lot with other formulas"*
"You can see the values and how they are worked out, that's great."
"It would help anyone willing to learn about spreadsheets"

In addition, tutors engaged during the sessions were positive:

- "I am sure that it could add value to the teaching of mathematics."*
"I think it would be very helpful"
"Absolutely brilliant when it comes to more complicated formulas for our learners. With regards to the IF statement, I particularly like the way it checks the condition and identifies whether it is TRUE or FALSE. Additionally, really good for formulas of non-adjacent cells."

Despite the lack of familiarity with the visualisation its presence and use there was no evidence of it distracting and thus impairing learners. In follow-on interviews all agreed that the visualisation approach had merit. Overall feedback was positive, with those interviewed seeing the potential to help "de-mystify" spreadsheets. Including trainee tutors themselves not confident with the technology:

"I struggle a lot with spreadsheets and find it hard to understand them. Seeing the spreadsheet visualisation prototype made it clearer to understand the formulas and feel that if I had chance to use a programme of that kind I would have a greater understanding and be able to pick up the skills I require much quicker. I feel that this product could help people like myself that struggle with spreadsheets."

An expert in mathematics education research commented:

"It will be very useful to many students to have a product that enables a better conceptual understanding of the equation format. There is a clear need for such a tool to be suitable for the many students who do not have high levels of mathematical skills and yet use mathematical symbolism every day in their studies. This will include students from Chemistry, Business, Economics, Psychology, Geography and many more."

The EQUIS tool is currently being developed as a robust commercially viable Excel plugin.

5.2 SQL

The SQL visualisation evaluation has been less extensive to date. Feedback has been gained from educators and professionals. The rationale and working examples of the tool have been received positively with formative feedback regarding the likely scale and complexity of leading examples. Compared with spreadsheets, SQL demands a more managed introduction and support mechanism and this is reflected in our visualisation that requires secondary support to ensure that end users understand how to interpret it. Although ideally we would like to have avoided that, both the power language and extent of its outputs (arbitrarily complex tables) do suggest a non-trivial introduction may be needed.

6. CONCLUSIONS

This paper has argued for the importance of non-expert data analysts who employ commonplace tools. In particular, the credibility and efficacy of non-expert analysis is supported by helping reduce the risk of errors in the notations and languages that get used. Our approach to this problem has been to develop visualisations for the language of data analysis (as opposed to the data itself). We would argue that this provides important support for the growing democratization of an ever-widening range of available data sources. The visualisation approach is described and illustrated for two tools, user evaluation and feedback for both are described.

Future work will involve gaining more feedback on our SQL visualization and use that experience to refine and strengthen our approach. In addition, refining our approach through applying it to other contemporary data analysis tools and languages.

ACKNOWLEDGEMENT

The work reported is in part supported by Innovate UK (project "SQL View") and in collaboration with Very Viz, Ltd. We are indebted to our industrial and educational participants in evaluation and feedback activities.

REFERENCES

- Ballinger, D., Biddle, R. and Noble, J., 2003. Spreadsheet structure inspection using low level access and visualisation. In Proc. Fourth Australasian User Interface Conference (AUIC2003). Biddle, R. and Thomas, B., Eds. ACS. 91-94.
- Blackwell, A.F., Britton, C., Cox, A. Green, T.R.G., Gurr, C.A., Kadoda, G.F., Kutar, M., Loomes, M., Nehaniv, C.L., Petre, M., Roast, C., Roes, C., Wong, A. and Young, R.M., 2001. Cognitive Dimensions of Notations: Design tools for cognitive technology. In M. Beynon, C.L. Nehaniv, and K. Dautenhahn (Eds.) Cognitive Technology 2001 (LNCS 2117). Springer-Verlag, pp. 325-341.

- Burnett, M., Atwood, J., Djang, R., Gottfried, H., Reichwein, J. and Yang, S., 2001. Forms/3: A first-order visual language to explore the boundaries of the spreadsheet paradigm. *Journal of functional programming* 11(2): 155-206.
- Burnett, M., Sheretov, A., Ren, A. and Rothermel, G., 2002. Testing homogeneous spreadsheet grids with the "what you see is what you test" methodology," *IEEE Trans. Softw. Eng.*, 28(6) pp. 576-594. <http://dx.doi.org/10.1109/TSE.2002.1010060>
- Campbell-Kelly, M., 2007. Number crunching without programming: The evolution of spreadsheet usability. *IEEE Annals Of The History Of Computing*, 29(3), pp.6-19, ISSN: 1058-6180
- Chambers, C. and C. Scaffidi (2010). Struggling to excel: A field study of challenges faced by spreadsheet users. *Visual Languages and Human-Centric Computing (VL/HCC)*, 2010 IEEE Symposium on, IEEE.
- Dix, A., Cowgill, R., Bashford, C., McVeigh, S. and Ridgewell, R. (2016). Spreadsheets as User Interfaces. *Proc. AVI2016*, ACM, pp.192-195. doi: 10.1145/2909132.2909271
- Davis, F.D., Bagozzi, R.P. and Warshaw P.R., 1989. User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35 (8) pp. 982–1003
- Gretton, H. and Challis, N., 2008. "Fundamental Engineering Mathematics; A Student Friendly Workbook" Woodhead Publishing, 2008, ISBN-13: 978-1898563655.
- Hendry, D.G. and Green, T.R.G., 1994. "Creating, comprehending and explaining spreadsheets: a cognitive interpretation of what discretionary users think of the spreadsheet model." in *International Journal of Human-Computer Studies*, 40(6), 1033-1065, DOI=10.1006/ijhc.1994.1047
- Igarashi, T., Mackinlay, J.D., Chang, B-W and Zellweger, P.T., 1998. Fluid Visualization of Spreadsheet Structures, *Proceedings of the IEEE Symposium on Visual Languages*, p.118.
- Leitão, R. and Roast, C., 2014. Developing visualisations for spreadsheet formulae: towards increasing the accessibility of science, technology, engineering and maths subjects. In: 9th Workshop on Mathematical User Interfaces, Coimbra, Portugal, 10 July 2014.
- Moody, D.L., 2009. The "Physics" of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering, *IEEE Transactions on Software Engineering*, 2009. 35(6): pp. 756-779.
- Myers, B. A. (1986). Visual programming, programming by example, and program visualization: a taxonomy. *ACM SIGCHI Bulletin* 17(4): 59-66.
- Panko, R. R., 2008. What We Know About Spreadsheet Errors. *Journal of End User Computing's Special issue on Scaling Up End User Development* 10, 15-21
- Panko, R. R. and Sprague Jr., R. H., 1998. "Hitting the wall: errors in developing and code inspecting a simple spreadsheet model" *Decision Support Systems* 22(4): 337-353.
- Peyton Jones, S., Blackwell, A. and Burnett, M., 2003. A user-centred approach to functions in Excel. In *Proceedings of the eighth ACM SIGPLAN international conference on Functional programming (ICFP '03)*. ACM, New York, NY, USA, 165-176. DOI=10.1145/944705.944721
- Prior, J. C., 2003. Online assessment of SQL query formulation skills. In *Proceedings of the fifth Australasian conference on Computing education - Volume 20 (ACE '03)*, Tony Greening and Raymond Lister (Eds.), Vol. 20. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 247-256.
- Rello, L., Kanvinde, G. and Baeza-Yates, R., 2012. Layout guidelines for web text and a web service to improve accessibility for dyslexics. *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*, ACM.
- Roast, C., Leitao, R. and Gunning, M., 2016. Visualising formula structures to support exploratory modelling, *International Conference on Computer Supported Education*, Rome, Italy, 21-23 April, 2016.
- Saariluoma, P., Sajaniemi, J., 1991. Extracting Implicit Tree Structures in Spreadsheet Calculation. *Ergonomics* 34(8), 1027-1046.
- Roast, C. and Uruchurtu, E. (2013). Formative evaluation for complex interactive systems. In: COLLAZOS, Cesar, LIBORIO, Andreia and Rusu, C., (eds.) *Human Computer Interaction : 6th Latin American Conference, CLIHC 2013, Proceedings. Information Systems and Applications, incl. Internet/Web, and HCI (8278)*. Springer Verlag, 47-54.
- Roast, C. and Uruchurtu, E. (2016). Reflecting on the Physics of Notations applied to a visualisation case study. In: CORNEJO, Raymondo and Mayora Ibarra, Oscar, (eds.) *MexIHC '16 : 6th Mexican Conference on Human-Computer Interaction*. ACM.
- Sorva, J., Karavirta, V. and Malmi, L., 2013. A Review of Generic Program Visualization Systems for Introductory Programming Education. *Trans. Comput. Educ.* 13, 4. DOI=<http://dx.doi.org/10.1145/2490822>
- Stein, J. and V. Walsh, 1997. "To see but not to read; the magnocellular theory of dyslexia." *Trends in neurosciences* 20(4): 147-152.
- Technology Insights 2012, e-skills UK (Available at <http://www.e-skills.com>).
- Sector insights: skills and performance challenges in the digital and creative sector 2015, Evidence Report 92, June 2015, ISBN 978-1-908418-81-4 © UKCES 1st Ed/06.15
- McKinsey Global Institute 2011, Big Data: the next frontier for innovation, competition and productivity. ©McKinsey & Company.

CORPORATE ATTITUDES TOWARDS DATA PHILANTHROPY IN SOUTH AFRICAN BANKING, RETAIL AND TELECOMMUNICATIONS

Kungela Mzuku and Jean-Paul Van Belle

University of Cape Town, Private Bag, Rondebosch, 7701, South Africa

ABSTRACT

This paper details an exploratory, qualitative study of the willingness, preparedness and concerns around the emergent phenomenon of Data Philanthropy by South African corporate organisations. The study examines the attitudes of corporations in respect of using data and associated technologies for social good, their chosen level of engagement in corporate social responsibility, and their willingness to change their current engagement to one that requires more efforts and tools to implement. The main research objective is to investigate the corporate attitudes towards Data Philanthropy in the South African banking, retail and telecommunications industries. The study includes a review which exposes gaps in literature that the research analysis and discussion will aim to uncover. The study used an inductive approach using grounded theory. Four main organising themes are used to analyse the data which was collected through interviews namely Corporate Social Responsibility, Willingness, Preparedness and Concerns. A remarkable diversity in detailed sub-themes was uncovered for each of the main themes.

KEYWORDS

Data Philanthropy, Corporate Social Responsibility, Open Corporate Data, South Africa, Data for Development

1. INTRODUCTION

Enterprises have access to vast amounts of data which is collected from customers on a day-to-day basis that can be useful to address problems that societies face today (Coulton, George, Putnam-Hornstein, & de Haan, 2015). Sharing this data, and sometimes also their own technical and human resources, with selected academic researchers and non-profit organisations for the purpose of data analysis in aid of social development, is known as data philanthropy. Data Philanthropy is a new form of Corporate Social Responsibility (Soldner, 2016). It was brought into light by the United Nations Global Pulse to combat the United Nations Sustainable Development Goals (Global Pulse, 2015).

The main research objective of this research paper is *to explore the corporates attitudes that corporations in South Africa have towards Data Philanthropy*. It will look at the level of willingness, preparedness as well as highlight the concerns that corporations have with regards to implementing Data Philanthropy. A sub-objective of this research is to explore *the level of involvement that corporations are willing to take on for social good*.

This research paper is significant in the field of information technology as it will help raise awareness about Data Philanthropy as a new phenomenon in South Africa. Apart from awareness, this research brings into light how corporations can use their data and resources to help South African communities in a more efficient way. It was revealed recently that more than 50% of South African citizens are living in poverty (Statistics South Africa, 2017). Ending poverty is one of the United Nations Sustainable Development goals for 2020 thus new methods, such as collaboration from various stakeholders of the country including but not limited to research institutions as well as private corporations, is needed to find new innovative solutions.

2. LITERATURE REVIEW

2.1 Definition of Data Philanthropy

Data Philanthropy is “*a term which describes a new form of partnership in which private sector companies share data for public benefit*” (Pawelke and Tatevossian, 2013, para. 1). Pawelke and Tatevossian (2013, para. 6) has identified four ways in which corporation could contribute towards Data philanthropy such as

1. “Share aggregated and derived data sets for analysis under non-disclosure agreements (NDA)
2. Allow researchers to analyse data within the private company’s own network, under NDA
3. Real-time data commons: data pooled and aggregated between multiple companies of the same industry to protect competitiveness
4. Public/private alerting network; companies mine data behind their own firewalls and share indicators”

Although data philanthropy appears to be similar to Open Data, it is important to understand that open data can be universally accessed and re-distributed; this is usually *not* the case for data released by corporate data philanthropists.

Although Pawelke and Tatevossian (2013) have identified above the various ways to engage in data philanthropy, Susha, Janssen, and Verhulst (2017) specify a clear difference in the way the data is shared amongst the corporations as well as the various ways in which it is used. Sharing data through collaboration requires one to consider the kind of data being shared, the actual contents of the data, how administrative the data is, where the barriers of access are. Additionally, how the data is used is dependent on the project objectives and the purpose of the project.

Literature brings into the light what is required from the corporations side in order to solve problems that could help the community using technology, skills and tools. Security is a concern due to weak security systems. Another relevant constraint is the South African Protection of Personal Information (POPI) Act which applies to all South African institutions “*that collect, store, process and/or disseminate personal information as part of their business activities*” (Bruyn, 2014, p. 1315).

2.2 Resources Used to Support Data Philanthropy

In order to adequately perform projects that involve big data, skills in data management and storage are needed. These skills include being able to analyse data and measure results effectively, present data visually, build a result-oriented database can be well utilized with the help of data experts such as data scientists are needed to perform projects which involve big data (Lucente, 2017). Data Philanthropy often includes the sharing, not only of data, but also of the necessary skills to analyze the data.

Apart from skills, various technical resources can be used to support Data Philanthropy. Social media can act as a platform for information collections, sharing and coordination in the face of disasters (Open Data Institute, n.d.). Although this is one of many, there is no set method or process to implement Data Philanthropy (Pawelke & Tatevossian, 2013). The United Nations Global Pulse expressed that they are in need of what they call a Data Philanthropist. This includes “*research and technology partners that will help design projects and build prototypes to test new tools*” as well as their need for Data Specialists that would be able to train new staff as well as financial support” (Kirkpatrick, 2013, p. 2).

2.3 Data Philanthropy in Practice

Data which is collected by the Telecommunications industry is amongst that which is suitable for Open Data projects because it can be collected systematically, it covers large proportions of the population, it contains rich information about individuals and it can be merged with data covering diseases in a specific area (Tomsu, Eggermont, & Snel, 2016). Table 1 blow shows the various case studies that involve Data Philanthropy and it is important to note the use of mobile data in the cases.

Table 1. Notable data philanthropy case studies

Author	Year	Country	Company	Focus (Impact) Area
Bloomberg Philanthropies, 2017	2017	Worldwide	Bloomberg	Public Health – Obesity prevention, Tobacco control, Road Safety, Maternal and Reproductive Health, Drowning Prevention Program
Open Data Institute, n.d.	2016	Indonesia	Open Data Institute	Open-source flood map in Indonesia.
Nielsen, et al., 2017	2014- 2015	Brazil	UN Global Pulse,	Analyzed social media data from Twitter which was used to inform communication campaigns to focus their target to HIV/AIDS work which increased the number of HIV tests done in the region.
Aimia, 2017	2014- now	United Kingdom, Canada	Amia	<i>“The company’s data analysts gave more than 10,000 hours of their time to work with more than 40 charities in the UK and Canada”</i>
Global Pulse, 2015	2015	France, Senegal	Orange	Allowed their mobile data to be analysed for patterns to combat climate change. Orange has also offered their skills to manage the data.

2.4 Corporate Social Responsibility

As mentioned above, Data Philanthropy is a form of Corporate Social Responsibility (CSR). Corporate organisations carry out Corporate Social Responsibility for various motives which include those that are value-driven and connected to the culture that the corporate upholds, those that are stakeholder-driven which are driven by the expectations of the stakeholders of the business and those that are performance-driven and are for financial gain (Carroll, 2016). Ali, Fynas, and Mahmood (2017) found that developing nations were more prone to pressure from external stakeholders such as foreign investors, international buyers and the media to practice Corporate Social Responsibility within corporations as opposed to developed nations where the pressure is from the public. Additionally, external factors can include government, activists as well as competitors but little attention has been given to the internal pressures (King & McDonnell, 2012).

There is a gap in literature when it comes to the attitudes that individuals in management feel towards conducting CSR. The Johannesburg Stock Exchange requires an integrated report (Johannesburg Stock Exchange, 2014), which means companies must conduct some form of CSR.

The United Nations found that the only companies that were most open to hearing about data philanthropy were those that understood that data philanthropy is not a charity rather it is vital for the well-being of the community in order for the business to continue doing business in that community (Kirkpatrick, 2013). In light of this requirement, if it was not there, would South African corporations still invest in CSR? This paper will aim to find out the following question: *“What are the attitudes towards Corporate Social Responsibility in South Africa within the Telecommunications, Retail and Banking Industry?”*

There are gaps in literature which raised questions that this study will aim to answer. By doing so study could further aid into research for IS researchers and practitioners. Understanding the willingness, preparedness, capabilities and attitudes towards this phenomenon could help researchers to find better ways to carry through the projects in various organizations in the Retail, Telecommunications and Banking industry.

3. RESEARCH METHODOLOGY

This study adopted a qualitative, exploratory approach to data collection which allowed for an in-depth exploration of corporate attitudes with regards to data philanthropy for social good. Selecting qualitative data helped create a holistic picture of the concerns as well as the level preparedness and willingness of the

organisations. Face to face interviews were conducted. If the participant were not able to take part in the interview due to unavailability or location, a telephonic interview was scheduled. Participants received questions and a short video describing Data Philanthropy beforehand to give them time to prepare for the interview. A sample size of 9 organisations was targeted because the nature of the study which requires in-depth verbal data. Unfortunately, during the course of the research phase, only 5 companies available. One interviewee was available in each company, occupying various positions in all of the industries targeted by the interviewee.

This study followed the grounded theory approach to research. In grounded theory, coding is a method used to label concepts and categorize them accordingly. Particularly, axial coding was used which is a systematic way of discovering the relationships among themes through relating “categories with their subcategories, test the relationships against data, and test the hypothesis” (Cho & Lee, 2014, p. 8). nVivo software was used to conduct the analysis.

4. RESEARCH FINDINGS, ANALYSIS AND DISCUSSION

As mentioned in the research methodology, an inductive approach was used. Upon exploring literature, various themes emerged which helped frame the main research questions that this project aimed to answer. A thematic analysis was used to analyse the data which was collected in this research paper (Braun & Clarke, 2006). The research findings are organized by the research questions which stemmed from the research objectives. The findings will be grouped by Corporate Social Responsibility, Preparedness, Willingness, Concerns and additional themes that arose from the study.

The following section details a list of companies who participated in the study.

Table 2. Organisations that participated in the interview

Company	Number of Employees	Industry
A	22	Mobile-to-Mobile, Telecomms
B	10 000+	Banking
C	160	Telecommunications
D	10 000+	Retail
E	10 000+	Banking

The research questions that this study explored revolved around willingness, preparedness as well as the concerns around Data Philanthropy. There were various gaps found in the literature which did not cover the following questions;

1. What are the attitudes towards Corporate Social Responsibility by South African corporations?
2. What level of involvement are South African corporations willing to take on for social good?
3. What is the level of knowledge that South African corporations have about Data Philanthropy?
4. What concerns do corporations in South Africa have towards participating in Data Philanthropy?

4.1 Attitudes towards Corporate Social Responsibility (CSR)

The participants were asked three questions which aimed to find out what CSR initiatives the company took part, if they did, how the projects were chosen as well what important factors were considered in making that decision.

Company A expressed that their participation in Corporate Social Responsibility (CSR) was based on relationships that they have formed with the school which they helped maintain: “*We knew them personally, we use that as a way to get insight to what they specifically need and we try to respond to that directly*”

Company B expressed that CSR was mandatory within the organisation and thus the following policy was implemented to ensure that all employees took part in the initiative: “*Each department gives feedback that goes up the ladder. Feedback of what you done during the year and project plans for next year*”

Company D stated that there were a number of projects that employees were encouraged to participate in throughout the year, although when asked of participation, the participant responded with: "*I haven't even got a chance to see what is where*".

Company C on the other hand, has been involved in the field of Data Philanthropy (DP) for the past recent years. Upon justifying why they had taken part in the DP, their response was as follows: "*There was a need for people to do this type of work so people were coming to us*".

In summary, the following themes emerged.

Table 3. Themes around Corporate Social Responsibility

Company	Summary of Theme
A	Relationship-based
B	Mandatory
C	Need
D	Mandatory
E	Mandatory

4.2 Preparedness

The second research question was: "What is the level of knowledge that South African corporations have about Data Philanthropy?" In order to answer this question, the participants were asked three questions which aimed to find which technologies the companies currently made use of, what the main uses of the data is as well as if the company was looking to hire any additional skills.

With regards to preparedness and capabilities on taking part in Data Philanthropy (DP), Company A stated that they lacked the skills and thus have been looking into hiring graduates to take on the roles of Data Science.

Company B stated expressed that they had a shortage in the skills required to perform data demands in the company and thus the current team is overworked: "*The team is like way overworked because there such a big demand and a lot of the stuff that we've been doing for the past year it's just basically putting out fires really or just trying to get things in place*".

In contrast, Companies C and E mentioned the shortage of skills that they had faced although they had created a solution to combat this problem: "*We've been taking a whole lot of interns as a percentage of our staff in Cape Town and as * said putting them in those training courses, and it's interesting to see when you actually educated then you come to the real world*". [Company E]: "*In terms of the specification, it's enough but then in terms of work load yes, we do need to get more people.*"

The following table summarizes the themes which emerged.

Table 4. Themes around Preparedness

Company	Summary of Theme
A	Skills
B	Data-driven, demand, shortage of skills
C	Stakeholders, training
D	Enough skills
E	Profits, costs, revenue, shortage of skills

4.3 Willingness

The third research question looked at what level of involvement South African corporations were willing to take on for social good?

The participants were given six various questions which aimed to find out what they thought of the Data Philanthropy concept, how much resources would the organisation be willing to set aside for Data Philanthropy and whether they thought the initiative was important.

Company A expressed that they would be willing to take part in a Data Philanthropy (DP) initiative only if a favourable environment existed. Should the environment not exist, the company will not partake: "...so long as where it gets shared is open, transparent, accountable and responsible. Otherwise, you know, no."

Company D proclaimed that due to the nature if the business, regulations which the South African Government imposed on the industry make it very hard to participate in DP: “*it there's a lot of regulations that come in it especially when you're dealing with medicine, so I don't know if it's something that I would have time for*”.

Speaking on behalf of the organisation, security and usability are the two main components that affected the willingness to participate for Company E. The participant, in their personal capacity however, felt that South Africa would gain in a positive way from the donation of technology skills in the form of hours: “*I think from a personal perspective yes, this is something that is valuable to get these skills out there. From a large corporation perspective, it becomes more difficult because a question mark has been around security and data. There needs to be very strict rules in place already which say how do we manage, how do we do it, what are we sharing, how often, where is it being stored*”.

Although security and data were the main factors to be considered, Company E expressed further what their objectives would be should DP be added to the day-to-day business operations: “*The main objective here is to say how do we increase the profitability of the customer overtime, and how do we increase our cost ratio*.”

Thus, the following themes emerged.

Table 5. Most Prominent Themes around Willingness

Company	Most Prominent Theme
A	Use
B	Security
C	Time, resources
D	Regulation
E	Security, practicability, time, profit, decision-making

4.4 Concerns

The last research objective looks at the concerns which South African corporations have towards participating in Data Philanthropy. The participants were directly asked what concerns they had with Data Philanthropy as well as consumer-data related regulations which data philanthropy would affect within their organisation.

Upon being questions about the concerns of which the participant had with Data Philanthropy (DP), Company A highlighted the concerns around the ownership of the data as well as whether the data would provide the value DP so promises: “*The danger of data is being sure about who owns it just to an extent, and then the second extent is this value in data*.”

Additional Company B highlighted concerns whether the data is secure and the effects this would have on the customer: “*it's always a bit of a challenge with customer data because for obvious reasons you can't just go ahead and share that*” and “*It becomes very questionable because it provides something public that links to our customer base*.”

Security was mentioned again by Company E, in addition to security, Company E mentioned the intended use of the data: “*So one would be security, two also the use. So saying that now we're moving our data we would definitely need to know what they want to do. So simply sharing for the sake of sharing doesn't really mean much in terms of the data. You need to have an objective*.”

In addition to security and usability, Company E further expressed their concern for the use of the data thus providing detail as to what form of use the organisation would consider: “*Depends on how do increase the skills, how do you increase the decision making, how do you increase the overall business making of business decisions for our customer*”

Company D expressed time as a concern: “*Yeah I would if it's something that I can easily do I can get involved it's just so busy and I'm the only one there that it's quite hard but yeah it's something that I would do*”

I'm just a little bit concerned about how other companies would take it in terms of the competitive advantage.”

Company C highlighted time and skills as factors that concerned the business when they chose to take part in DP. Company C further went to highlight the kinds of skills that were needed to embark on DP, which

were not only technical skills: “*these things take lots of time lots of resources and take into account what historically, let’s call it the foundation of our skills, and bringing it too this it’s a huge change*” [Company C]: “*we had to get lawyers to look at the terms and conditions when you upload the data*”

The following table summarizes the themes which emerged.

Table 6. Themes around Corporate Social Responsibility

Company	Summary of Theme
A	Employee Competency, data ownership, value
B	Security, legacy-systems
C	Regulations
D	Time, competition
E	Usability, Regulations

5. RESEARCH CONCLUSIONS

The overall objective of the research paper was to explore the corporate attitudes were towards Data Philanthropy in South African Banking, Retail and Telecommunications. The literature review of this paper opened up several questions which were probed through the various participants. In answering the following questions, it was discovered that the companies responded to Data Philanthropy with the same approach used for their Corporate Social Responsibility (CSR). It was also found that in the sample of the study, all large corporations who had thousands of employees performed explicit CSR as opposed to small companies who instead who performed implicit CSR.

Which concerns do South African corporations have towards Data Philanthropy?

Small South African corporations who performed more implicit Corporate Social Responsibility (CSR) are concerned about the ownership regulations of data, the value that the data will provide to the community and if it will live up to that promise, not having enough skills to participate in Data Philanthropy as well as the limited time available during business operations to dedicate to Data Philanthropy.

Large South African corporations who performed explicit Corporate Social Responsibility (CSR) are concerned about the regulations around consumer data, security around the data during Data Philanthropy, limited time to perform Data Philanthropy and lastly the regulations and policies which the corporations follow that hinder their freedom to take part in projects such as Data Philanthropy.

What level of involvement are South African corporations willing to take for social good?

Small South African corporations are willing go as far as share data in a secured setting, willing to donate hours and willing to hold the data while being used. Large South African are overall reluctant to be involved heavily in such an initiative.

What are the capabilities do South African corporations lack which would not allow them to easily implement Data Philanthropy today?

South African corporations, both large and small lack the time to as well as the required skills to participate in Data Philanthropy.

What are the attitudes towards Corporate Social Responsibility in South Africa within the Banking, Retail and Telecommunications?

South African corporations both large and small are active in Corporate Social Responsibility, some making it mandatory for their employees to dedicate time to partake in it. Overall South African corporations, specifically those with a high number of employees are reluctant when it comes to taking part in Data Philanthropy due to concerns around consumer data security although the technology and capabilities are available. Small corporations on the other hand in South Africa are more willing to participate in Data Philanthropy but are hindered by the limited skills that they possess and time.

This study was exploratory given the newness of the phenomenon under investigation; indeed most organisations are not aware of Data Philanthropy. Thus only a very small number of organisations could be sampled and their views may not be representative. Indeed, if anything, sample bias means that most companies have even fewer capabilities or less positive attitudes towards Data Philanthropy.

In future, the field would benefit from looking into the specific concerns around Data Philanthropy participation accompanied by situations where South African corporations would and would not participate in Data Philanthropy. It would benefit this study to research whether introducing incentives to corporations such as tax breaks would incentivize large corporations enough to take part in Data Philanthropy. Both suggestions to future researchers would help provide a starting point for the growth of Data Philanthropy. By finding out what a favourable environment would be to house such an imitative, it could incentivize the South African government to consider modern ways to better service delivery in the country.

REFERENCES

- Aimia. 2017. Data Philanthropy. Retrieved May 20, 2017 from: <https://www.aimia.com/en/social-purpose/data-philanthropy.html>
- Ali, W., Frynas, J. G., & Mahmood, Z. 2017. Determinants of Corporate Social Responsibility CSR Disclosure in Developed and Developing Countries: A Literature review. *Corporate Social Responsibility and Environmental Management*, 24 2.
- Bloomberg Philanthropies. 2017. Public Health. Retrieved May 20, 2017, from Bloomberg Philanthropies: <https://www.bloomberg.org/program/public-health/#intro>
- Braun, V., & Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2) 77-101.
- Bruyn, M. 2014. The Protection of Personal Information POPI Act - Impact on South Africa. *International Business & Economics Research Journal*, pp. 1315-1339.
- Carroll, A. B. 2015. Corporate Social Responsibility: The center of competing and complementary frameworks. *Organisational Dynamics*, 44, 87-96.
- Cho, J. Y., & Lee, E.-H. 2014. Reducing Confusion about Grounded Theory and Qualitative Content Analysis: Similarities and Differences. *The Qualitative Report*, 19 (32), 1-20.
- Coulton, C., George, R., Putnam-Hornstein, E., & de Haan, B. 2015. Harnessing Big Data for Social Good: A Grand Challenge for Social Work. *American Academy of Social Work & Social Welfare*.
- Fiegerman, S. 2017. Earnings. Retrieved June 19, 2017, from CNN Technology: <http://money.cnn.com/2017/05/03/technology/facebook-earnings/index.html>
- Global Pulse. 2015. Leading Global Businesses, Policymakers Share Commitment to Work with UN to Enable a Data Revolution for Climate Action. United Nations. Paris: Global Pulse.
- Johannesburg Stock Exchange. 2014. Retrieved from JSE: <https://www.jse.co.za/content/JSERulesPoliciesandRegulationItems/JSE%20Listings%20Requirements.pdf>
- Kilian, T., & Hennigs, N. 2014. Corporate social responsibility and environmental reporting in controversial industries. *European Business Review*, 26 (1), 79-707.
- King, B. G., & McDonnell, M.-H. 2012. Good Firms, Good Targets: The Relationship between Corporate Social Responsibility.
- Kirkpatrick, R., 2013. Big data for development. *Big Data*, 1(1), pp.3-4. DOI:10.1089/big.2012.1502
- Lucente, I. 2017. Blog: Data Philanthropy - The Analytics Skills for the Community. Retrieved from John Snow Labs: <http://www.johnsnowlabs.com/dataops-blog/data-philanthropy-the-analytics-skills-for-the-community/>
- McWilliams, A., & Siegal, D. 2001. Corporate Social Responsibility: A Theory of the Firm Perspective. *Academy of Management Review*, 26 (1), 117-127.
- Nielsen, R. C., Luengo-Oroz, M., Mello, M. B., Paz, J., Pantin, C., & Erkkola, T. 2017. Social Media Monitoring of Discrimination and HIV Testing in Brazil 2014-2015. *AIDS and Behaviour*, 1-7.
- Open Data Institute. n.d. ODI Showcase: Peta Jakarta Real Time Flood Mapping. Retrieved from Open Data Institute Showcase: <https://theodi.org/odi-showcase-peta-jakarta-real-timeflood-mapping-jakarta>
- Pawelke, A., & Tatevossian, A. R. 2013. United Nations Global Pulse. Retrieved from <http://www.unglobalpulse.org/data-philanthropy-where-are-we-now>
- Soldner, M. Writer. 2016. Your company's data could help end world hunger [Motion Picture].
- Statistics South Africa. 2017. Poverty Trends in South Africa. Statistics South Africa.
- Susha, I., Janssen, M., & Verhulst, S. 2017. Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Developmen. *50th Hawaii International Conference on System Science*, pp. 2691-2700.
- Tomsu, K., Eggermont, A., & Snel, N. 2016. *Telecom data for efficient malaria interventions*. Bloomberg Data for Good Exchange Conference. New York: Bloomberg.

THE EFFECTIVENESS OF USING SOCIAL MEDIA IN GOVERNMENT COMMUNICATION IN UAE

Elsayed B. Darwish
Zayed University, UAE

ABSTRACT

The purpose of this paper is to explore the effectiveness of using social media by the government entities in UAE. It is a network analysis study seeking mainly quantitative data. 100 UAE government accounts, belong to (25) entities, were examined using web-based analytical tools. The general findings of the study indicate that the UAE government communication entities' accounts convey a rich picture of how these entities interact with their stakeholders on their social media. Most of them are performing an active role in terms of publicity, reach, marketing and transparency.

KEYWORDS

Social media, Government Communication

1. INTRODUCTION

Nowadays, social media became a huge virtual community, with highly interactive and collaborative environment among its members, which led many governments to move from e-government services to social government; i.e. governments provide their services over social media (Khasawneh, R. & Abu-Shanab, E. (2013: 10-17)

Using social media in e-government will enhance (1) the transparency, by releasing the information that citizens are always checking; (2) Participation by maintaining the citizens' engagement with their governments, and allowing citizens to express their opinions, experiences and wisdoms; (3) Collaboration, where citizens participate by creating the content of government topics and the government use and follow the content generated by citizens to fulfill government mission. (Dareen A. Mishaaal & Emad Abu-Shanab (2015) Zavattaro, S. M., & Sementelli, A. J. (2014: 257-264).

Generally, governments have a threefold social media strategy: push, pull, and networking. These strategies are often guided by the governments with social media policies and guidelines for properly using social media. While many government agencies are experimenting the use of social media, however very few recognize the differences between these strategies, utilize the whole capacity and types of social media and actively measure the impact of their digital interactions (Khasawneh, S., Jalghoum, Y., Harfoushi, O., & Obiedat, R. (2011: 568-582).

The widespread recognition of the potential of social media technology for achieving public outcomes does not match our understanding about how and why specific tools are being used for specific purposes. (Lee, M., & Lee Elser, E. (2010: 3-4).

in the same time, there has been little research to understand how and why these tools are becoming integrated in the public sector and it is not clear how successful and effective they are. (Mergel, I. (2010).

Many governments have problems in their communication, due to low budget and because they put communication with stakeholders as a low priority. Also, governments use the traditional methods of communication such as newspapers, radio and television; this one-way communication reflects the low feedback of stakeholders to communicate with government, which leads to low participation from stakeholder's side. In addition, many governments have multiple deficiencies which include the low trust in government, limited accountability, lack of transparency and low quality of service related to deficiencies in accessibility to the different services (Hofmann, S., Beverungen, D., Räckers, M., & Becker, J. (2013: 387-396).

This paper proposes a framework for measuring the effectiveness of the government communication, and tests this framework which consists of sets of missions and indicators on different governmental social media platforms in UAE. The UAE presents a unique example of using social media in government communication. The eGovernment program in the UAE has achieved remarkable success in enabling the technology-based transformation of public services UNDESA (2012). However, this digital interaction needs to be analyzed and interpreted to understand to what extent they support government's mission.

Since 2011, the UAE has produced guidelines for using social media to be integrated with the government communication. Although such policies and guidelines are useful for understanding the general principle of social media use in government, it remains unclear how each entity is working in cooperative efforts through social media based on those strategies and guidelines, and to what extent they improve on the delivery of citizen services using social media in conveying the entity's message and information, transparency, participation, engagement etc., and what is the most dominant role played by UAE government communication entities in using social media; Do they focus on disseminating information externally or promoting and marketing their entities or getting feedback on service quality or facilitating participation by citizens or external stakeholders or seeking more engagement from their stakeholders?

2. LITERATURE REVIEW

This paper explores the literature to better understand the environment of social media and its utilization in e-government communication effectiveness. The differences between the conventional media and social media, the importance of social media as a communication channel, and the reasons that lead governments to adopt it in their communication will be also investigated in the literature section. Also, some models of e-government based on social media will be reviewed and the strategies of governments adopted to communicate over the social media by UAE government entities. The paper seeks to set criteria for measuring the effectiveness of using social media in government communication based on the literature, and the practical uses of social media.

2.1 Government Communication between Conventional Media and Social Media

The use of social media applications by government organizations can be called an extension of the current digitization efforts of government services as a new wave of the e-Government era. (Bertot, John C., Jaeger, P. T., & Grimes, J. M. (2010: 264–271).

In most cases, social media applications have not replaced existing offline or even e-Government services. On the contrary all social media applications are used to complement the existing communication mechanisms in government. (Khasawneh, S., Jalghoum, Y., Harfoushi, O., & Obiedat, R. (2011: 568-582), these deficiencies can lead to the public's regarding of government communication as little more than propaganda (Mergel, I. (2013). Social media applications are used to create additional channels for governments' interactions with its stakeholders (Bretschneider, S. I., & Mergel, I. (2010: 187–203).

2.2 Governments' uses of Social Media

Research on government communication in social media and social media in general is limited at best. Most studies focus on particular application areas and refer to communication only peripherally. Brainard and McNutt (2010) analyzed the structure of the discussion between the police in Washington, D.C., and the public in online discussions forums and classify the interactions as informational, transactional, or collaborative, finding that most of the content is informational, although the platform allows more transactional and collaborative interactions.

Panagiotopoulos and Sams (2012) in their two-stage study on the use of Twitter by the UK local government, found that over 296,000 tweets were collected from the accounts of 187 UK Local Authorities. They also found a positive relation between the signup date and the number of tweets as well as the number of followers. Mundy and Umer (2012: 502-511) concluded that the use of Twitter has not yet matured and

provided a set of recommendations, including understanding the nature of Twitter as a communication channel, as well as the importance of developing a social media policy.

Dalakiouridou et al. (2012) shows that Twitter is the preferred method of communication with the European public and is used by these institutions for short notifications, link-sharing and cross-dissemination.

Alasem (2015: 67-74) found that the level of maturity of using Twitter by Saudi public authorities in general has not matured yet, and there is a significant difference between the performances of these accounts.

Sara Hofmann et al (2013) found that multimedia features like photos and videos contribute to the success of communication. They also saw a tendency for citizens to prefer topics related to leisure activities.

2.3 Government-Social Media Based Models

Previous research is short on theoretical models and managerial insights into the effectiveness of governments' online communication strategies and techniques.

Chm and Luna (2012) proposed three models. The first model assumes that social media is a catalyst that transforms citizens, government and data. It discusses how social media interacts with these three dimensions. Social Media-Based Citizen Engagement Model focuses on social media used as a tool to enable users to express their opinions, emotions, behaviors and interactions. Governments can use such media to transform their citizen to participate in good governance and to enjoy democracy. Social Media-Based Data Sharing Model focuses on the data that a citizen generated using social media when he/she participated in a political topic. Social Media-Based Real-Time Collaborative Government Model focuses on the idea that using social media in e-government starts to enhance the communication between governments and citizens to be nearly in real time.

Lee & Kwak [2012] proposed a model for open government that is based on public engagement. The model contains five levels, which is based on the benefits of social media that provide transparency, participation and collaboration by public engagement. Mishael & Abu-Shanab (2015) proposed a framework for investigating the success of communication between governments and other stakeholders utilizing social media. The framework is founded on a set of proposed factors that lead to communication success and they are: transparency, participation, collaboration, comfort, and the posted topic. Mergel, (2013) proposed a framework for measuring social media interactions in the public sector contains from mission, goals, tactics, mechanisms, and outcomes. The main three missions are: transparency, participation and collaboration. The goals are: information education, engagement and cross-boundary action two way interactives. Mergel (2013a) distinguishes between three possible strategies for social media use: 1) representation; 2) engagement; and 3) networking.

3. RESEARCH METHODOLOGY

The paper seeks to answer the following questions:

To what extent the government entities in UAE use the social media effectively?

To what extent the government entities in UAE are effective in using social media in communicating with their stakeholders? What are the main roles of social media in government communication?

A proposed framework has been used to conduct a comparative empirical study that compares UAE government entities' social media features, strategies and techniques (quantitatively and qualitatively), by using web-based analytical tools such as quintly, zoom sphere, union metrics, etc., It has employed a wide range of webometric and social network analysis techniques for measuring the effectiveness of using social media.

Using the UAE government portal, 25 entities have been selected, with the total of 100 social media accounts (25 Facebook pages, 25 twitters pages, 25 YouTube sites and 25 Instagram accounts). We found that some entities do not have either Facebook or twitter or YouTube or Instagram). The total of active accounts were (65).

To examine the effectiveness of using social media by the UAE government communication entities, group of missions and indicators have been constructed based on the previous studies, and criteria used by web- based analytical software and websites.

The study has investigated the following missions: publicity & communication, marketing, reach, transparency, participation, and collaboration & engagement over Facebook pages, twitter, Instagram and YouTube accounts of these entities, to measure the effectiveness of their governmental communication

The study proposes the following: (1) The higher the publicity indicators used, the lower the communication effectiveness (2) The higher the marketing indicators used, the lower the communication effectiveness (3) The higher reach indicators used, the moderate the communication effectiveness (4) The higher the transparency indicators used, the moderate the communication effectiveness (5) The higher the participation indicators used, the higher the communication effectiveness (6) The higher the collaboration & engagement indicators used, the highest the communication effectiveness.

To measure each mission of government communication, the total of the following indicators of each governmental Facebook, twitter, YouTube and Instagram will be determining and then will be compared as a group with other indicators of another mission (as shown in Table 1).

This is an exploratory network analysis study, more than a content analysis, which uses quantitative data covering the aspects of the number of posts, tweets, followers, fans, interactions etc., using a multi-method approach.

4. DATA

4.1 Governments' Materials Properties in Social Media

The automatically evaluated indicators/ features (posts, fans, followers, likes, shares, links, photos, videos, comments, replies, tweets, impressions etc.) contain data from the 25 entities that have Facebook, twitter, Instagram and YouTube accounts as of Nov and Dec 2016.

The remaining indicators/ features (links to services, speed of the response, posting events, notes, announcements, etc.) had to be coded manually. (As shown in Table 2). For anonymity reasons, we combined the data of these entities together in some case and classified it to three categories in other case (entities based services, production and authority). Table 2 contains the aggregated results, the manual assessed categories are marked with an asterisk. Some indicators were not included in the combined factors which we created to compare between the entities regarding performing their missions (shown in table 2). It provides us extra indicators of to what extent these accounts are effective.

The rows of Table 2 contain data from all entities with Facebook, twitter, Instagram and YouTube accounts in the columns “Total”, “Mean”, and “Standard deviation”. The total Facebook indicators are (35), twitter (17), YouTube (9) and Instagram (5)

The results show that most UAE government communication entities are using twitter, then Facebook, then Instagram and lastly YouTube. The total number of Facebook fans of the UAE government communication entities was 1926357, the mean number of fans was 77054, with a standard deviation of 161592, twitters followers was 2320187, the mean number of twitter followers was 92807, with a standard deviation of 150757, YouTube subscribers was 5757, the mean number of subscribers was 230, with a standard deviation of 880. Instagram Followers was 260014, the mean number of Instagram followers was 10400, with a standard deviation of 46886. While all of the entities have twitter accounts, some of them do not Facebook accounts and most of them they do not activate their Instagram accounts.

Table (2) shows that the number of tweets is 162246, Instagram posts is 100365, using links in Facebook is 313, videos 769, photos 767, YouTube videos are 927, which confirms the above results of depending mainly on Tweeter to communicate with the stakeholders.

The UAE government communication entities have reached many stakeholders; through shares 11786, Facebook likes 171876, twitter likes 8756, retweets 4523. The potential reach of Tweeter was 14540585, impressions 28405971 and mentions 2393, which reflects their interest in performing reach mission.

The total numbers of indicators related to engagement with the stakeholders show less focus in performing this task, the number of user posts on Facebook was 129, replies to users in Tweeter was 470, and the total number of contributors in Tweeter was 842. The mean of response rate on YouTube was (.4) in Instagram, (.16), on twitter (.0), and the same of Facebook, which shows the lack of attention giving to the engagement mission.

Analyzing these accounts manually shows few interest in replying to fans' comments or questions on the Facebook, the slowness in responding to questions, the overwhelming of the entities' posts and tweets, not the stakeholders' contributions, the no use of applications available completely on Facebook, or applications directing to gov. website or live stream. Few of them have been posted notes, announcements, surveys, and contests.

The manually analysis shows that these entities are mainly using one one-way push tactics with their stakeholders, focusing on conveying the entity's messages, promoting and marketing its identities and activities more than seeking stakeholders' participation and engagement

Table 1. Coding scheme for governments' materials properties in Social Media

Indicator	Code	Indicator	Code
Fans (Facebook)	(C)	Notes*	(A) (B) (D)
Fan's Change (Audience change over the time)	(C)	No of Followers (twitter)	(C)
N of likes (Facebook)	(C) (E)	No of tweets	(A) (B) (D)
N of comments (Facebook)	(C) (E)	Following	(C)
N of shares (Facebook)	(C) (D) (E)	Twitter Likes	(C) (E)
interactions/ Total (Facebook)	(C) (D) (E)	Twitter List	(C)
admin posts/ status	(A) (B)	AVG Likes	(C) (E)
admin posts/ Links	(A) (B) (D)	Retweets	(C) (E)
admin posts/ videos	(A) (B) (D)	AVG retweets**	(C) (E)
admin posts/ photos	(A) (B) (D)	Own Tweets	(A) (B) (D)
admin posts/ total**	(A) (B) (D)	Retweeted Tweets	(C) (E) (F)
	(C) (D) (E)	Replies to Users (Twitter)	(C) (D) (E) (F)
Reactions (Facebook)			
User Posts (Facebook)	(C) (E) (F)	Mentions	(C)
I-Rate (Facebook)	(C) (D) (E)	I-rate	(C) (E)
Response rate	(D) (E) (F)	Response rate (Twitter)	(C) (D) (F)
number of people talking about the entity	(C) (E) (F)	Contributors	(C) (F)
N of replies (Facebook)	(C) (D) (E) (F)	Estimated & potential reach	(C)
Links to services & media & Governments' sites*	(A) (B)	Impressions	(C)
Speed of response (Facebook)*	(C) (D) (E)	YouTube subscribers	(C) (E)
Government site enables Live Stream*	(A) (B)	YouTube Videos	(A) (B) (D)
Govt. site enables posts that include calendar & events*	(A) (B) (D)	YouTube Views	(C)
Govt site enables posts that include survey & pool	(A) (B) (D) (E) (F)	YouTube Views changing Rate	(C)
Govt. site enables their stakeholder to post*	(C) (D) (E) (F)	YouTube Likes	(C) (E)
Govt. site enables their stakeholder to co-create*	(C) (D) (E) (F)	YouTube Dislikes	(C) (E)
Page's policy (Facebook)*	(A) (B) (D)	YouTube Comments	(C) (E)
Read more (Facebook)*	(A) (B)	YouTube I-Rate	(C) (D) (E)
		YouTube Views Subscribers rate	(C)
Background info (Facebook)*	(A) (B)	Instagram posts	(A) (B) (D)
Reports and Announcements*	(A) (B) (D)	Followers	(C)
Up-to-date (Facebook)*	(A) (B) (C) (D)	Following	(C) (E)
Applications completely on Facebook*	(A) (B) (C)	Average like**	(C) (E)
Applications directing to gov. website*	(A) (B) (C)	Average Comments**	(C) (E)
Visitors' publications & opinions & Reviews*	(B) (C) (D) (E) (F)		
Contests (Facebook)	(B) (C) (E) (F)		

(A) Publicity & Communication & Access (B) Marketing & Comfort (C) Reach (D) Transparency (E) Interactions & participation (F) Collaboration & Engagement

4.2 Effectiveness of Government Communication with regard of accomplishing their Missions

This section contains the results concerning the effectiveness of the UAE governments' entities posts properties on social media. We present the analysis of the effectiveness of using social media in performing their missions towards their entities; publicity, marketing, reach, transparency, participation and engagement. It has been evaluated in terms of total used of indicators, such as the frequency of posts, comments, shares, likes, posts, tweets, etc., of the four types of social media: Facebook, YouTube, twitter and Instagram accounts.

In general, the results differ significantly among the entities in terms of missions accomplished we analyzed in depth, which we ascribe to the governments' differing social media strategies

Table (3) shows that the UAE government entities were interested in performing the reach (Mean = 4062867), transparency (Mean = 2085988) and publicity mission, (Mean = 11079), then the marketing and participation missions. The least mission was the engagement (Mean = 642). A big mean difference in terms of accomplishing the reach mission comparing to other missions. While the minimum of reach is (105317), it is (9) in terms of participation mission and 17 in engagement mission. Similarly to Mergel's findings (2013a), the one-way representation strategy is the most prevalent.

As per of the higher used of the publicity, marketing, reach and transparency indicators, and the less use of the engagement indicators, we can conclude that the effectiveness of using social media by UAE government entities is moderate in general.

Table 2. The aggregated results of indicators used by social media of government's entities

Indictor	Sum	Mean	Indictor	Sum	Mean	Std. Deviation	Indictor	Sum	Mean
Fans	1926357.00	77054.2800	To create*	.00	.0000	.00000	Replies to users	470.00	18.8000
Fans Change	26.84	1.0736	Page Policy*	3.00	.1200	.33166	mentions	2393.00	95.7200
Likes	171876.01	6875.0404	Background*	19.00	.7600	.43589	Tweeter_L_rate	.00	.0000
Page Like	1821645.00	72865.8000	Read more*	19.00	.7600	.43589	Tweeter_R_Rate	159.16	6.3664
Comments	5946.00	237.8400	Reports& announcements	20.00	.8000	.40825	Contributors	842.00	33.6800
shares	11786.00	471.4400	Up-to-date	20.00	.8333	.38069	Potential reach	14540585.00	581623.4000
Interactions	184581.00	7383.2400	App	.00	.0000	.00000	impressions	28405971.00	1136238.8400
statutes	560.00	22.4000	Apps	1.00	.0417	.20412	YouTube subscriber	5757.00	230.2800
links	313.00	12.5200	Visitors' Reviews	760.00	31.6667	70.22304	YouTube videos	927.00	37.0800
videos	769.00	30.7600	Contests	1.00	.0417	.20412	YouTube views	2.50	.1000
photos	767.00	30.6800	Notes	6.00	.2500	.44233	Views change	8.60	.3440
Reactions	148934.00	5957.3600	Followers	2320187.00	92807.4800	150757.80114	YouTube likes	4.00	.1600
user Posts	129.00	5.1600	Tweets	162246.00	6489.8400	8480.05839	YouTube dislikes	.00	.0000
Interactive rate	.03	.0010	Following	3557.00	142.2800	139.80687	YouTube comments	.00	.0000
Response Rate	.00	.0000	Tweets likes	8756.90	350.2760	946.87715	Y_I_rate	4.00	.1600
Share Quota*	.99	.0397	Lists	331.00	13.2400	62.88301	Y_views_sub	26.30	1.0520
People talking about it*	12021.00	480.8400	AVG Likes	101.40	4.0560	3.58563	Instagram posts	100365.00	4014.6000

Replies*	3.00	.1200	Retweets	4523.10	180.9240	262.56224	Inst. followers	260014.00	10400.5600
Links to services*	23.00	.9200	AVG retweets	101.20	4.0480	4.84158	Inst following	719.60	28.7840
Speed of response*	23.00	.9200	own tweets	1705.00	68.2000	78.76547	Inst_AVG_like	1239.80	49.5920
Livestream*	1.00	.0400	Retweeted tweets	1015.00	40.6000	72.31759	Ins_AVG_comments	4.00	.1600
Survey*	3.00	.1200							

* counted Manually ** did not include in the combined factors

4.2.1 Publicity & Communication & Access

To assess the effectiveness of performing the publicity mission by the UAE government entities by using social media, many indicators have been analyzed such as posts, links, videos, photos, links to services & media & governments' sites, government site enables Live stream, government site enables posts that include calendar & events, publishing the page's policy (Facebook), number of tweets, YouTube videos and Instagram posts, etc. (Table 2) shows that they mainly used posts, likes, photos, and tweets. Many indicators or features were missing such as enabling posts on applications, live stream, including events, calendars, and notes or conducting surveys etc. They were not interested in posting their policies regarding posting on the accounts or providing adequate information about their entities. Some of them post inactive features or links and do not provide enough links to other services or sites. However most of them posted photos and videos, and post updated news which shows their official activities.

4.2.2 Marketing

The most indicators or features which have been used by these entities to market the entities were using photos, videos, posts and tweets. Some of YouTube accounts either are not active or do not post enough videos. Most of the UAE government entities do not activate their Instagram accounts. Only one account uses live stream. Few of them are conducting pools or surveys. Most of them do not post update events on their calendars. However, some of them post reviews from their users.

4.2.3 Reach

The “reach” mission was the most accomplished mission. As Table (2) shows, the number of Facebook fans is 1926357, the number of twitter followers is 2320187, potential reach of Tweeter is 14540585, the impressions of Tweeter is 28405971 and Instagram followers is 260014. However, mean of fan's change is 1.07, which means that the same stakeholders are using these accounts.

4.2.4 Transparency

The UAE government communication entities have used several features and indicators of transparency through using their own social media. They mainly focus on sharing photos, videos, links, posts, and retweets. However, many indicators and features which could enhance the transparency were few or missing such as replies, speeding the response, allowing user to post, publishing more visitors' reviews and post more reports, notes, announcements and providing their stakeholders with more update news and events. Also, publishing posts and tweets without quick response, reply, and comment could lead to negative impacts on the effectiveness of the use of social media.

4.2.5 Participation

Allowing stakeholders to participate through comments, shares, posts, replies, retweets, tweets, to create posts, to like or dislike came in the fifth rank of UAE government communication entities' interests or mission accomplished. The most noticeable indicator was “likes” in Facebook, with 6875 mean, and reactions (mean= 5957). However, the mean of people who are talking about the entities' Facebook was 480, and visitors' reviews was 31. However, most of these indicators reflect part of participation aspects.

4.2.6 Collaboration & Engagement

The least accomplished mission by UAE government communication entities in their social media accounts was engaging their stakeholders in producing contents, posts, tweets, etc., with a mean = 642. It was noticeable the less interest in allowing users to post or create their own content, the slowness of responses and replies and the few numbers of contributors on tweeters. In spite of allowing stakeholders to post their own comments, some of these comments are not related to the posts itself and do not build on it to create series of replies and discussions and some of them should be filtered or removed.

Table 3. Effectives of government posts with regard to mission accomplished

Mission	Total	Minimum	Maximum	Mean	Std. Deviation
Publicity	265917.00	150.00	49428.00	11079.8750	13581.41177
Marketing	266677.00	266.00	49428.00	11111.5417	13588.79054
Reach	97508811.34	105317.82	15089756.00	4062867.1393	4495150.04982
Transparency	50063735.59	69455.82	8035548.00	2085988.9831	2304770.34735
participation	215282.03	9.00	106514.00	8970.0848	21818.05879
Engagement	15425.19	17.00	7923.00	642.7161	1590.64890

5. CONCLUSION

The results of analyzing the UAE government communication entities' accounts convey a rich picture of how these entities interact with their stakeholders on their social media. Most of them are using twitter, then Facebook, then Instagram and lastly YouTube. The same result was reached by Alasem (2015) (Munday and Umer (2012). They tweet than posting on Facebook or YouTube or Instagram. While all of the entities have twitter accounts, some of them do not have Facebook accounts and most of them do not activate their Instagram accounts. They have reached many stakeholders; through shares, Facebook likes, twitter likes, and retweets.

There is a significant difference among the entities in terms of missions accomplished. The UAE government entities were interested in performing the reach, transparency and publicity missions, then the marketing and participation. The least mission was the engagement. A big mean difference in terms of reach mission comparing to other missions.

Analyzing these accounts manually showed less interest in replying to fans' comments or questions on the Facebook, slowness in responding to questions, no use of applications available completely on Facebook, or applications directing to Gov. Website or live stream. Few of them have been posted notes, announcements, surveys, and contests and the overwhelming of the entities' posts and tweets, not the stakeholders'.

While the UAE government communication entities differ in terms of reach, participation and transparency missions "mean". There is no difference between them in terms of performing on publicity, marketing and engagement missions. The services based entities are willing to be more reachable, participatory and transparent. They mainly depend They are currently focusing mostly on push techniques and use social media channels to provide information and marketing their own entities by having online presence and trying to reach many stakeholders.

Many of the features published by government entities used posts, tweets, photos, videos, etc. to reach many stakeholders, but they do not focus on using other features, which could lead to more participation and engagement. While the used features could lead to more publicity and marketing the entities among the stakeholders, but most of them belong to the one-way push tactics, not pull or networking tactics which could create more participation and engagement. (Mergel, 2013).

Many features on these accounts are not active, such as notes, calendars, links, applications, etc. some of them do not update their data and news. They publish the same materials and content which they publish on the conventional media or the website itself such as their press releases, publications, reports without realizing the social media need different content with different treatment.

While Tweeter accounts are active in most entities, most of other accounts, especially in YouTube and Instagram are not active. Examining the comment, and share features in the Facebook does not show that there is enough interest in communicating, and replying to their users.

In addition, they rarely make infrequent use of the specific data types offered by Facebook, such as calendars and events.

We operationalized the effectiveness of a feature/ post/ indicator in terms of the frequency and the polarity of the reactions (replies, comments, retweets, shares, etc., it evoked, as we defined the communication goals as publicizing the entity's information, creating awareness of its role and image, increasing the reach of social media to its stakeholders, transparency, participation and engagement. In general, this evaluation worked well, but other measures and levels of effectiveness should be applied to evaluate the communication effectiveness in terms of other goals. A concrete definition for each factor/ mission should be clearly distinguishable from other factors or missions. Some missions such as publicity and marketing could be combined together.

The analysis of how UAE government communication entities use social media reveals that UAE government entities is using this new technology, effectively to some degree. Although social media offer potential that could overcome the defects of governments' traditional offline communication, this potential is not well-exploited by these entities. The analysis reflects that the online communication behavior of UAE government entities is based on disseminating information in a traditional way, without getting more engaged with their stakeholders. While there are interested in reaching more stakeholders, through transparency, marketing and publicity, they are less interested in efforts of using social media in empowering more participation and engagement.

In order to answer the question guiding the research concerning which communication behaviors of governments is effective on social media, we measured the effectiveness of communication in terms of the publicity, marketing, reach, transparency, participation and engagement.

We found that UAE government communication entities were interested in using one-way push tactics more than using pull and network tactics, the same conclusion reached by several international studies, applied on different governments. However, they should recognize the most important benefits of social media, which focus on engaging the citizens and allow them to participate in the decision making process.

The study has a number of limitations that lead to suggestions for further research. First, our sample size for the quantitative analysis, which consisted of almost more than half of government administrations. Nevertheless, we recommend enlarging the number of analyzed government accounts in future studies. Second, our methods may not be without bias. Using different network analysis tools could produce different data and results. It is better for the government to work on creating its own matrix for measuring the effectiveness and build its own analysis tools to be customized with their needs. Third, we did not apply sentiment analysis or qualitative analysis. An analysis of posted topics can be used to reveal what topics attract stakeholders and encourage the entities to focus on them.

In sum, these suggestions indicate that social media should be seen as a communication channel in addition to the governments' traditional channels like letters, press releases, and static websites. They need to rethink their role, as their traditional understanding of themselves as information and service providers (Sandoval-Almazan & Gil-Garcia, 2012) no longer fits the paradigms of the new media, and they need to professionalize their behavior in social media (Mergel, 2012).

ACKNOWLEDGEMENT

This paper is a working paper, funded by Zayed University. It has been published by Journal of Arab and Muslim Media Research. Volume 10 issue 1. Cover Date: April 2017.

REFERENCES

- Abdulrahman Alasem. (2015) e-Government on Twitter: The Use of Twitter by the Saudi Authorities. *Electronic Journal of e-Government* Volume 13 Issue 1, pp 67-74
- Al-Dalou', R. & Abu-Shanab, E. (2013). E-Participation Levels and Technologies. The 6th International Conference on Information Technology (ICIT 2013), 8-10 May, 2013, Amman, Jordan, pp.1-8.
- Bertot, John C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3), 264–271.
- Brainard, L. a., & McNutt, J. G. (2010). Virtual government–citizen relations: Informational, transactional, or collaborative? *Administration & Society*, 42(7), 836–858.
- Bretschneider, S. I., & Mergel, I. (2010). Technology and public management information systems: Where have we been and where are we going. In D. C. Menzel, & H. J. White (Eds.), *The state of public administration: Issues, problems and challenges* (pp. 187–203). New York: M.E. Sharpe, Inc.
- Camacho, R., & Kumar, M. (2012). Social Media on e-Government. Accessed in 2014 from: http://www.academia.edu/1958732/Social_Media_on_e-government, visited on 29/11/2016.
- Chun, S., & Luna Reyes, L. F. (2012). Social media in government. *Government Information Quarterly*, vol. 29(4), pp. 441-445.
- Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly*, 11(3), 355–366.
- Dareen A. Mishaal & Emad Abu-Shanab (2015), The Effect of Using Social Media in Governments: Framework of Communication Success, ICIT 2015 The 7th International Conference on Information Technology doi:10.15849/icit.2015.0069
- Elbadawi, Ibrahim. (2011). Cloud computing for e-government in UAE: opportunities, challenges and service models. In Proceedings of the 5th international conference on theory and practice of electronic governance (ICEGOV '10) (Tallinn, Estonia, September 26-28, 2011). ACM Press, New York, NY, 2011, 387-388.
- Fichter, D., & Wisniewski, W. J. (2008). Social media metrics: Making the case for making the effort. *Online*, 45–57.
- Fisher Liu, B., & Horsley, J. S. (2007). The government communication decision wheel: Toward a public relations model for the public sector. *Journal of Public Relations Research*, 19(4), 377–393.
- Khasawneh, R. & Abu-Shanab, E. (2013). E-Government and Social Media Sites: The Role and Impact. *World Journal of Computer Application and Technology*, Vol. 1(1), July 2013, pp. 10-17.
- Khasawneh, S., Jalghoum, Y., Harfoushi, O., & Obiedat, R. (2011). E-Government Program in Jordan: From Inception to Future Plans. *International Journal of Computer Scienc Issues(IJCSI)*, vol. 8(4),pp.568-582.
- Mergel, I. (2012). The social media innovation challenge in the public sector. *Information Polity*, 17, pp. 281–292.
- Mergel, I. (2013 a). Social media adoption and resulting tactics in the US federal government. *Government Information Quarterly*, vol. 30(2), pp.123-130.
- Mergel, I. (2013 b). A framework for interpreting social media interactions in the public sector. *Government Information Quarterly*, vol.30(4), pp.327-334.
- Mergel, I. (2010): The use of social media to dissolve knowledge silos in government, in: O'Leary, R., Kim, S. and Van Slyke, D. M. (Editors): *The Future of Public Administration, Public Management and Public Service Around the World: The Minnowbrook Perspective*, Georgetown University Press, pp. 177-187.
- Panagiotopoulos, P. and Sams, S. (2011) Twitter in local government: a study of Greater London Authorities. [Online]. Available from: <http://bura.brunel.ac.uk/bitstream/2438/6239/2/Fulltext.pdf>. [Accessed 19th November 2016].
- Panagiotopoulos, P. Sams, S. (2012) An overview of Twitter in the UK local government. [Online]. Available from: http://www.academia.edu/1543872/An_overview_study_of_Twitter_in_the_UK_local_government .[Accessed 19th November 2016].
- Storck, M.(2011). The Role of Social Media in Political Mobilization: a Case Study of the January 2011 Egyptian Uprising. Accessed at (http://www.culturaldiplomacy.org/academy/content/pdf/participant-papers/2012-02-bief/The_Role_of_Social_Media_in_Political_Mobilisation_-_Madeline_Storck.pdf), visited at 29/11/2016.
- Zavattaro, S. M., & Sementelli, A. J. (2014). A critical examination of social media adoption in government: Introducing omnipresence. *Government Information Quarterly*, vol.31(2), PP.257-264.

DIGITAL ANALYTICS IN BUSINESS PRACTICE – USAGE, CHALLENGES AND RELEVANT TOPICS

Darius Zumstein and Seraina Mohr

*Lucerne University of Applied Sciences, Institute for Communication and Marketing
Zentralstrasse 9, 6002 Lucerne, Switzerland*

ABSTRACT

Digital analytics has become an essential data collection and analytics method of online marketing and website management. In business practice, different digital analytics jobs, processes, roles and tasks are well established. This is shown in a survey among 504 digital analytics specialists. In 2011, in 80% of the cases, digital analytics was a subsidiary corporate task and few resources have been invested. In contrast, today every eighth company employs a full-time digital analyst. One of five companies has a small digital analytics team, and every tenth firm has a larger team of more than five analytics experts, who exploit various benefits from web data. Notably, nine of ten companies profit from web, campaign or event tracking and digital business performance management. In the e-society, research and teaching at universities and research institutes have to catch up: Most digital analytics experts and managers described educational offers as inadequate. 76% of respondents want to learn more in the field of conversion optimization, (big) data analytics, data integration, data-based management and decision support. Moreover, they want to gain insights on e-shop and digital marketing analytics.

KEYWORDS

Digital analytics, business analytics, analytics, web data, data-driven marketing, digital marketing, digital business.

1. INTRODUCTION AND PROBLEM STATEMENT

Digital analytics – the collection, measurement and analysis of web data for controlling and optimizing digital platforms and marketing campaigns – has become an essential instrument of digital business. Nowadays, it is no longer a question of whether a company does digital analytics, but how and what for.

The fields of applications of digital analytics and its potential benefits are large and manifold, as this paper shows in details. It ranges from classic web analysis, analyses of apps or customer portal usage to analyses and management of online marketing campaigns, using different digital marketing instruments like search engine advertising, display advertising, newsletter marketing or social media advertising.

Every organization who does not track its digital platforms, assets and marketing initiatives, is flying blind without instruments. Research shows, that companies who are not investing in digital analytics are less successful in achieving digital business goals (Zumstein et al. 2012, Phippen et al. 2004, Järvinen 2016).

In business practice, digital analytics has become established and professional in a short time-period. It has outgrown its infancy, as this research project clearly shows. In most European and American companies, digital analytics has a professional set up and – especially in larger and digital companies – a lot of resources are invested in digital analytics tools and manpower.

However, in many places, trainings and (further) education of high schools and universities have not yet adapted to the requirements of the companies and labor markets. This contribution answers following four research questions of information management, or business analytics respectively:

1. **Potential benefits:** In which areas of the digital value chain is digital analytics applied in business practice and what are the main benefits for companies?
2. **Challenges in digital analytics:** What are relevant application problems in digital analytics?
3. **Methods of further education:** How do digital workers and managers specialize in digital analytics, and what for? Which are important forms and types of further education?
4. **Contents of further education:** How do practitioners of digital business evaluate the existing course offers of further education and on which topics do they want to deepen their knowledge?

To answer these research questions, a detailed online survey was carried out among 504 digital analytics experts. The study results are presented in three main chapters: First, the **second chapter** explains the research methodology. Looking to the results, section 2.2 discusses the experience and manpower in digital analytics. Section 2.3 explains the departments, in which digital analytics is assigned to. Section 2.4 discusses 27 different application fields, in which digital analytics generates value for the respondents. Challenges in digital analytics are covered in Section 2.5. The **third chapter** deals with the forms and methods of further education used in digital analytics and the satisfaction with existing education offers. Section 3.3 shows the topics of interest for education and the **fourth chapter** concludes with detailed recommendations and an outlook.

2. DIGITAL ANALYTICS IN BUSINESS PRACTICE

2.1 Methodology, Proceeding and Sample

To explore the usage, benefits and the challenges of practical web analytics, or digital analytics respectively, an online survey among 740 experts was carried out for the first time in 2011. The results were published in different research papers (Meier & Zumstein 2011, Meier & Zumstein 2012, Zumstein et al. 2012).

End of 2016, the online survey was repeated by addressing 504 digital (analytics and marketing) experts, using special interest groups in professional networks like XING and LinkedIn, and by different newsletters.

2.2 Professionalization and Experience in Digital Analytics

The surveys, conducted in 2011 and 2016, clearly show that digital analytics has emerged from its infancy in recent years. It has been recognized as a new, own discipline. Job titles, roles, competences, activities and tasks of digital analysts, digital analytics managers and digital analytics consultants have established in many organizations. Thousands of new and open positions in the labor markets confirm this trend too.

In 2011, Web analytics was run as a side job in four of five companies (with 0 to 0.3 Full Time Equivalent, short FTE, in Figure 1a). In 2016, this share halved to 43%. In contrast, the proportion of companies that employed a full-time digital analyst doubled from 7% to 15%. The share of companies with an own digital analytics team (1 to 5 FTE) has even quintupled from 4% in 2011 to 19% in 2016.

In addition, digital analytics teams are getting bigger and more specialized: in 2016, every tenth company has a large analytics team with more than five FTE. Mostly of them are companies with digital business models.

Another indication of the increasing digitization and professionalization is the growing experience with and in digital analytics. The number of years, in which companies have been practicing digital analytics, has increased significantly (see Figure 1b): On the one hand, the proportion of inexperienced organizations with zero to one year of experience decreased from 17% to 13% and these ones with two to three years of experience from 33% to 26%.

On the other hand, the proportion of experienced companies with seven to nine years of experience increased to 18%, companies with ten or more years doubled from 10% to 18% (Zumstein 2011, 2017).

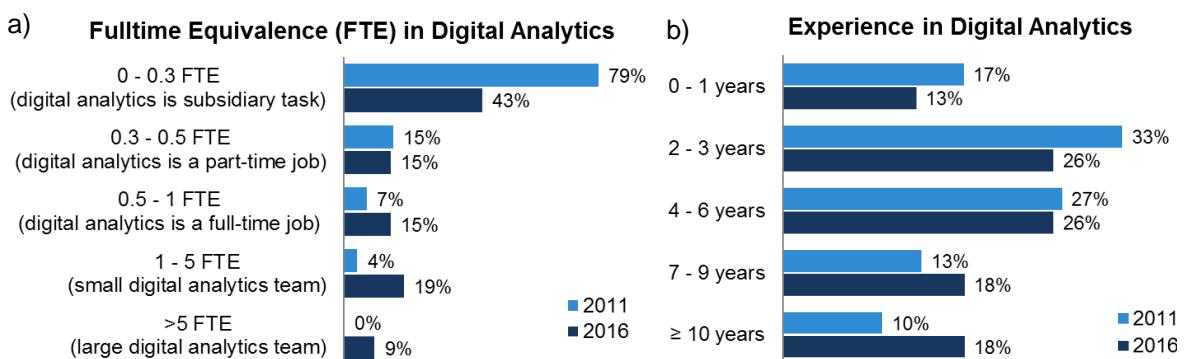


Figure 1. a) Full Time Equivalence and b) experience in digital analytics in 2011 and 2016

2.3 Organization of Digital Analytics

Researching organizational issues, the results in Figure 2 show that digital analytics is mostly a matter of (*digital marketing or digital business*, also called digital or electronic commerce, shortly e-commerce).

In business practice, digital analytics is therefore usually the task of business departments, and rarely of business intelligence (at 12%) or of information technology (IT; 7%). To allocate digital analytics in business-oriented departments makes sense. For instance, website, online shop and content managers should analyze the use of their website or online shop by themselves. Accordingly, digital marketers should analyze their online campaigns of search engine marketing (e.g. Google AdWords), newsletter marketing, digital advertising (e.g. display or native ads), or social media advertising (e.g. Facebook, Instagram or Twitter ads).

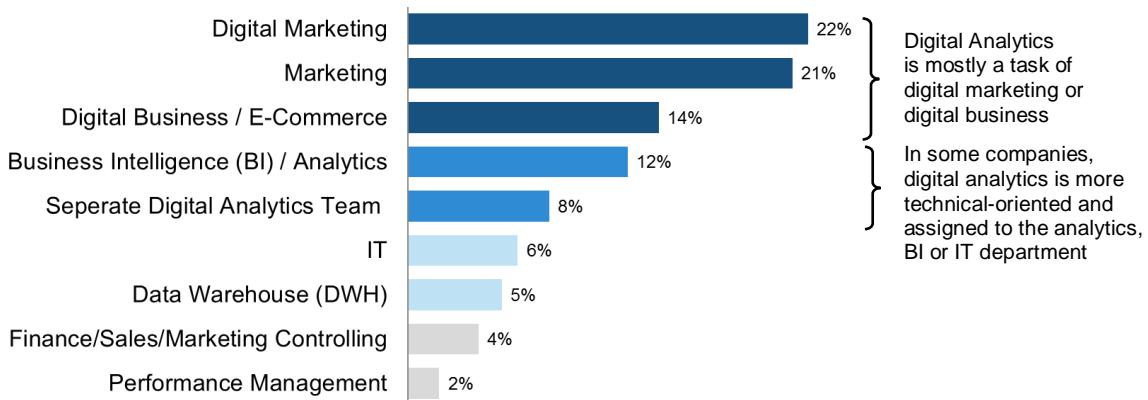


Figure 2. Departments of digital analytics

There are other reasons, why web, app or digital analytics is mostly located in business departments: in the e-society, sales and marketing are becoming more IT-, software-, and data-driven. In addition, information systems resp. software tools products like Google Analytics, Adobe Analytics or PIWIK are very user-friendly and quite easy to handle. Finally, data, metrics and key performance indicators (KPIs) of digital analytics have to create added value. In doing so, this requires close and interdisciplinary cooperation between different departments. Nowadays, direct information, communication and actions between employees are important, because decision-making and action measures have to be undertaken in a short time and in high quality.

2.4 Benefits of Digital Analytics

For 98% of the asked specialists, the benefits of digital analytics lies in *website usage analyses* (web tracking), followed by *campaign and event tracking* (86% of the respondents in Figure 3).

Nine out of ten specialists use the data to analyze *digital touchpoints* between the company and customers (event tracking), and to check the achievement of business goals. Three out of four respondents use digital analytics to analyze user behavior across multiple channels (so-called customer journey and multi-channel tracking), and for *user or customer segmentation*.

In e-commerce, every Euro is earned online, and respondents use digital analytics tools to monitor the online shop, including tracking all products, services and conversion rates. Furthermore, digital analytics helps marketing departments to analyze *digital and search engine advertising* (SEA; in 80% of the cases) and search engine optimization (SEO; 76% of the respondents in Figure 3).

Three-quarters fully or partly agreed that they use digital analytics to streamline *business processes* and to plan, decide and control (digital) business. The implementation of *mobile or app tracking*, often with Google Analytics or Adobe Analytics, is an issue for companies, around 55% of the respondents, who distribute native apps through the Apple App Store and/or Google Play.

More than half uses digital analytics tools for content or website management to improve navigation (80%), usability (77%), web content (73%), customer or user experience (70%) and design (63%) of a website and/or an app.

Surprisingly, only a minority of the asked specialists profit from digital analytics to accelerate digital transformation, to increase customer acquisition, to bind customers and to track electronic payment processes.

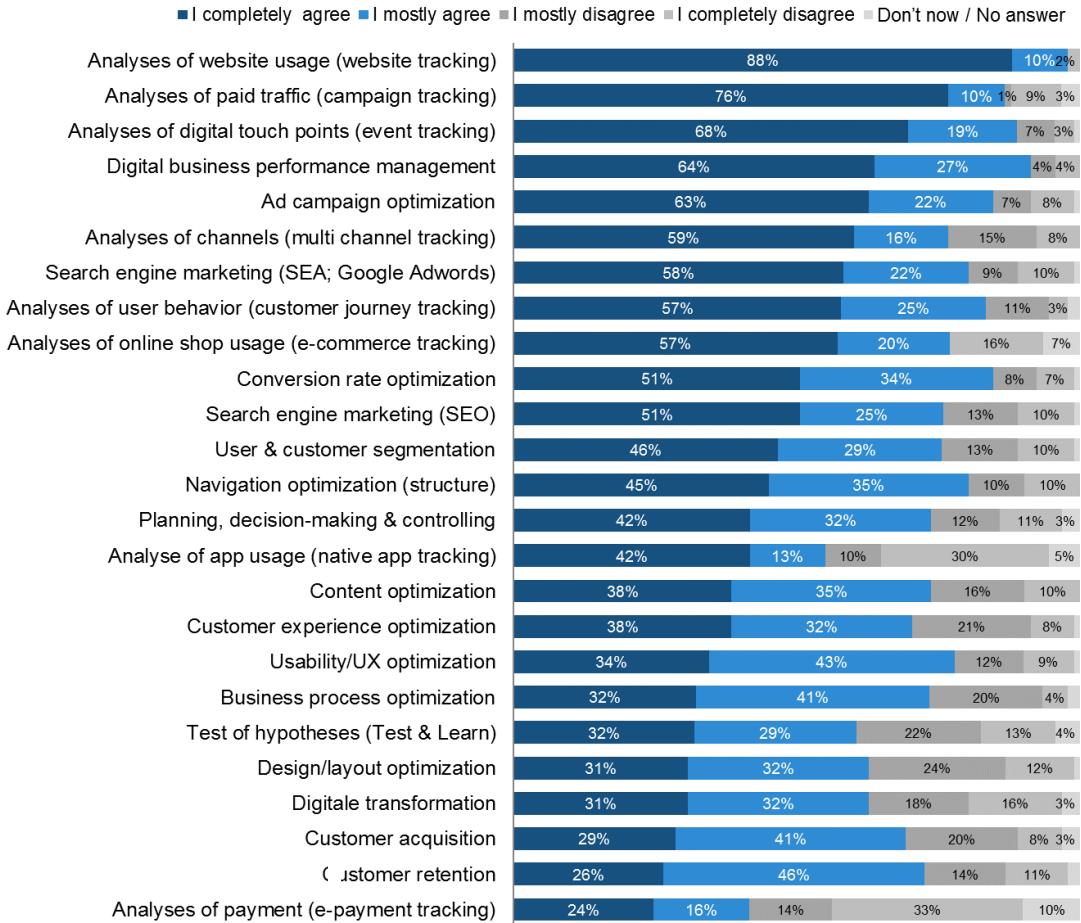


Figure 3. Benefits of digital analytics

2.5 Challenges in Digital Analytics

Digital business and analytics managers are faced with many strategic and operative challenges or big problems (see Figure 4). First and most critical is the *shortage of skilled workers*: 84% of respondents agreed that digital analytics specialists and experts are difficult to find on the labor market. Accordingly, in most countries, there is a shortage of qualified university and other graduate students in the job market. Most digital analytics expert acquired the required knowledge "on the job" through learning-by-doing (see Chapter 3).

Secondly, digital analytics is an *interdisciplinary discipline and task* with many interfaces and contacts inside and outside of the organization. No wonder, interdisciplinary cooperation becomes a challenge for 65% of the respondents. Obviously, in digital analytics it is not easy to speak the same language with all employees of IT, analytics, business, product, marketing and management and other, who all have different goals, interests, backgrounds and understandings.

Another challenge for the majority of the companies is the *integration and linking of digital analytics data* with other company data, e.g. data in database or core systems, data of marketing, customer relationship management (CRM), data warehouse (DWH) or enterprise resource planning (ERP) systems.

Data protection, *privacy and data security* challenges the half of the companies. This is especially the case, when it comes to capturing, storing and sharing personal or sensitive data stored in the cloud, or in other countries and continents. Finally, in digital analytics, data quality and the interpretation of the data is a minor problem, as the numbers at the bottom of figure 4 show.

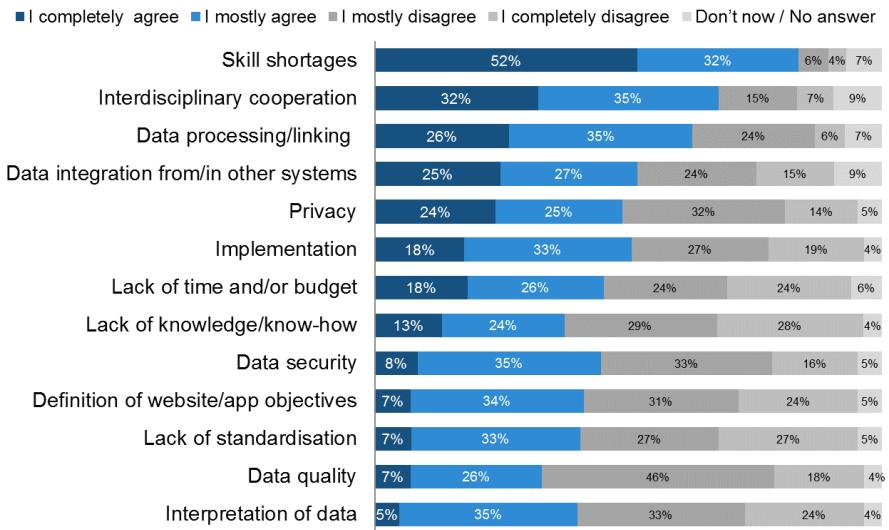


Figure 4. Challenges in digital analytics

3. DIGITAL ANALYTICS IN TEACHING AND EDUCATION

3.1 Forms of Further Education in Digital Analytics

How do specialists master the challenges and problems discussed above? One possibility is the acquisition of the necessary competences and knowledge through education. Therefore, the experts were asked which forms and methods of further education they used. Figure 5 shows a surprising picture.

First, the most important way is *learning-by-doing* in 94 cases. Nine out of ten specialists acquire the necessary know-how by themselves on the job. Since there are few courses or training opportunities in digital analytics, employees have to teach themselves. They often use (*expert*) *blogs*, which rank second with 64%.

Another important form of knowledge transfer are *textbooks*: every second uses it. For further readings see for instance Sterne 2002, Phippen et al. 2004, Weischedel et al. 2005, Burby & Atchinson 2007, Jansen et al. 2008, Kaushik 2009, Haberich 2016, 2018, Neckel and Knobloch 2016 or Hassler 2016.

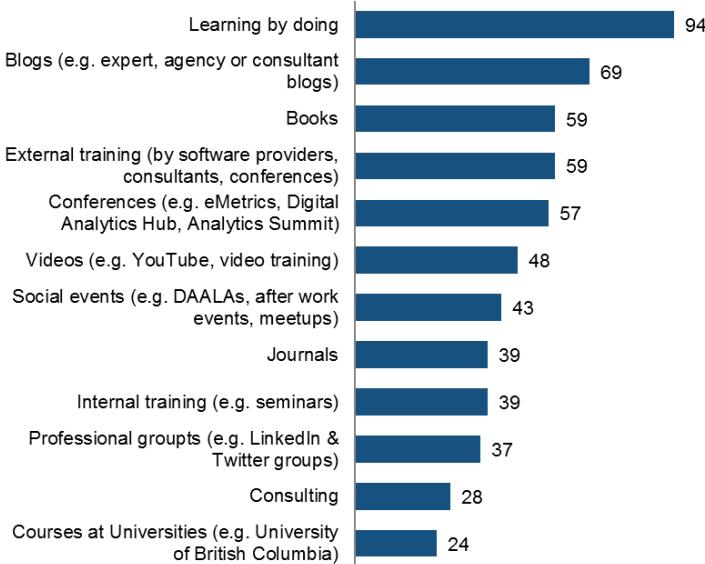


Figure 5. Methods of further education in digital analytics

Business conferences (e.g. eMetrics, Analytics Summit, Digital Analytics Hub, Adobe Summit) are visited by the half, as well as *external trainings*, e.g. from software vendors, consultants or workshops at conferences.

For many of the respondents, *videos* are very popular. For instances, this includes video training on YouTube, which provide facts, applications, tutorials and concrete solutions to every digital analytics problem.

Informal forms of further education are *events* such as DAALAs (Digital Analytics Late Afternoons), after work events, or events of associations and stakeholders, for example organized on platforms like meetup.com.

Journals and technical *magazines* like Internet World Business, Website Boosting or T3N, are read by a third. *Internal training, in-house seminars and courses* by colleagues are used by 39 of the respondents, as well as specialist interest groups of *professional social networks* like XING, LinkedIn and ResearchGate. Courses and webinars on digital analytics at *universities and high schools*, for example offered by the University of British Columbia, were used by a quarter only. However, they are becoming more important. The number and variety of courses offered, for instances CAS (Certificate of Advanced Studies) or MAS (Master of Advanced Studies) courses, are increasing.

3.2 Further Education in Digital Analytics

The research results show that education in digital and business analytics has to be improved: Two-thirds of the experts evaluate the existing training and course offers on digital analytics as poor or very poor (see Figure 6a). In a digital society, these negative evaluations and the lack of education is alarming.

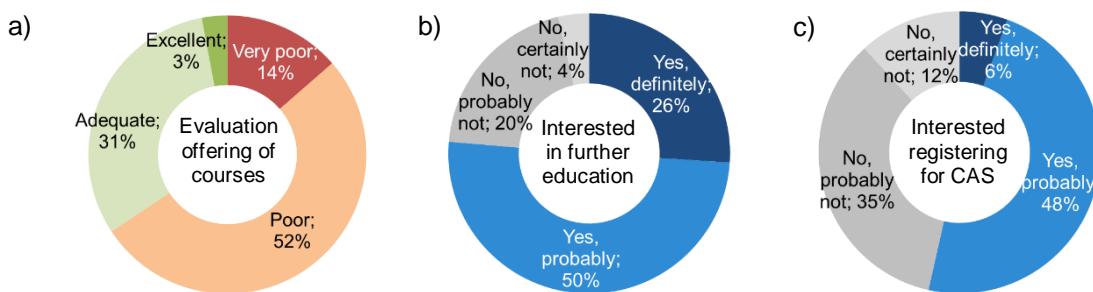


Figure 6. a) Evaluation of education offer, b) interest in further education and c) interest in CAS in digital analytics

After all, there is high attention and interest in further education in the domain of digital analytics: a quarter has great and the half hast at least some interest in further education (see Figure 6b). Over half would (probably) attend a CAS course on digital analytics at the university of applied sciences (in Figure 6c).

3.3 Topics of Further Education in Digital Analytics

At the end of the survey, the specialists were asked, which *learning content* they would prefer for their personal education in the field of digital analytics. *Conversion optimization* is a hot topic, i.e. all methods and counter measures, which leads to higher goal achievement of digital marketing and digital business.

The most popular conversion goals mostly are purchases (transactions) and leads (contact data of potential customers). In marketing and user experience (UX), a new and own discipline has been established that deals with conversion optimization. Another important field of education is *big data analytics* and *data integration* of digital analytics data with other data of the company (such as ERP, product and customer databases mentioned above). For this purpose, raw data from digital analytics systems like Google or Adobe Analytics are extracted from the cloud via APIs (Application Programming Interfaces) and sFTP (Secure File Transport System) directly to the server of a company. Web and big data can be integrated using common information technologies, such as Hadoop, and processed with program languages like Python and tools like R or SPSS.

On the third position is a classical topic of business informatics: *management and decision support*. Data, prepared and presented by analysts, data scientists and data managers, should help business managers and executives to make better decisions and to control businesses.

The respondents who work in the field of e-commerce are particularly interested in the topic *shop analytics*. Derived from digital analytics data, analyses, counter measures and concrete actions help marketing, online shop managers and decision makers to optimize their business (results).

Many of the respondents were also interested in further education on marketing analytics topics, such as *campaign analysis and attribution modeling*. Another field of interests is advanced web and app analytics, i.e. all methods and processes to gain insights and know-how about website and app usage.

The topics of *web mining*, which uses statistical analysis methods, are attractive for many respondents too, as well as *A/B and multivariate testing*. In these tests, different types of design, text, banner or website components are tested and optimized for the highest possible click and conversion rate.

With the increasing complexity and specialization in analytics, executives are faced to complex analytical facts, processes, relationships and values of different data, metrics and key figures. Therefore, data has to be visualized and internally communicated clearly, so that the executives understand data-driven facts and stories. Consequently, some of the digital analysts want to educate themselves in *visual analytics and reporting*.

Finally, in organizational, legal, technical or social topics, the respondents were less interested.

4. CONCLUSIONS AND OUTLOOK

Technological and economic conditions as well as business (best) practices of informatics are developing fast and dynamically. Many universities and high schools in different countries seem to insufficiently aligning their teaching and research focus, content and methods with the needs of professionals.

To provide study programs with the knowledge, competences and expertise which are demanded by the digital economy, the authors derived *following recommendations* for further education at universities in the fields of digital analytics and data science:

- 1. Interdisciplinary collaboration:** In the near future of the information and knowledge society, faculties, departments, institutes and competence centers of higher education should cooperate in interdisciplinary ways. Isolated *working, teaching and research in silos* probably does not work anymore and generates less impact and value for the science, society and the economy.
- 2. Cooperation with companies:** Particularly in the field of digital or business analytics, academic and practice-oriented research and teaching needs to be checked to their *practical relevance*. Universities, higher education and research institutes should not conduct basic and applied research only, but also work together with small, medium and large enterprises from business practice. Consequently, practice-relevant and application-oriented research and solutions emerge. Forms and ways of cooperation may be manifold, ranging from joint ventures, lectures, conferences and research projects, studies, alumni, advanced studies programs, funding programs and agencies, up to sponsored chairs and institutes.
- 3. Practice-oriented theses:** Term, bachelor or master theses and dissertation/PhD projects at universities should be evaluated for their practical relevance. When students work in business practices during writing their thesis, this often produce more relevant results and value for companies, and students gain an understanding and know-how about what is really needed in companies. In addition, academic professors, lecturers and assistants receive inspiration and feedback regarding the value and practical relevance of their teaching and research when Theses are created in cooperation with enterprises.
- 4. Practice-oriented teaching:** Particularly in the field of information and marketing management, it makes sense to hire lecturers at universities and high schools in a practice-oriented manner. This applies especially to digital business and analytics, where knowledge, skills and know-how must be acquired through active making, developing and testing, not just passive listening and reading. Practical orientation is increased by conducting student presentations (e.g the concept of flipped class room), practical group work, exercises, practical projects, case studies, labs, lectures by practitioners and application-oriented open-book exams.
- 5. Development of existing and new education offers:** Regarding curriculum, modules, content, teaching and research methods, Bachelor, Master and PhD Programs at universities and all their courses, lectures or seminars must be practice-oriented. This means that content, goals and the defined competences in degree programs must be reviewed annually for their *quality, actuality and practical relevance*.
In a digital age, practical requirements for graduates are rapidly changing and evolving. In higher education, professors and teachers of institutions, programs and courses must always stay on the ball and focus on the technological, methodological, procedural and skills, which are required on the job market.

At innovative, agile, flexible, open-minded and growing universities and high schools, it may be easier and more effectively to redesign, discuss and introduce new courses aligned to the labor market, rather than in old, static, conservative and bureaucratic institutions.

6. **Rapid implementation processes:** The time-to-market of new technologies, applications, (digital and physical) products and services has grown rapidly, not only in the private sector. Universities also need to rapidly bring their education programs to education systems. When Universities develop new courses, modules or courses, they should be quickly approved and implemented. Long-lasting bureaucratic and administrative processes and barriers have to be avoided. This applies in particular to further education, where participants have to pay high course or study fees, e.g. for MAS or EMBA courses. These students may have very high expectations on the timeliness, quality and practical benefits of the course content.
7. **Guaranteeing higher education standards:** When university programs have to be quickly designed and implemented to meet the expectations of the dynamic education market, the quality of teaching should be guaranteed. For example, frameworks like EDISON in the field of data science will help program leaders in designing, introducing and quality assuring of new university courses (EDISON 2018).
8. **Agile didactics:** Even in the case of well-functioning and evaluated courses and education systems that meet participants' expectations and requirements, it is becoming increasingly important for professors, lecturers, and assistants to be agile and flexible in meeting the needs and interests of students. This enables an advanced education in analytics that the future labor market and e-society demands.

REFERENCES

- Burby, J., Atchinson, S. 2007. Actionable Web Analytics: Using Data to Make Smart Business Decision, Wiley, New York.
- EDISON, Building the data science profession, Available: <http://edison-project.eu>, Accessed the 5th of January 2018.
- Fasel, D., Meier, A. (Hrsg.): Big Data: Basics, Systeme & Usage Potentials (in German), Springer, Berlin (2016)
- Haberich, R. (Ed.), 2016. Future Digital Business, mitp, Heidelberg
- Hukemann, A., 2004. Controlling im Online Sales – A Performance-oriented Approach for Online Shops (in German), Logos, Berlin.
- Huiizingh, E., 2002. The antecedents of Web site performance, In: European Journal of Marketing, Vol. 36, No. 11/12, pp. 1225-1247.
- Jansen, B., Spink, A., Taksa, I., 2008. Handbook of Research on Web Log Analysis, Idea Group, Publishing, Hershey.
- Järvinen, J., 2016. The use of digital analytics for measuring and optimizing digital marketing performance, Dissertation, Jyväskylä University School of Business and Economics.
- Kaushik, A., 2009. Web Analytics Wiley, Wiley & Sons, New York
- Hassler, M. 2016. Digital and Web Analytics: Analyzing Metrics, Understanding User Behavior, Optimizing Websites (in German), mitp, Frechen.
- Meier, A., Zumstein, D., 2010. Web controlling – Analysis and Optimization of the Digital Value Chain with Web Analytics (in German), In: Multikonferenz Wirtschaftsinformatik (MKWI 2010), Göttingen, Germany, pp. 299-311.
- Meier, A., Zumstein, D., 2012. Web Analytics & Web Controlling – Web-based Business Intelligence for Success (in German), dpunkt, Heidelberg.
- Meier, A., Zumstein, D., 2016. Business Analytics, HMD: Praxis der Wirtschaftsinformatik, Vol. 3, Springer, Berlin.
- Neckel, P., Knobloch, B. 2016. Customer Relationship Analytics – Practical Application of Data Mining in CRM (in German), dpunkt, Heidelberg.
- Phippen, A., Sheppard, L., Furnell, S., 2004. A practical evaluation of Web analytics, In: Internet Research, Vol. 14, pp. 284-293.
- Sterne, J., 2002. Web Metrics, Wiley, New York.
- Weischedel, B., Matear, S., Deans, K., 2005. The Use of Emetrics in Strategic Marketing Decisions: In: International Journal of Internet Marketing and Advertising, Vol. 2, No. 1, pp. 109-125.
- Zumstein, D., 2017. Digital Analytics – Data a Part and Driver of Digital Business (in German). In: Nappo M. (Eds.): Digital Yea(h)rbook 2017, Center for Digital Business, HWZ, buch & netz, Zurich, pp. 237–250.
- Zumstein, D., 2017. Digital Analytics in Action – Interdisciplinary Cooperation in the Information Management Practice (in German), In: Portmann, E. (Ed.): Wirtschaftsinformatik in Theorie und Praxis, Springer, Berlin, pp. 85–105.
- Zumstein, D., Meier, A., Myrach, Th., 2012. Web Analytics – Empirical Research on the Usage, Benefits and Challenges (in German), In: Multikonferenz Wirtschaftsinformatik (MKWI 2012), Braunschweig, pp. 917-929.
- Zumstein, D., Gächter, I., 2017. Digital Analytics – Implementing Digital Strategies and Controlling with KPIs (in German), In: Meier, A., Zumstein, D. (Eds.): Business Analytics, Volume 3, Springer, Berlin, pp. 371–388.

CONCENTRIC DIVERSIFICATION: GROWTH WITHOUT LOSING COMPETITIVENESS FOR A SMALL IT FIRM

Erick A. Arroyo-Blanco¹, Alexis E. Boza-Monge¹ and Roberto Cortés-Morales²

¹*Information Systems Master Program, Instituto Tecnológico de Costa Rica*

²*Computer Graduate Department, Instituto Tecnológico de Costa Rica*

San José, Costa Rica

ABSTRACT

This article presents a study case applied to the company “RidgeRun Embedded Solutions”. Previous research has described the core value activities system of this firm as a base of an e-business model of such system. Deepen the analysis, we keep on refining the model, but envisioning an expansion of its capabilities and position on the embedded systems market. Using the concentric diversification approach, such expansion is intended to preserve core capabilities without compromising coherence and consistency of activities. By applying the analysis of the industry’s forces, an activity system for an enhanced e-business model was proposed which responds to the expansion of RidgeRun.

KEYWORDS

E-Business, Enterprise Architecture, Embedded Systems Industry, Industry Forces, Strategy, Diversification

1. INTRODUCTION

RidgeRun was founded in 2003 in Costa Rica by two former Hewlett Packard engineers (RidgeRun, 2017). They envisioned a company that could create products for a specific sector in the software market. The main engineer workforce is located in Costa Rica mostly formed by electronics and computing engineers. Since its inception, the company has specialized in the development of software for embedded multimedia systems. RidgeRun develops its own software products that can be used in different systems on chip (SoC) and also offers engineering services to provide advice and help to other companies in the process of developing their multimedia product. During its 14 years in the market, the company has offered its products and services to medical, automotive, aeronautical, military and videogame industries.

2. INDUSTRY CONTEXT

The importance and trends of this industry are described in (Global Industry Analyst, Inc., 2015) where a 240 billion dollars market is anticipated by 2020. In this context, embedded systems’ designers are under pressure to reduce the cost and size of electronic devices. At the same time, industry requires high robustness on the products since they are on the core of solutions of the smart technologies for communication, transport and others. In turn they need to improve the time to market and the overall reliability of their systems. Based on this need, RidgeRun wants to complement its current competitive advantage with the software support it provides to the newest system on chip platforms in the market, by creating and offering a set of carrier boards for each of these platforms which can be used directly in multimedia products. The description of the system activities has been already analyzed (Rodríguez-Picado et al., 2017). In such article, authors found that RidgeRun emphasizes on research activities as a main source of client value. As a result, customers benefit from faster time to market, reduced total cost of ownership, product longevity, risk reduction during development and fully integrated applications within a single package. However, as a small firm that response to big industry customers' needs (such Texas Instruments or Nvidia), a question arises as how RidgeRun can expand its capabilities for other type of products without eroding current capabilities. As

described in (Rodríguez-Picado et al., 2017), RidgeRun usually designs and implements the solution directly to such customers. But what if the firm can adopt other position into the supply chain of the industry becoming a buyer and integrating supplier parts for its products. In (Lambert & Cooper, 2000), authors describe the complexities of adding levels in the Supply Chain both in the buyer and supplier perspective. As it can enhance the competitive position of RidgeRun by becoming an important buyer, new challenges arise on managing the relationships for buyers and suppliers to provide products on time with the quality required.

3. THEORETICAL APPROACH

However the firm looks to enhance its competitive position, it adopts a perspective where the coherence of the current activities system is not largely impacted. Coherence (as well as reinforcing activities and effort optimization) are the type of links that are described for achieve a distinctive activity system (Porter, 1996). Such controlled expansion based on the core activities is what is described as “concentric diversification” (Rijamampianina et al., 2003). Both perspectives, coherence and expansion are the main drivers for adapting Ridge Run’s system. e-Business as a classic concept integrating ICT, customer value, infrastructure and financial sustainability (Osterwalder & Pigneur, 2002) can be used to design such activities system.

4. METHODOLOGY DESCRIPTION

Methodological strategy was based on a case study research (Yin, 2003). Here we want to analyze a specific organizational context without pretending generalize the findings, but willing to stimulate the research on this kind of area providing knowledge that can be applied to small firms in an industry such as embedded systems. A structure industry analysis was made using the approach of the Porter’s Five Forces (Porter, 2008). Then the respective value activities were defined based on the Porter’s system activities model (Porter, 1996). Finally, an e-business proposal is designed based on the concepts of enterprise systems and enterprise architecture of applications.

5. CASE ANALYSIS

The case was conducted applying the different models of competitive assessment described by Porter.

5.1 Industry Forces Analysis

The first analysis performed concerns the structure of the industry of this new growth opportunity, taking as a starting point the five forces described by Porter. For each of them, we analyze how they affect RidgeRun and what characteristics are advantageous or disadvantageous. It is necessary to recall that if forces are intense, less competitive is the industry. However even if some forces are intense, as a whole the industry could be considered profitable. The figure 1 shows the results of this analysis.

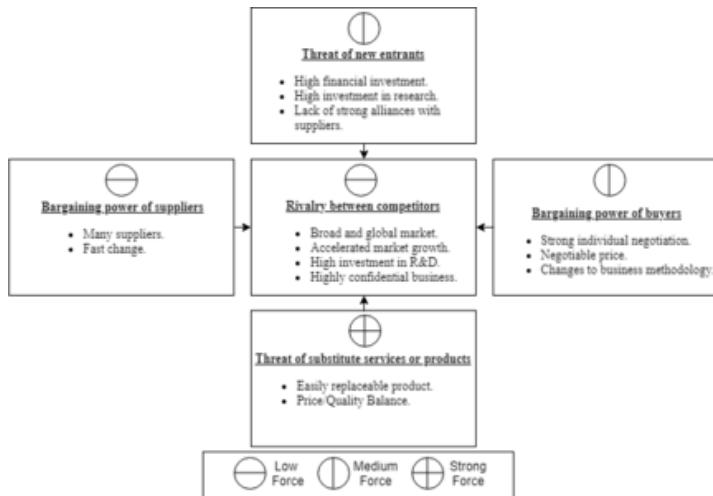


Figure 1. RidgeRun's five forces analysis

5.1.1 Threat of New Entrants

This force was assessed as “medium” because important industry barriers related to massive financial investments and technological knowledge that a new entrant must incur to enter the market. In addition, due to its global scale of operations, RidgeRun gains cost advantages from economies of scale and this advantage is not available to new entrants, at least during the initial stages of operations. Established companies have proprietary products and knowledge that can only be generated with a large number of engineer hours. On the other hand, there are certain risks with new competitors with ideas of innovation for technological products and services who could ensure the necessary financing of investors and venture capitalists to enter the industry. In addition, they can negotiate with suppliers the access to confidential information necessary for the development of new products.

5.1.2 Bargaining Power of Suppliers

Companies within the hardware industry are able to switch between their suppliers easily, especially when they own hardware specifications. Changing suppliers does not represent a high cost in this industry. On the contrary, companies aim to obtain from suppliers what best fit both for their economic and productive needs. Large number of suppliers within the industry allows availability and contingency among them. On the other hand, the power of the suppliers is greater when negotiating access to confidential information of the system on chip. An opportunity that companies can exploit, both to maintain a good relationship with suppliers and keep the intensity of this force low, is to promote the use of supplier’s SoC.

5.1.3 Bargaining Power of Buyers

This force is considered as medium within the industry. For RidgeRun, client's individual bargaining power is a very strong force, since their buying capacity represents a significant amount of revenue for the company. Clients sometimes demand certain requirements, which involve making changes to RidgeRun's methodology. This increases the power that buyers have over the operation of the company. Depending on the number of devices requested, it would be possible price negotiations between the buyer and the company.

5.1.4 Rivalry between Competitors

The market of embedded systems is broad, global and in accelerated growth. These characteristics decrease the intensity of this force. To profit from this force depends on RidgeRun investment in Research and Development, innovating with new products. No evidence of copying among companies is observed.

5.1.5 Threat of Substitute Services or Products

Hardware is an easy replaceable product. Usually customers will look for a balance between price and quality. Most of times and due to the advancement of technologies is not very difficult to find new products.

However the idea proposed covers a complete customer support package, a set of high quality features of hardware and software. RidgeRun should manage this strong force by making significant investments in Research and Development, which would allow it to continue developing new and unique products building significant brand loyalty (Rodríguez-Picado et al., 2017).

5.2 Activities System

Based on the previous Porter's Five Forces Analysis, an appropriate activity system was defined to take advantage of this growth opportunity. Figure 2 shows the primary and secondary activities accordingly with the Porter's Value Chain.



Figure 2. RidgeRun's value chain

The proposed system is composed by 6 primary activities, Research and Development, Purchase Management, Strategic Product Design, Production and Logistics, Marketing and Sales and “End to End Services”.

5.2.1 Research and Development

The Research is the core of the RidgeRun activity system. This is the primary activity that gives greater value and differentiation to its products. Intensive and pertinent research aims to be always at the forefront of the embedded systems market keeping the company unique from rivals with a sustainable competitive advantage. Under the framework of research and development we can analyze a set of activities that define the flow to follow within the company. Figure 3 shows these activities.

5.2.2 Purchase Management

As a consequence of the new package that the company intends to offer to its clients, purchase management becomes a primary activity. This activity must be analyzed from two main perspectives: one is the management of the carrier boards suppliers and the other is the access management to confidential information of new SoC. For the management of carrier boards suppliers, the following general procedure is recommended: 1) Identification of needs. 2) Identification of suppliers 3) Selection of suppliers 4) Supplier's negotiation 5) Generation of purchase orders and/or contracts with suppliers 6) Follow-up of the purchase order 7) Control of receipt of goods 8) Payment to the supplier 9) Renegotiation with suppliers.

SoC's suppliers generally offer public information about the characteristics of the new SoC, evaluation boards, and sample software. This information is very important during the research and evaluation stage of the new technology. However, being a homogeneous resource and public access is not useful information to create differentiated products by companies. At this point it is when it becomes necessary to enter into negotiations with the supplier of the new SoC. In fact, one of considerations to determine if it is feasible to support the new SoC, is how easy it is to negotiate with the supplier the access to such information.

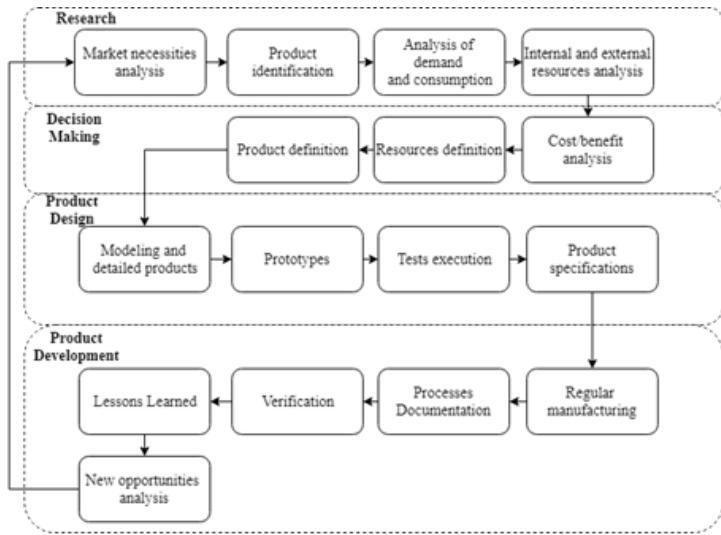


Figure 3. RidgeRun's research and development activity

5.2.3 Strategic Product Design

Within this activity system, the design stage has a great weight. From this activity the company must define where the effort and resources should be spent. It is always important to be cautious in what the company is spending engineer hours. This activity is based on the information generated by the Research and Purchase Management activities. Decisions must be made in these other activities, before proceeding with the Product Design stage. The product design stage should be divided into two sections, which are: the design of the software products that will be developed for the new technologies and the design of the specifications of the carrier boards for the new SoCs.

5.2.4 Production and Logistics

Part of the challenges that the company has to deal with the offer of this new package is to make changes in its production activities. Software development must be supported and a set of activities must be added to allow the development of the carrier boards and the logistics involved in the management of these physical resources. In addition, for the success of the package, synchronization must be achieved between the launch of the developed software products and the carrier boards for the new SoC. This synchronization must be carefully planned at the stage of Strategic Product Design. The idea is to parallelize the two main processes. RidgeRun should develop the software as usual and the supplier company specialized in hardware is responsible for the development of the carrier boards.

5.2.5 Marketing and Sales

The way RidgeRun make marketing is through innovation. To be the first company that shows demos using edge technologies is how the company advertises its brand. Offering a full package that contemplates the software and the carrier boards for the most recent SoC's is also a way to do marketing. The idea of this package is to retain current customers giving a more complete service looking also to attract new ones.

5.2.6 End to End Services

RidgeRun should support its customers in each of the stages of the development of their products. The objective must be that clients perceive an integrated experience of its services. Focus on customer satisfaction from the first contact, as well as during and the end of the contract should be the mission of each of the team members. The main idea of providing an integrated service during the whole cycle is to create customers loyalty with the brand, using the degree of satisfaction as a network to attract new customers.

6. SOLUTION PROPOSAL

6.1 Business Architecture

Business architecture defines how systems can organize the processes of an organization based on their final goals. Thus resources are optimized and also ensure compliance with predefined objectives. The idea of using this approach is to align the different information technology components of the organization, based on a strategic vision making them useful for decision activities. Based on the value proposition of "reduce the time to market", the architecture proposed for RidgeRun is composed by 3 fundamental pillars of information: ERP / SCM, CRM and KM. Using the concept of Enterprise Architecture of Applications (O'Brien & Marakas, 2008), Figure 4 shows these business architecture pillars. Because the proposed activity system does not contemplate external distributors, the need to maintain a Partner Relationship Management (PRM) system was not considered.

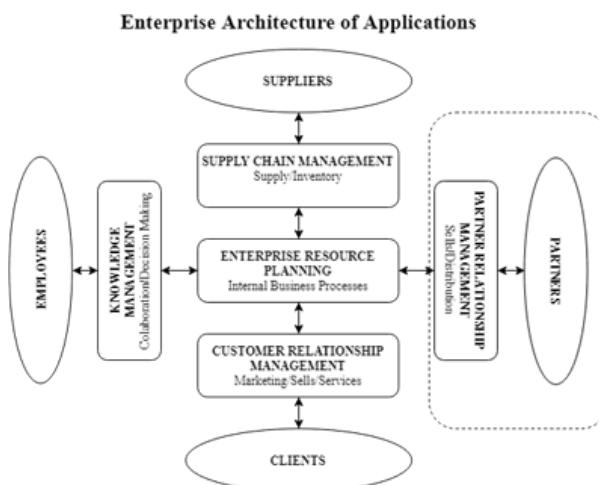


Figure 4. RidgeRun's business architecture proposal based on (O'Brien & Marakas, 2008)

6.1.1 ERP/SCM

The main axis of the architecture is undoubtedly the ERP, which provides control to the business processes of the organization, and helps to solve multi-platform IT issues and improve performance with the availability of information in real-time and faster response time. Furthermore, ERP would increase the interaction between RidgeRun employees and external organizations and can streamline workflow. It is clear that the implementation of an ERP system entails a considerable investment in time and money. However, a correct implementation based on the analysis of the processes of the organization and also under the premise of unifying the systems of the organization and the processes of this so that the information flows through all the departments will contribute to the improvement of the organization. Mainly the organization will be able to benefit in several aspects: 1) Supplier: Improving supplier performance (tying them to the ERP), 2) Operations and Logistics: Generating product differentiation, 3) Financial: Finance information accuracy and faster decision making capability will improve corporate governance and transparency, 4) Sales and Marketing : Improving interaction with customers, 5) Human Resources: Improving learning, training and development, 5) Customer: Increasing the interaction between organization with customers and suppliers in a long term plan.

Reinforcing the ERP, the SCM is necessary to have a mean for controlling the suppliers with little interference for the organization. The value promise of reducing the time to market, RidgeRun must take responsibility for the delivery of their products not only at the best price but also at the right time, keeping inventory levels to a minimum. This can only be carried out with the control of the chain. The SCM system allows tracing the passage of the pieces (traceability) between the different participants of the supply chain. It

includes the management of the processes of the chain of business value from the moment it is supplied, to the point of consumption. The lack of an information system that executes an efficient transfer of information between the parts of the logistics network makes the transfer of physical products more expensive, more complicated and slower. RidgeRun must support its SCM under the premise of just-in-time manufacturing (JIT), in which the items are created to meet the demand, they are not created in excess or before the need.

6.1.2 Knowledge Management

The core of the RidgeRun's activities system is the research. There is evident of the need to manage the company's knowledge in an appropriate manner. An adequate knowledge management at RidgeRun has an impact on both the effective operation of its activities and strategic decision-making (Seidler-de Alwis & Hartmann, 2008). Internally it is important to correctly manage this knowledge, in order to:

- **Avoid redundant effort in research and development activities.** It is very important to document the knowledge generated during the research and development activities. Many times different clients need similar products or services. Through an appropriate knowledge management, the engineer-hours can be used more efficiently. For example, avoiding the repetition of procedures to obtain the knowledge that was previously generated in another project. From the point of view of research, knowledge management prevents different engineers from investing their time researching on the same subject. In addition, one can easily determine which investigations. To support decisions based on knowledge is relevant to RidgeRun.

- **Reduce the tacit knowledge of employees and transform it into explicit and shared knowledge within the company.** Normally, each engineer is working on different projects, serving different clients or participating in R&D projects. This type of operation of the company causes its engineers to specialize in some technology and generate a lot of tacit knowledge that must be converted into explicit knowledge. The general culture of RidgeRun tries to avoid that the knowledge concentrates in few people. Knowledge should be shared among to ensure that engineers have basic knowledge about the pertinent technologies.

- **Improve the accessibility of information.** One of the main benefits that RidgeRun would obtain with the implementation of a KM system is to improve the accessibility and visibility of the knowledge generated on new technologies. Both developers and project coordinators may review previous work on specific technologies, determining implications and recommendations. It also would allow improvements to the process of training new members, showing them explicitly the type of work developed into the company.

- **Increase the quality of the products and services.** Managing knowledge aims to increase the quality of products and services. Regarding the services, it can provide better advice to the clients in their projects, more accurate recommendations of which type of technology best fits the multimedia product that the client wants to develop and provide better estimations of cost, time and complexity of a project. Regarding the product development, engineers can use prior knowledge to determine what are the best development and testing techniques when working on a technology.

A KM system at RidgeRun would become a tool for strategic decision-making. RidgeRun invests constantly in research and development projects, carried out both by employees of the company and by students through partnerships with the academy. From these projects, a large amount of information is generated about new technologies relevant to the company. As previously determined in the analysis of Porter's forces, the threat of substitute products is one of the strongest forces in the case of RidgeRun. So it is necessary to constantly evaluate and analyze the information generated in the research projects to determine market trends, what technologies the company must invest more resources, using this intensive force for the benefit of the company.

6.1.3 CRM

Providing a comprehensive service where the client is guided from the start of their project until the end of this, arise the need to have the support of a CRM to meet this premise by coupling all departments under this proposal. Allowing the formation of personalized relationships with RidgeRun customers, with the objective of improving customer satisfaction and maximizing benefits of using a CRM with goals like: 1) to identify the most profitable customers and offer them the highest level of service, 2) to provide employees with the information and processes necessary to know the clients, 3) to have the necessary capacity and maturity in the process to understand and identify their needs and forge relationships between the company and its customer base, 4) to accurately log and record sales activities from start to finish, ensuring that active sales orders and potential leads are all tracked and allocated in the most efficient manner possible.

7. CONCLUSIONS

Based on the results of the Porter's Five Forces Analysis, RidgeRun must continue investing in research and development projects to maintain its pioneering and innovative company image in the embedded systems market. But in order to approach this new market opportunity the company must focus efforts also on marketing and sales encouraging the idea of the "end to end" service and taking possession of some point of the supply chain. The new alliances should enable RidgeRun to open new opportunities for its software and hardware. The enterprise architecture of e-Business would allow complying with the established positioning by using an ERP as a central system reinforced by a KM and a CRM. ERP must be a central focus for the organization and improvement of the internal processes of the company. SCM should support the supply chain and never lose sight of the initial objective (reduce the time to market) and the premise of "Just in time". KM system has an impact on both the effective operation and strategic decisions of RidgeRun. The best use of the information generated through research and development projects can only take place through an adequate knowledge management.

Through the analysis carried out, RidgeRun's incursion to the carrier boards market is considered feasible, as long as the company uses the proposed activity system. In summary, the Concentric Diversification proposed points to maintain Research as the core of the activity system adding the provision of an End to End service to its clients.

REFERENCES

- Global Industry Analyst, Inc., 2015. *The Global Embedded System Market*. [Online] Available at: http://www.strategyr.com/MarketResearch/Embedded_Systems_Market_Trends.asp [Accessed 28 February 2018].
- Lambert, D. & Cooper, M., 2000. *Issues in Supply Chain Management*. *Industrial Marketing Management*, (29), pp.65-83.
- O'Brien, J.A. & Marakas, G.M., 2008. Introduction to Information Systems. 14th ed. McGraw Hill.
- Osterwalder, A. & Pigneur, Y., 2002. *An eBusiness Integrated Ontology for Modelling eBusiness*. In 15th Bled Electronic Commerce Conference. Bled, 2002. Association for Information Systems.
- Porter, M.E., 1996. *What is Strategy?*. Harvard Business Review, (Nov-Dec), pp.61-78.
- Porter, M.E., 2008. *The Five Competitive Forces That Shape Strategy*. Harvard Business Review, (1).
- Rijamampianina, R., Russell, A. & February, Y., 2003. *A Framework for Concentric Diversification Through Sustainable Competitive Advantage*. *Management Decision*, 41(4), pp.362-71.
- Rodríguez-Picado, C., Cortés-Morales, R., Lopez, R. & Pizarro, S., 2017. *An e-Business Model for Small Technological Firms: the case of Ridge Run*. In 15 e-Society Conference. Budapest, 2017. IADIS.
- Seidler-de Alwis, R. & Hartmann, E., 2008. *The use of tacit knowledge within innovative companies: knowledge management in innovative enterprises*. *Journal of Knowledge Management*, 12(1), pp.133-47.
- Yin, R., 2003. *Case Study Research Strategy*. 5th ed. London: SAGE.

THE DEVELOPMENT OF B2C E-COMMERCE IN MACEDONIA: CURRENT SITUATION AND FUTURE POTENTIAL

Shkelqim Nazifi and Adrian Besimi
South East European University, Tetovo, Macedonia

ABSTRACT

The main intention of this paper is to study the development of e-commerce on small sized companies (SSCs) in Macedonia. There are millions of people online as potential customers in the online market. The most important thing for companies is to understand the customers and their needs. This paper gives information about the current situation as well as future potential for e-commerce implementation for Macedonia by analyzing the current situation, legal aspects, financial aspects, technical barriers and potential for the future.

KEYWORDS

E-commerce technologies, influence, obstacles, benefits, small sized companies, Macedonia

1. INTRODUCTION

E-commerce refers to all electronically performed transactions between the company and other party. The new way of e-commerce has changed the global economy in many aspects and its power is increasing year by year (Chris G. Christopher, 2011). The impact on developing countries could be even stronger than that on developed countries because the scope for reducing inefficiencies and increasing productivity is much larger in the developing countries (Machado, 2011). The reasons for carry out an e-commerce strategy can differ from company to company. The leading individuals in the small sized companies need to rise above the obstacles and understand the benefits from e-commerce presence (BBC, 2014). E-commerce can be of great importance to small sized companies (SSCs) by enhancing its advancement and development particularly in developing countries. This is inspired by the benefits of IT possibilities in reducing costs of business processes. Companies in developing countries are not faced with same problems and challenges as in developed countries. The development is quite low since they are still slow to follow the emerging IT solutions (Lacka, 2014). The scale of commercial activities taking place into Macedonian market and the rapid development of IT makes it imperative for adoption by Macedonian SSCs, which could contribute on economic growth and stability. The globalization pressure on SSCs has compelled many of them to implement e-commerce solutions to survive the increasingly competitive supply networks. Macedonian companies started to use e-commerce for many purposes, like enhancing efficiency and reduction of costs, and for providing better services to customers (USAID, 2010). This research is significant in terms that will highlight the factors influencing the SSCs.

2. B2C E-COMMERCE IN THE LITERATURE

E-Commerce can be any transaction over a computer network, which includes the transfer of ownership or right to use goods and services, including service and support after sale (Gao, 1999). The company performance supported by e-commerce can gain in increasing the profitability, growing market share, improving customer related service and quicker delivering of products or services (Watson R, 2008). B2C e-commerce is the most popular application of e-commerce. Online businesses will stock goods, post them online and sell directly to the customer. The customer will reach the online store, browse and buy the items

that are posted. When this occurs, the operational team is notified and they are starting to prepare the sold products for shipping to the customer's house. In the Figure 1, below is represented the B2C business model that most of the bigger online stores are focused on.



Figure 1. Example of B2C model (Image credits: Netonomy.com)

B2C is not just for “big players”, many small companies and startups are also using this business model (Sundheim, 2017). In B2C there are five major activities involved in conducting it. It starts with (I) info sharing, using different applications and platforms to distribute information to customers like online advertisements, e-mail, company web site, online catalogues, message board systems. Second comes (II) ordering where the customer may use electronic e-mail or online forms available on the company's online store. Next comes (III) the payment which can be realized with credit cards, electronic checks, and digital cash. After the payment is realized arrives the (IV) fulfilment, responsible for physical delivery of product or service from the seller to the customer. Finally the service (V) and support where e-companies must provide timely qualitative support to maintain current customers and attract new (METU, 2017). Living in digital age, it is no longer enough to provide just a functional online store, today's e-companies are presenting multiple buying opportunities. The increased usage of mobile devices requires the B2C shops to work on corresponding platforms. Being price competitive and use technology that is flexible and able to accommodate innovation is crucial (Crimsonwing, 2014). E-commerce offers opportunities to tap into global markets. Any sized company beginning an online presence, immediately enters into the global marketing. Small companies can sell to the entire Internet world and maximize their impact (Watson R, 2008).

3. ANALYSIS OF MACEDONIAN E-COMMERCE AND IT INFRASTRUCTURE

E-commerce stands as an extraordinary potential to improve the global market efficiency and is important to integrate these developments into the developing countries (UNCTAD, 2015). A considerable rise of e-commerce has been taking place in Macedonia in the last decade but remains quite new for the most. The successful adoption of the ICT in e-commerce is dependent on the way it has been used by the implementers, not always being as successful in Macedonia. The understanding and implementation of e-commerce is becoming a very important component of development in Macedonia. The number of internet users' is increasing notably, and this affects the increase of e-commerce. Shopping via credit card has become a reality a decade before, and the companies are increasingly applying it. The number of companies that sell online is still not satisfied. The existing online companies are selling computer equipment, textile products, books and other products and services. The companies put the blame on the banks by saying that they are the biggest reason for the lack of development of e-commerce, because of huge commissions. Confidence among customers must be built by adopting appropriate regulations for e-commerce that will precisely define the conditions for security and confidence in the performance of existing electronic services (USAID, 2010).

3.1 The Current Situation in the Field of e-Commerce in Macedonia

Macedonia relatively late has brought the law for e-commerce ("Official Gazette of the Republic of Macedonia" 133/07). This emphasizes the fact that Macedonia needs to undertake obligation for harmonizing national legislation with EU regulations on e-commerce. One of the reasons for the delay lies in the law on

e-commerce in Macedonia. This law must include clear defined conditions of privacy and security when performing e-commerce services and protection elements for customers who use such services. So far, the Ministry of Economy, as a body which is working on preparing the law and supervising its implementation, is concentrated on analyzing the application of this law and taking proper measures that will positively influence on all involved parties, the companies, the banks and the customers (ExportGov, 2017)(Hadzieva, 2008). According to the data of the State Statistical Office of the Republic of Macedonia, as of January 2016, 93.8% of the companies with 10 or more employees had broadband connection to the Internet (via fixed or mobile broadband connection). Access to the Internet via a portable device using mobile telephone network (3G/4G) was used by 69.3% of the companies. Around 53% of the companies had official website, of which 83% provided on their website descriptions of goods or services, price lists, 59% had links or references to their social media profiles, and 20% provided online ordering or booking. Cloud computing services are used by 6.9% of the companies. In 2015, 23.5% of companies sent or received electronic invoices in a structure suitable for their automated processing. During 2015, 11.4% of companies with 10 or more employees had e-commerce, i.e. buying or selling goods or services over Internet, 7.4% of enterprises had e-sales, and 5.6% of enterprises had e-purchases (State Statistical Office, 2016). These numbers have the potential to reshape Macedonian economy by the next decade.

3.2 Conditions for opening an Online Store in Macedonia

To conduct e-commerce, internet merchant (legal entity) must first be registered as a company in the Central Registry. Those companies that are already registered and operate as offline retailers, but want to expand online, they do not need re-registration. Unlike physical stores, for online stores is required fulfillment of certain minimum technical conditions. In physical stores, in accordance with the regulations, is required a certain number of minimum employees who depend on working hours. Online stores can work with just one employee and therefore can satisfy legal requirements, no matter what the hours are 24/7. E-companies do not need to possess any cash registers, if payments are made through the Internet. Based on Eurostat statistics, the proportion of enterprises having access to the Internet, 2010 and 2016 in % are shown on Table 1 (Eurostat, 2017). By this the biggest precondition is not an obstacle for companies.

Table 1. Proportion of enterprises having access to the Internet, 2010 and 2016 in % (Statistics: (Eurostat, 2017))

Location/Area	2010	2011	2012	2013	2014	2015	2016
European Union (28)	94	95	95	96	97	97	97
Republic of Macedonia	84	82	88	92	93	93	94
Serbia	/	/	/	/	100	/	/
Turkey	91	/	/	/	/	92	94

3.2.1 Privacy Policy, Shipping Policy and Refund Policy

Adopting a privacy, delivery and refund policy is key aspect of e-commerce but not clearly reflected on Macedonian law for e-commerce. Such request comes directly from banks and indirectly from the international card organizations. In fact, these are preconditions to open an online account. These policies are a guarantee for the safety of customers.

3.2.2 Online Payment

The most common method of payment in e-commerce, worldwide and Macedonia, it is with payment cards. Increasingly, in the world frames, are used electronic money, but in Macedonia this is not practiced. Some companies that operate online allow the payments to be done in cash upon the delivery of products. Payment processor as PayPal is currently available but not for money withdrawal (PayPal, 2017).

3.2.3 Payment Cards

In the last few years in Macedonia drastically increased the number of issued payment cards. This can be spotted from Table 2, on statistics performed by the National Bank of Republic Macedonia - NBRM (Updated on: 20.10.2017) for the year 2017. The statistics includes all cards issued in Macedonia by domestic payment services providers. Cards are classified by ownership of the card (natural persons or legal entities), by brand (MasterCard, VISA, American Express, Diners and other domestic cards), by function (cash withdrawal/deposit cards, cards with a payment function - credit and debit and cards with an e-money function), by manner of execution of payments (contact, contact-contactless and internet cards) and by

payment terminal (mobile phones). For the statistic only, valid cards are counted, and the cards issued by merchants are not included, except in instances when they are issued in collaboration with the payment services provider, such is the case with the co-branded cards. Gift cards, as well as cards with other functions are included as "cards with other functions" (NBRM, 2017).

Table 2. The number of issued payment cards in Macedonia for 2017 (credits: (NBRM, 2017))

Payment card statistics Months	Total number of cards (irrespective of the number of functions on the card)			Cards with cash function					Brands	
	Last updated: 20.10.2017	Total	NP*	LE**	Total	NP	LE	Master	Visa	Other***
Jan	1,868,414	1,838,706	29,708	1,850,439	1,820,703	29,736	663,300	1,107,866	79,273	
Feb	1,877,606	1,847,514	30,092	1,858,651	1,828,591	30,060	667,246	1,111,938	79,467	
Mar	1,878,991	1,849,267	29,724	1,859,951	1,830,263	29,688	666,978	1,114,383	78,590	
Apr	1,887,946	1,858,051	29,895	1,868,838	1,838,979	29,859	672,254	1,117,727	78,857	
May	1,854,076	1,824,090	29,986	1,834,951	1,805,003	29,948	660,482	1,096,409	78,060	
Jun	1,848,880	1,818,322	30,558	1,829,682	1,799,164	30,518	656,871	1,094,587	78,224	
Jul	1,851,636	1,820,703	30,933	1,832,401	1,801,510	30,891	660,251	1,094,347	77,803	
Aug	1,853,421	1,822,360	31,061	1,834,127	1,803,111	31,016	663,329	1,093,388	77,410	
Sep	1,862,240	1,830,948	31,292	1,842,892	1,811,648	31,244	667,907	1,096,707	78,278	

* Natural Person ** Legal Entities *** Other brands: American Express, Diners and Others

4. MACEDONIAN PAYMENT INFRASTRUCTURE

Payment system is established infrastructure for transfer of funds between participants, composed of institutions, instruments, rules, procedures and technical means for secure transmission. Nowadays, payments are usually committed using bank deposits. To make payment, the payer must issue instructions, usually to the bank. The instruction can be in a paper form or an electronic instruction using modern technological solutions (credit card, PC or mobile device). Furthermore, the payment instruction is processed and is usually settled without the payer's involvement. Hence, although the payment system is essential for providing payment services to the end users, they are not direct participants and its payment is executed indirectly through a limited number of direct participants in the payment system. On Figure 4, is shown the actual working process of the payment system in Macedonia (NBRM MPI, 2017).

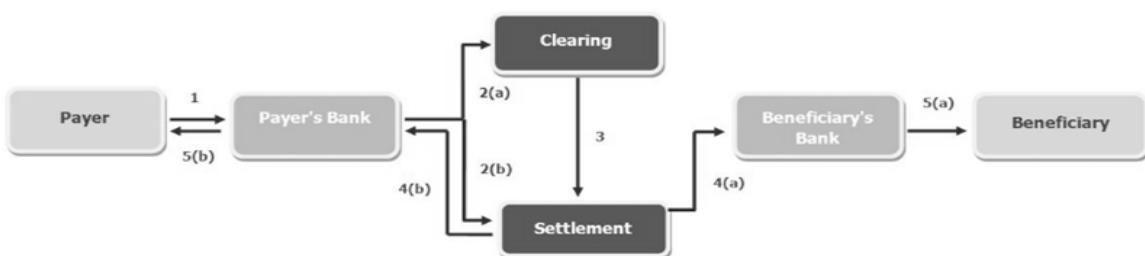


Figure 4. Macedonian payment system (credits: (NBRM MPI, 2017))

4.1 Macedonian Interbank Payment System – MIPS

MIPS is a system for real time settlement of large-value and urgent payments in denars on gross basis (RTGS - Real Time Gross Settlement System). MIPS began operating in July 2001. The NBRM is the owner and operator of the system. Payments settled in MIPS are final and irrevocable, and can be divided into three categories: *Customer-related payments* where users (payer and/or receiver) are clients of the bank and NBRM, which are participants in MIPS; *Bank to bank payments* (both parties are banks and MIPS

participants), and *Service payments* (where a third party, e.g. clearing house, card processor, is allowed to debit the account of an MIPS participant and to credit the amount to another MIPS participant).

The legal basis for the functioning of the MIPS is defined by the Law on Payment Operations. The rights and obligations of NBRM and system participants are regulated by the Operating Rules of MIPS and relevant documentation derived from these rules. The MIPS ensures settlement of other payment systems, implementation of monetary policy, and maintenance of soundness of financial institutions, markets and economic development. Therefore, MIPS is a systemically important payment system and should be harmonized with internationally accepted standards for the payment system operations (NBRM MIPS, 2017).

4.2 Clearing Interbank Systems – KIBS

The Clearing House - Clearing Interbank Systems AD Skopje (KIBS) operates the system for multilateral deferred net settlement for processing retail payments. Participants in KIBS' are banks, Treasury System and the Health Insurance Fund. The settlement of net balances is carried out once a day, between 14:30 and 15:30 in MIPS. The operation of KIBS is regulated by the Law of NBRM, the Law on Payment Operations and regulations prescribed by NBRM. The NBRM is authorized to oversee the safety, soundness and efficiency of payment systems for clearing and settlement (KIBS, 2017)

4.3 International Card System AD – CaSys

International Card System AD Skopje (CaSys) is a system of multilateral deferred net settlement of payments with domestic brands of payment cards. CaSys manages the network of ATMs and banks' POS terminals and on their behalf executes authorization/diversion and processing of transactions carried out on these terminals. Settlement of net balances of participants is performed once per day, between 09:00 and 11:00 in MIPS. The legal basis of the operation of CaSys is the Law on Payment Operations under the supervision of NBRM (CaSys, 2016). Domestic banks that use CaSys system are Uni Banka, Stopanska Banka AD, Sparkasse Banka Macedonia, TTK Banka AD Skopje, Capital Banka, Eurostandard Banka, Centralna Kooperativna Banka, Komercijalna Banka, Stopanska Banka AD Bitola, ProCredit Bank Macedonia , FiBank and others.

4.4 Transaction Fees in e-Commerce

Generally, the e-commerce processing fees in Macedonia are between 3 to 5% of transaction value. One part of it is for international payment cards organization, another for the payment system and the last part for the bank as issuer of the card. The amount is under agreement by all parties and it is based on policies and tariffs. In Macedonia as in any developing country, there is neither a government policy on these fees. These fees increase the cost of e-commerce, at least for the Internet companies. Actually, the parties are working on this issues and a kind of stimulus measure are starting to be visible (MIO, 2007) (ETrgovija, 2014).

4.5 Cash, issuing Invoices and recording Sales by Internet Merchant

The essence of e-commerce is a non-cash payment via Internet. The Internet allows merchant and customer to decide for the payment to be made in cash upon delivery of the products, as is the case with some of our online merchants. Current legislation does not stipulate electronic form of the invoice, the only one that is accepted for tax purposes is the printed form. The law on e-invoice is delayed for unknown purposes.

4.6 Delivery of Products

Speed, reliability and cost of delivery may be a key segment for the development of e-commerce in Macedonia. Online retailers often use the services of specialized courier companies like FedEx and DHL. Online retailers that have their own network of offline stores use the same transport and distribution. This is an advantage in terms of speed and reliability of the delivery and in terms of price. The deadline for delivery outside Skopje (capital city) is 24 to 48 hours, while in the territory of Skopje is shorter. The advantage of delivery service within "Macedonian Post Offices" in relation to other companies is having the most developed network (branches and vehicles) in the country. On the other hand, they are less flexible in

negotiating with online retailers who are treated like any other customer. Smaller logistic companies are ready to deliver the products for low price and other favorable conditions. Usually the cost of delivery ranges from 80 to 150 denars (1.3 to 2.5 euros) for delivery within territory of Skopje, while delivery to other cities ranges from 120 to 250 denars (2 to 4 euros) (Trajanov, 2007).

4.7 Legal

One key reason for the slow development of e-commerce in Macedonia is that there is not an overall policy framework covering all aspects in enabling e-commerce environment. Many researches point on the following aspects: *Reliable encryption and decryption techniques* are necessary for processing credit card information, such as authentication, authorization, confidentiality and integrity to services, increasing the security of e-commerce transactions; *Relevant digital signatures and electronic contracts* in all e-commerce transactions; *Trusted certification authorities* for secure electronic transactions; *Clear customer protection mechanisms*; *Efficient electronic payments* and certainty of the bank regulations; *Certainty in legislation on copyright and intellectual property rights*.

These aspects could stimulate the development of e-commerce. Not all of them are present in Macedonian Commercial Law that is undergoing a harmonization process with European Union (EU) legislation (Law on e-commerce, 2007) (Ma'aruf L, 2012) (Actuary, 2017)(El Gawady, 2005).

5. RECOMMENDATIONS FOR IMPROVING THE SITUATION OF E-COMMERCE IN R. MACEDONIA

The most essential and widely recognized factors for the development of e-commerce can be considered:

- **Trust** that motivates customers on future purchases and in parallel it reduces customer worries about privacy (Pittayachawan S, 2008) (Baubonienė, 2015).
- **Security**, one of widely recognized obstacle to the expansion of e-commerce. Factors like security and privacy, innovativeness, knowledge, and awareness level increase the acceptance of e-commerce services among Macedonian customers (Shahibi Mohd Sazili, 2011).
- **Knowledge and Recognized Value**, are significant on attitude towards the use of e-commerce. The use of online banking services for instance is new knowledge to many customers, and the lack of awareness of online banking is a crucial factor in preventing customers from adopting it (Mohammed J.A, 2013).
- **IT Competence**, refers to skill and knowledge of company's staff who know the function of e-commerce, ICT and computer software in their daily work to support business objectives. ICT literate employees with are more likely to adopt e-commerce (Jantarajaturapath P, 2011).
- **Accessibility** is a critical factor for e-commerce success. Easy web accessibility can attract people to use, understand and interact with the web (OBP, 2016)(Sambhanthan A, 2013).
- **Perceived Quality** of a service is technological dimension in terms of what is delivered, and as functional dimension in terms of how the service is delivered or the customer perception of the quality of information about the product/service that is provided by a certain website (Sebastianelli R, 2007).
- **Role of Government** and its vision supports the essential requirements for the development of e-commerce through promotion and awareness, HR and skills development, liberalization of telecom, regional strategy and financial services (AlGhamdi R, 2013) (Wikibooks, 2017).

6. CONCLUSION

The future of e-commerce looks promising for small sized companies in Macedonia. Having customers that are easy adopting new technologies, can be a challenge for companies considering adapting to this change in customer behavior or even predicting it. Small sized companies are job creators. For instance, AliBaba Group in China and their e-commerce platforms (the B2C platform), are major job creators. Their prediction is that by 2035 they will create 30% of the jobs in China's Digital Economy (Tom Brennan, 2017). This is a motivational example that Macedonia needs to investigate. It clearly shows how can help even small shop

owners or individual sellers to reach greater market and increase the revenues. By analyzing the world trends in e-commerce practice, the companies that are already online or are planning to be, can focus their efforts to survive the next advance in technology. The pace of technology change is very fast, and all parties involved in e-commerce need to be synchronized with it. Investment in mobile commerce, availability of interactive content (Puri, 2015), Social selling advertising or social media commerce (Lordahl, 2017), digital couponing (Couponing in the digital age!, 2011), Artificial Intelligence (AI) Systems (Goel, 2017), even delivery drones (Apparatus, 2016) are the trends for companies all over the world and inclusively in Macedonia they need to focus on. Factors such as Trust, Security, Knowledge and Recognized Value, IT Competence, Accessibility, Perceived Quality and Role of Government are the most important ones that need to be considered to support the advancement of e-commerce implementation. Macedonia needs to encourage, support and improve the e-commerce developments especially for B2C. Macedonian citizens wants to fit in the move for e-commerce because it has the potential, infrastructure and at the same time cannot afford to be left out. The future of e-commerce for small businesses will be smarter, faster and more promising. Maybe for some of them Macedonian online market will have to wait longer.

REFERENCES

- Actuary. (2017). *The E-Commerce Directive and Regulations*. Retrieved from TheActuary.com: <http://www.theactuary.com/archive/old-articles/part-1/the-e-commerce-directive-and-regulations/>
- AlGhamdi R, N. A. (2013). A study of Influential Factors in the Adoption and Diffusion of B2C E-Commerce. *International Journal of Advanced Computer Sciences and Applications (IJACSA)*.
- Apparatus. (2016). *Delivery Robots The Future of Ecommerce*. Retrieved from Apparatus.io: <http://www.apparatus.io/Delivery-Robots-The-Future-of-Ecommerce>
- Baubonienė, Ž. (2015). E-COMMERCE FACTORS INFLUENCING CONSUMERS' ONLINE SHOPPING DECISION. *ISSN 2029-7564 (online)*, 74-81.
- BBC. (2014). *Effects of e-commerce on businesses*. Retrieved from BBC Technology News: <http://www.bbc.co.uk/schools/gcsebitessize/ict/implications/2workpatternsrev4.shtml>
- CaSys. (2016). Retrieved from NBRM: <http://www.casys.com.mk/mk.html>
- Chris G. Christopher, J. (2011). *The economic impact of e-commerce*. Retrieved from supplychain quarterly: <http://www.supplychainquarterly.com/columns/scq201102monetarymatters/>
- Couponing in the digital age!* (2011, October 26). Retrieved from <http://thekrazycouponlady.com/2010/08/16/couponing-in-the-digital-age>
- Crimsonwing. (2014). *Introduction to B2C E-Commerce*. Retrieved from crimsonwing.com: <https://www.crimsonwing.com/wp-content/uploads/2014/11/B2C-eCommerce-brochure.pdf>
- El Gawady, Z. M. (2005). The Impact of E-commerce on Developed and Developing Countries Case Study: Egypt and United States. *International Conference of Globalization,Technology and Sustainable Development* (pp. 21-32). UAE: World Association for sustainable Development.
- ETrgovija. (2014, January 21). *Провизија во Интернет Трговија*. Retrieved from ETrgovija: <https://etrgovija.wordpress.com/2014/01/21/провизија-во-интернет-трговијата/>
- Eurostat. (2017). *Internet access Stats*. Retrieved from EUROSTAT: http://appsso.eurostat.ec.europa.eu/nui/show.do?wai=true&dataset=prc_hicp_manr
- ExportGov. (2017, September 22). *Macedonia – eCommerce*. Retrieved from export.gov: <https://www.export.gov/article?id=Macedonia-ECommerce>
- Gao, J. (1999). *Introduction to E-Commerce*. Retrieved from San Jose State University;Computer,Information&Systems Engineering: <http://www.engr.sjsu.edu/gaojerry/course/cmpe296u/296z/introduction.pdf>
- Goel, S. (2017, March 31). *The role of artificial intelligence in the E-commerce sector and its scope in future*. Retrieved from Deccan Chronicle: <http://www.deccanchronicle.com/technology/in-other-news/310317/the-role-of-artificial-intelligence-in-the-e-commerce-sector-and-its-scope-in-future.html>
- Hadzieva, A. (2008, Spring). *The Macedonian law on electronic commerce*. Retrieved from Polenak.com: http://polenak.com/wp-content/uploads/2015/05/AmCham_EM_Spring-2008.pdf
- Ibls. (2017). *Internet Law - Macedonian E-Commerce Law*. Retrieved from Internet Business Law Service: https://www.ibls.com/internet_law_news_portal_view.aspx?s=articles&id=E8D3E770-4AEC-49E0-BBFB-4D412AEF0E7E

- Jantarajaturapath P, K. S. (2011). Factors affecting e-commerce adoption and business success of Thai SMEs. *European Journal of Management ISSN:1555-4015*.
- KIBS. (2017). *NBRM*. Retrieved from <http://www.nbrm.mk/kibs1e32d708-en.nspx>
- Lacka, E. C. (2014). *E-commerce Platform Acceptance*. Springer.
- Law on e-commerce. (2007, October 26). Retrieved from mio.gov.mk: http://www.mio.gov.mk/files/pdf/en/Law_on_%20e-Commerce_2007_REV_final.pdf
- Lordahl, E. (2017, March 28). *Is social selling the future of ecommerce?* Retrieved from conversocial.com: <http://www.conversocial.com/blog/is-social-selling-the-future-of-ecommerce>
- Ma'aruf L, A. K. (2012). An overview of e-commerce implementation in developed and developing country; A case study of United State and Nigeria. *International Journal of Modern Engineering Research (IJMER)*, 3068-3080.
- Machado, A. (2011, Feb). *Usability:Impact on e-commerce*. Retrieved from University of Lisbon Repository: https://www.repository.utl.pt/bitstream/10400.5/2959/1/TESE_Versao_Definitiva_2011.pdf
- METU. (2017). *Introduction to E-commerce*. Retrieved from METU, Middle East Technical University: http://ocw.metu.edu.tr/pluginfile.php/352/mod_resource/content/0/Lecture_3.pdf
- MIO. (2007, October 26). *Degree for promulgation of the law on electronic commerce*. Retrieved from mio.gov.mk: http://www.mio.gov.mk/files/pdf/en/Law_on_%20e-Commerce_2007_REV_final.pdf
- Mohammed J.A, A. M. (2013). The Factors That Affects E-Commerce Adoption in Small and Medium Enterprise: A Review. *Australian Journal of Basic Applied Sciences*, 406-412.
- NBRM. (2017). *Payment cards issued domestically*. Retrieved from http://www.nbrm.mk/platiezhna_statistika-en.nspx
- NBRM MIPS. (2017). *MIPS*. Retrieved from <http://www.nbrm.mk/mipsdd77afa6-en.nspx>
- NBRM MPI. (2017). *Payment Infrastructure*. Retrieved from http://www.nbrm.mk/makedonska_platiezhna_infrastruktura-en.nspx
- OBP. (2016). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. Retrieved from ISO,Online Browsing Platform: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:dis:ed-2:v2:en>
- PayPal. (2017). *PayPal Offerings Worldwide*. Retrieved from https://www.paypal.com/cgi-bin/webscr?cmd=_display-country-functionality-outside&dispatch=5885d80a13c0db1f8e263663d3faee8dc3f308defb7330dd8d0b0a9f21afdf7d3
- Pittayachawan S, S. M. (2008). A multitheoretical approach for solving trust problems in B2C e-commerce. *International Journal of Networking and Virtual Organizations*, 369-395.
- Puri, R. (2015, November 9). *The Future of Ecommerce: 7 Trends That Are Changing How We Shop*. Retrieved from <https://www.shopify.com/enterprise/70543555-the-future-of-ecommerce-7-trends-that-are-changing-how-we-shop>
- Sambhanthan A, G. (2013). Implications for Improving Accessibility to E-Commerce Websites in Developing Countries- A study of Hotel Websites.
- Sebastianelli R, T. N. (2007). How Shopping Frequency and Product Type Affect Consumers' Perceptions of E-Tailing Quality. *Journal of Business&Economics Research*.
- Shahibi Mohd Sazili, F. S. (2011). Security Factor and Trust in E-Commerce Transactions. *Australian Journal of Basic and Applied Sciences ISSN 1991-8178*, 2028-2033.
- State Statistical Office. (2016, October 20). *Користење на информатичко - комуникациски технологии во деловните субјекти*. Retrieved from stat.gov.mk: <http://www.stat.gov.mk/pdf/2016/8.1.16.29.pdf>
- Sundheim, K. (2017). *Which Is The Best Business Model For Your Startup – B2B Or B2C?* Retrieved from Entrepreneurs-Journey.com: <https://www.entrepreneurs-journey.com/8101/what-is-b2c/>
- T.Watson, R. (n.d.). *Electronic Commerce:The Strategic Perspective*. Retrieved from BC Open Textbooks: <https://opentextbc.ca/electroniccommerce/chapter/introduction/>
- Tom Brennan. (2017, March 27). *ALIBABA TO GENERATE 30% OF JOBS IN CHINA'S DIGITAL ECONOMY: BCG*. Retrieved from Alizila: <http://www.alizila.com/alibaba-china-digital-economy-bcg/>
- Trajanov, D. (2007). *E-commerce in Macedonia*. Retrieved from academia.edu: http://www.academia.edu/3710713/E-commerce_in_Macedonia
- UNCTAD. (2015). Information economy report. *United Nations Conference on trade and development*. Geneva: United Nations UNCTAD Division.
- USAID. (2010). *Analysis on the e-commerce situation in the Republic of Macedonia*. Skopje: Ministry of Information Society and Administration.
- Watson R, B. P. (2008). *Electronic Commerce: The Strategic Perspective*. The Global Text Project.
- Wikibooks. (2017). *E-Commerce and E-Business/E-Commerce in Developing Countries*. Retrieved from Wikibooks.org: https://en.wikibooks.org/wiki/E-Commerce_and_E-Business/E-Commerce_in_Developing_Countries

ENABLING SEMANTIC INTEROPERABILITY OF DISEASE SURVEILLANCE DATA IN HETEROGENEOUS HEALTH INFORMATION SYSTEMS IN NAMIBIA AS A STANDARD APPROACH

Nikodemus Angula and Nomusa Dlodlo
Namibia University of Science and Technology, Namibia

ABSTRACT

In the Namibian health sector silo information systems are run by the Ministry of Health and Social Services despite the 14 regions across the country. The Silo information systems came into existence through non-governmental organisations donation. Furthermore, District health information system (DHIS-2) for every region there is a main DHIS-2 at Khomas regional office under the MoHSS. As it stands, health information systems (HIS) such as the DHIS-2 and silo systems work separately from one another, therefore this study is centered on finding a framework to enable semantic interoperability of data in these heterogeneous health information systems (HIS). As a result this would enable the DHIS-2 and these silo systems in the Namibian public hospitals to act as an integrated platform that shares and exchanges disease-surveillance information. With the use of integrating the Healthcare Enterprise (IHE) standard, this research guides on how to integrate these heterogeneous health information systems through the adoption of established health standards. In addition, IHE and Health Level Seven (HL7) are implemented to interface DHIS and silo systems at a data level. Moreover, an Enterprise Master Index that uses the Patient Identification Segment is adopted. The purpose of this study was to design and develop a framework for data semantic interoperability of DHIS and these other health information silo systems so that they can exchange disease-surveillance information. The outcomes of this study are a framework to enable the semantic interoperability of disease surveillance data in Namibian hospitals through the adoption of IHE and HL7 standards in the form of an artifact that demonstrates how disease surveillance data can be integrated in the Namibian health domain.

KEYWORDS

Integrating the Healthcare Enterprise, IHE, Health level Seven, HL7, Semantic Interoperability, Health Information Systems, Disease Surveillance, Enterprise Master Patient Index

1. INTRODUCTION

In the Namibian health sector silo information systems are run by the Ministry of Health and Social Services despite the 14 regions across the country (Angula & Dlodlo, 2018). The Silo information systems came into existence through non-governmental organisations donation. In addition, is important that semantic interoperability of data is adopted in this heterogeneous environment so that the DHIS and these silo systems in the Namibian public hospitals can act as an integrated platform to share and exchange health-related information. The study employed disease surveillance as a case study, this research sought to interface DHIS-2 and silo systems at a data level. The Integrating the Health Enterprise (IHE) is regarded as a consortium that advocates the sharing of health information from distributed and heterogeneous health information systems (Siegel, 2001). This is done standards that specify protocols through which integrated systems can communicate as profiles. In this study, the disease surveillance information is collected from patient data that is captured when patients visit health facilities. The IHE, through its information Technology Infrastructure Framework Volume 2b (Integrating Healthcare Enterprise, 2018), provides a Patient Identification Segment which provides a way in which patient demographic details can be exchanged among HIS. The Namibian health domain has not capitalized on these standards yet. In this research, a framework that uses IHE format is proposed for semantic interoperability of data in the Namibian health domain, and specifically the Health Level Seven (HL7) standards for meaning. The Enterprise Master Index that uses the

Patient Identification Segment is adopted. The outcome is an artifact that illustrates how disease surveillance data can be integrated in the Namibia health domain. Furthermore, District health information system (DHIS-2) for every region there is a main DHIS-2 at Khomas regional office under the MoHSS. As it stands, health information systems (HIS) such as the DHIS-2 and silo systems work separately from one another, therefore this study is centered on finding a framework to enable semantic interoperability of data in these heterogeneous health information systems (HIS). As a result this would enable the DHIS-2 and these silo systems in the Namibian public hospitals to act as an integrated platform that shares and exchanges disease-surveillance information. With the use of integrating the Healthcare Enterprise (IHE) standard, this research guides on how to integrate these heterogeneous health information systems through the adoption of established health standards. In addition, IHE and Health Level Seven (HL7) are implemented to interface DHIS and silo systems at a data level. Moreover, an Enterprise Master Index that uses the Patient Identification Segment is adopted. The purpose of this study was to design and develop a framework for data semantic interoperability of DHIS and these other health information silo systems so that they can exchange disease-surveillance information. The outcomes of this study are a framework to enable the semantic interoperability of disease surveillance data in Namibian hospitals through the adoption of IHE and HL7 standards in the form of an artifact that demonstrates how disease surveillance data can be integrated in the Namibian health domain.

2. LITERATURE REVIEW

2.1 Health Information Systems in Namibia

Namibia's National Health Information System (NHIS) falls under the Primary Care Directorate in the Ministry of Health and Social Services (MoHSS, 2014). It is charged with the responsibility of providing a comprehensive source of data on a large number of health-related indicators. The NHIS was designed to improve service delivery in terms of quality and effectiveness of strategies and to monitor the trends in disease occurrence. In addition, it also provides information for national policy makers, socio-economic and health personnel, as well as the public at large (HaosesGorases, 2015). The organisational structure of the NHIS system is fragmented across different directorates and institutions. The challenges to the system are that there is a shortage of human resources to coordinate, analyse, and report on the information in a comprehensive and timely fashion. In the Namibian health environment, a large number of systems, databases, and processes are fully manual, paper-based, or only partially electronic, and to a large extent formats are either fragmented or nonstandard. This adds significantly to work burdens and seriously undermines efficiency. The Namibian HIS as a routine health and management information system was first introduced in Namibia in 1992 (Chotard, 1992), (Chotard, Conkle & Mason, 1992).

2.2 Data Interoperability

Barbarito et al. (2012), stated that Information technologies (ITs) have now entered the everyday workflow in a variety of healthcare providers with a certain degree of independence. This independence may be the cause of difficulty in interoperability between information systems and it can be overcome through the implementation and adoption of standards. Here we present the case of the Lombardy Region, in Italy, that has been able, in the last 10 years, to set up the Regional Social and Healthcare Information System, connecting all the healthcare providers within the region, and providing full access to clinical and health-related documents independently from the healthcare organization that generated the document itself. This goal, in a region with almost 10 million citizens, was achieved through a twofold approach: first, the political and operative push towards the adoption of the Health Level 7 (HL7) standard within single hospitals and, second, providing a technological infrastructure for data sharing based on interoperability specifications recognized at the regional level for messages transmitted from healthcare providers to the central domain.

2.3 Heterogeneous Distributed Systems

In the Namibian health sector heterogeneous distributed systems come into play through donation from non-profit organisations. Different donors have donated different silo systems running on different software. However, heterogeneous distributed systems have become highly heterogeneous (Godfrey, 2009). Rather than running on clusters or supercomputers composed of identical nodes, today's distributed systems have wide variation in participants' failure rates, bandwidth, processing speed, security, and other dimensions (Godfrey, 2009). Lima & Faria (2016), stated that heterogeneous and distributed system of systems are involved in the provisioning of end-to-end services to user (Lima, 2016). Multibase is a software system for integrating access to pre-existing heterogeneous, distributed databases.

3. RESEARCH METHODOLOGY

The aim of this research was to design and develop a framework for disease surveillance data semantic interoperability of DHIS and other health information silo systems so that they can exchange health data and information.

The main research question was "How can semantic interoperability of data in heterogeneous health information systems in Namibian hospitals be achieved"?

The sub-questions are:

- i. How is healthcare information currently shared between existing DHIS and silo systems in Namibia?
- ii. How can a data interlink protocol that would govern heterogeneous health information system in Namibian public hospitals be developed?
- iii. How can information relating to healthcare be interfaced between existing DHIS and silo systems in Namibia by adopting technologies in the distribution of health information?
- iv. How can a model to enable semantic interoperability of data in heterogeneous health information systems in Namibian hospitals be designed and developed?

The research is in two phases. The first phase was the collection of data on the status of integration of HIS in the Namibian healthcare sector. The study used a qualitative approach. The second phase was the design and development of the semantic interoperability framework. The study adopted a case study setting looking at 8 hospitals in Namibia. The case studies selected in this case was the Khomas regional office (KRO) and public hospitals. Therefore, the study used the grounded theory to guide in data collection and Design Science Research Theory was used to develop the artefact. Expert reviews were used to validate the artefact developed.

4. DATA ANALYSIS

The data was analyzed to get a better understanding on how disease surveillance information is exchanged and communicated between public hospitals and the Ministry of health and Social Services in Namibia. The data gathered were analyzed according to the two Technician interviewed using semi-structured interview.

According to the two IT technician at the local hospital in Windhoek there is no management system for advanced staff. In addition, they further stated that they do not have a system or framework that aggregates data from remote health systems for management staff to view and analyse information and also they do not have a system that exchanges data automatically. The system or framework developed in this study would aggregates data from remote health systems for management staff to view and analyse information for hospitals decision making. The two technicians further, stated that they do not have a system or framework for semantic interoperability of data in Namibian public hospitals. Currently, what system administrators do is login remotely and access a file in another computer on the same network which does not allow silo systems data to be accessed, communicate, and exchange health data and information from other silo systems hosted in other Namibian public hospitals.

Table 1. Technicians interviews

Technicians Interviews				
Type of job	Number of Technicians	Silo systems	No exching of data in the current system	Technician move from one hospital if they need data
Technician	1	stand alone systems	No exching of data in the current system	No exchanging of data
Technician	1	Silo systems	Exchanging of health data	The current process of exchanging data in hospital

The table 1 below symbolises the responses from the two technicians interviewed in the local hospital in Windhoek who clearly stated that at the present moment hospitals in Namibia have standalone silo systems that works in isolation and they do not communicate and exchange health related information to one another. The two technicians highlighted that this is a main challenge in the Namibian health sector simply because if a technician requires data from a remote silo systems in another hospital they are required to go physically to that specific hospitals in order to acquire health related data. As a result the study developed a protocol for the exchange of data from remote systems and aggregation of that data into meaningful information.

5. FINDINGS

5.1 Architecture of the Integrated System

Figure 1 shows an example of four HIS of DHIS-2, ETR.net, EDT and EPI Info which are hosted of different public hospitals in Namibia. When a patient visits a hospital, they are allocated a hospital identifier for the day. This identifier is unique to the hospital on the particular day. For example, it may indicate the patient number in the queue in that hospital for that particular day. The local identifier, plus the demographic details of the patient are transferred to the common registry for all hospitals for documenting the history of the treatment of the patient. If the patient visits a different hospital, the same procedure takes place. The additional event is that the demographic data from the second hospital is compared against the demographic data already in the registry to extract the medical history of the patient.

An example of how the processes work can be explained using the following example. If a patient visits Windhoek Central Hospital, for example, the DHIS-2 sends a query to the Health Information Exchange (HIE). HIE is middleware that facilitates exchange of information. The Health Information Exchange (HIE) has a registry (the equivalent of EMPI) and a repository. In the registry, the patient visit to Windhoek Central is allocated a unique local Windhoek Central hospital-related number to the patient. The unique Windhoek Central hospital-related unique number and the demographic details of the patient go into the registry. When the ETR.net that is housed in Katutura hospital sends a query to the registry about a patient whose details have just been captured as a result of their visit to Katutura hospital, demographic details from this hospital are compared with what is already held in the registry in order to extract the history of the patient from the registry. This means that mobile patients can visit any hospital and there will be continuity in their treatment by virtue of the fact that their history can easily be extracted from the system. The repository on the other hand keeps documents such as MRI scan, CDA documents on each and every patient, which can be referenced.

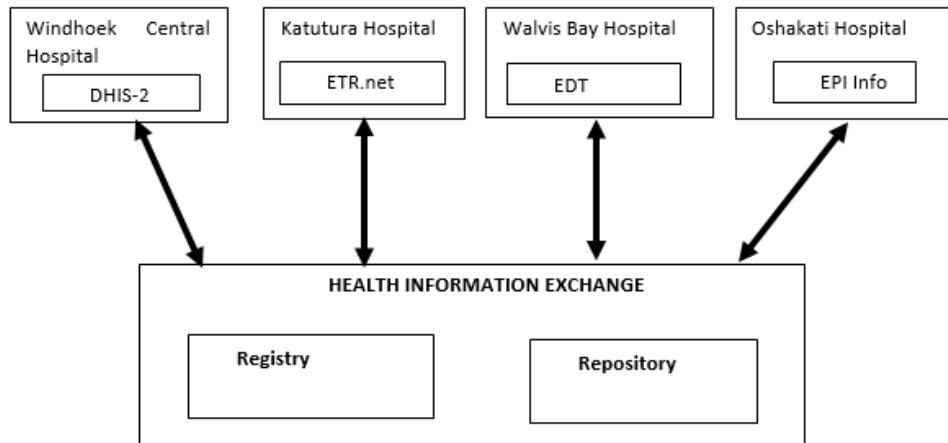


Figure 1. Architecture of integrated system

5.2 Profile

The IHE IT Infrastructure Technical Framework has different profiles (IHE Technical Frameworks, 2018). The profile selected for Namibian disease surveillance data case to comply with IHE standard corresponds to Transaction ITI-21 of the IHE Technical Framework. Transaction ITI-21 and is used by the Patient Demographics Consumer and Patient Demographics Supplier Actors.

This transaction involves a request by the Patient Demographics Consumer Actor for information about patients whose demographic data matches data provided in the query message. The request is received by the Patient Demographics Supplier Actor. The Patient Demographics Supplier Actor subsequently processes the request and returns a response in the form of demographic information for matching patients.

The general use case role of the profile is as follows (see Figure 2):

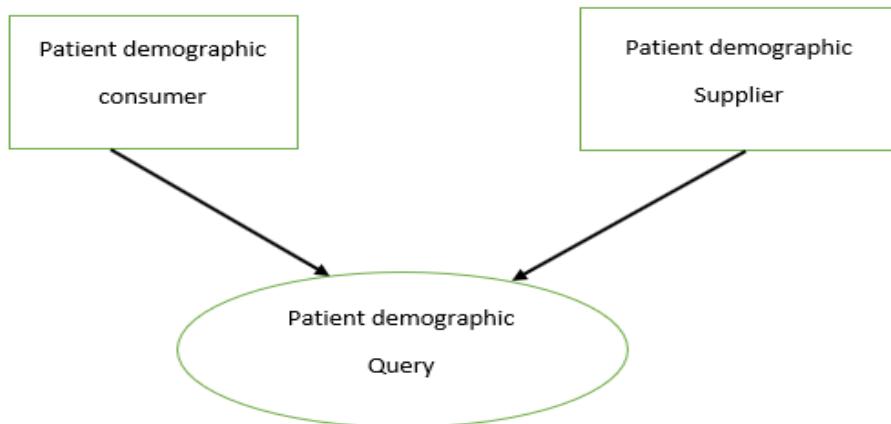


Figure 2. Patient Demographics Actor. Adapted from IHE Technical Frameworks

Therefore the processes are as follows:

- Actor: Patient Demographics Consumer role requests a list of patients matching a minimal set of demographic criteria (e.g., ID or partial name) from the Patient Demographics Supplier. It populates its attributes with demographic information received from the Patient Demographics Supplier. The consumer in this case is the HIS that is housed in the hospital, such as DHIS-2, ETR.net, EDT or EPI Info. The supplier in this case is the Registry.

- Actor: Patient Demographics Supplier Role returns demographic information for all patients matching the demographic criteria provided by the Patient Demographics Consumer.

5.3 HL7 Code Explanation

HL7 is a communications standard protocol. The protocol consists of a header plus the message. The message in this case is the demographic details of the patient which are sent from the querying HIS. The messaging segments that fit the Namibian health sector in HL7 is a stored procedure request which enables an application on one system to execute a stored procedure on another system, which is coded to extract specific data.

The Health Information Exchange Structure's registry should match the HL7, i.e. the HL7 fields to match the registry. The queries from DHIS-2, ETR.net, EDT and EP Info should be standardized to suit the registry. For example, all the HL7 codes of AD-1, AD-2, etc. should be available in the queries. Looking at the following example:

Let us suppose that the DHIS-2 systems captures the following data:

Table 2. DHIS-2 Systems Capture

Name (NN-Windhoek Central)	Surname (SS-Windhoek Central)	Address (AA-Windhoek Central)	National ID (ID-Windhoek Central)
John	Smith	5 Kleine Kuppe	6578823

Let us also suppose the ETR.net system in Katutura Hospital captures the following data

Table 3. ETR.net Systems Capture

Name (NN-Katutura)	Surname (SS-Katutura)	Address (AA-Katutura)	Next-of-kin (NK-Katutura)	Passport number (PN-Katutura)
Andrew	Gonzalez	123/56 Katutura	Bedford	DF4279QD

The registry should have all the fields that are available in the HL7 code, in the standard code syntax.

Table 4. Registry Data

Person ID (XCN-1)	Family Name (XCN-2)	Given Name (XCN-3)	Address (AD-1)	Telephone (XTN-1)	Designation (AD-2)	State (AD-4)
6578823	Smith	John	5 Kleine Kuppe	-	-	-
DF4279QD	Gonzalez	Andrew	123/56 Katutura	-	-	-

This is all about matching the HIS in the different hospitals with the registry at the MoHSS. The queries for DHIS-2 (DHIS-2, 2018), ETR.net (ETR.net, 2018), EDT (EDT, 2018) and EPI Info (EPI Info, 2018) are standardized to suit the registry. Match the names NN-Windhoek Central and NN-Katutura with the name in the registry XCN-3. Match the surnames SS-Windhoek central and SS-Katutura with the family name XCN-2 in the registry. In this example, the next-of-kin (NK-Katutura) does not have a representative code in the registry. Therefore when updating the systems, it is only the local ETR.net in Katutura hospital that is updated and the information is discarded without error when it comes to updating the registry. If not given any value, default values will be put in the registry.

The information should be traceable, that is, obtained from the header which has the sending facility, sending software, etc. The sequence in the message header is standard. Whenever a message goes out, it follows the established format of the header in terms of the sequence of the fields and the length of each field. The user need not know the format of the header.

There should also be privileges on who can update the registry. There is also the issue of how much information is availed to the person querying. There should be different access rights for the different medical personnel. For example, when HL7 provides clinical information, suppress certain information if it's a nurse accessing the information as opposed to a doctor accessing.

5.4 Technology Demonstrator



Figure 3. Distribution of disease

The map is for the user to navigate and check from which geographical location the patient is from. It checks the type of disease they are suffering from and relates it to geographic area.

6. CONCLUSION

Data interoperability has been the main challenges in health sector across the globe. One of the most challenging problems in the healthcare domain today is providing interoperability among health care information systems. In order to tackle this problem, semantic interoperability data model will be designed and developed among public hospitals and clinics in Namibia. The use of information and communication technologies (ICT) by healthcare service providers is often driven by the need to achieve effective and efficient services delivery to other public hospitals in Namibia. Public hospitals in Namibia have been challenged in the way they carry out their operations, processes and how health related information is distributed and accessed by other public hospitals in all the 14 regions in Namibia. The challenge include distribution and information flow between public health institutions and other public hospitals operating in the regions. Unfortunately, these challenges and gaps hampers and negatively impact healthcare service delivery between public hospitals operating within the same health sector environment. This study will design and develop a data interface for health information sharing between different public hospitals in Namibia.

ACKNOWLEDGEMENT

The authors are grateful for the support rendered by the Namibia University of Science and Technology through funding and availing facilities to undertake the research.

REFERENCES

- Angula, N., and Dlodlo, N. 2018. Towards a Framework to Enable Semantic Interoperability of Data in Heterogeneous Health Information Systems in Namibian Public Hospitals. In International Conference on Information Theoretic Security (pp. 835-845). Springer, Cham.
 ASTM, <https://www.astm.org/>.

- Barbarito,F.et al, 2012.Implementing standards for the interoperability among healthcare providers in the public regionalized Healthcare Information System of the Lombardy Region retrieved from <https://doi.org/10.1016/j.jbi.2012.01.006>.
- Chotard, S. J. 1992. Nutrition and health information and surveillance systems retrived from <https://www.scribd.com/doc/969821/Nutrition-Information-inNamibia-Situation-Analysis-the-Way-Forward..>
- Godfrey, P. B. 2009. Designing Distributed Systems for Heterogeneity retrieved from <http://digitalassets.lib.berkeley.edu/techreports/ucb/text/EECS-2009-82.pdf>.
- Haoses-Gorases, L. 2015. Utilisation of Health Information System (HIS) in Namibia: Challenges and Opportunities Faced by the Health Care Delivery System/Health Sector. Strategies for permanet access to scientific information in South Africa: focus on environmental information.
- Integrating the Healthcare Enterprise, http://www.ihe.net/uploadedFiles/Documents/ITI/IHE_ITL_TF_Vol2b.pdf
- Lima.J, &. F. 2016. Testing Distributed and Heterogeneous Systems: State of the Practice.In Proceedings of the 11th International Joint Conference on Software Technologies (ICSOFT 2016)ISBN 978-989-758-194-6, pages 69-78. DOI: 10.5220/0005989100690078.

Short Papers

STUDENT RETENTION AS AN INDICATOR OF QUALITY ASSURANCE IN ECUADORIAN UNIVERSITIES

Marco Veloz¹, Mayra Albán¹, Idalia Pacheco¹, Gloria Vizcaíno¹ and Mónica Salazar²

¹*Technical University of Cotopax*

²*University of Armadas Forces
Ecuador*

ABSTRACT

The quality assurance allows development processes to establish a proper management in universities. One of the ways to ensure quality is to establish strategic policies that guarantee the permanence of students, the consequences of desertion in higher education institutions generate economic and social consequences for IES. For this reason, a linear probability model is proposed by applying least squares based on obtained data from results of evaluations carried out by control agencies in Ecuadorian universities. As a result was obtained that the college retention indicator influences in the quality assurance of universities.

KEYWORDS

Student's Retention, Quality Assurance, University Evaluation, Box Jenkins

1. INTRODUCTION

For the Ecuadorian system of higher education quality is a principle which consists in the constant and systematic pursuit of the excellence, relevance, optimal production, knowledge transfer and development of thought through self-criticism, external criticism and continuous improvement " (Ecuador National Assembly, 2010). In this way, means the quality of universities and polytechnics schools as the degree to which, in accordance with its mission, framing in the purpose and functions of higher education system in Ecuador, these achieve the objectives of teaching, research and linking up with society, through the execution of processes that observe the principles of the system and seek continuous improvement.

(Larrea, 2015) One of the problems facing the higher education system has to see with the rate of terminal efficiency which is expressed in students who are titled tardily or desert of their professional training, the same that presents difficulties from the basic levels, the problem lies in high desertion showed in the higher education system, there are many causes to which the problem is attributed, however this document tries to explain the factors that influencing empirically to the rate of retention of higher education Institutions (HEI).

(CEAACES, 2015)The assessment committee, accreditation and quality assurance of higher education (CEAACES) has designed different models to evaluate and accredit institutions of higher education in Ecuador in the last 10 years, the last model used in 2015 presents six criteria: organization, academy, research, link up, resources, infrastructure and students. This research aims to analyze the relationship or impact indicators in academy student's retention rate, the explanatory variables considered are: number of full-time teachers at the institution (FTT); number of teachers with PhD degree (PhD); number of teachers who only have third level degree (FTT_{III}); number of teachers who have fourth level (FTT_{IV}) and average number of students per teacher full time (SFTT). While the dependent or explained variable is the student retention rate (RR) indicator of student's standard, the problem formulation: In which way the academic standard variables influence in the student retention rate?

2. BODY OF PAPER

The structure of the evaluation model issued by the CEAACES is organized around six evaluation criteria: Organization, Academy, Research, Linking up, and students who consider broader aspects of quality, and are related to the basic functions of universities and polytechnics schools as well as processes, the conditions and the resources that allow proper execution of them.

On a second level, the structure of the evaluation model considers sub criterion that address on more specific conceptual aspects and attributes of the criterions; the sub criterions are likely to be measured through evaluation indicators. These indicators are of two types: qualitative and quantitative indicators.

For quantitative indicators, standards are constituted of mathematical functions that determine the performance of the institution on a scale of 0 to 1, where 0 represents a null fulfillment of the standard and 1 corresponds a totally satisfactory standard fulfillment.

(CEAACES, 2015) The indicator tax rate level presents a standard that evaluates the level of permanence and indirectly the level of desertion of the students of the institution at the beginning of their career. The expected retention rate is 80% whereas low value to this are considered useful logistic function as shown in Figure 1.

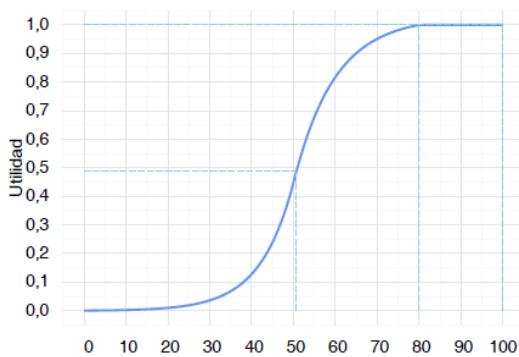


Figure 1. Function Utility retention rates

Source: Evaluation Model universities and polytechnics schools, 2015

The calculation of the TR indicator is an average that considers the number of students enrolled in first level versus the number of enrolled two years after starting. So we have: $RR = \frac{NTEAD}{NEMPO} * 100$

Variables: NEMPO: Number of students enrolled during the regular academic period in which the evaluation is carried out of the institution, which were admitted two years earlier. NTEAD: Total number of students who were admitted in the career two years before the evaluation period.

2.1 Proposed Method

The methodology used in this research is based on econometric elements, on the one hand a regression analysis is applied through ordinary least squares, so too methodology Box Jenkins also is applied since it is about data of time series.

2.1.1 Ordinary Least Squares (OLS)

(Gujarati & Porter, 2010) OLS method is the most common in the regression analysis, especially for being much more intuitive and mathematically simpler than the maximum likelihood method. Furthermore, as discussed after in the context of linear regression, usually the two methods give similar results.

The ordinary least squares method is attributed to Carl Friedrich Gauss, German mathematician. From certain assumptions, the least squares method presents very attractive that have become in one of the most effective and popular regression analysis. To understand this, first we explain the principle of least squares.
$$Y_i = \beta_1 + \beta_2 X_1 + \mu_1$$

(Novales Cinca, 2000) They are considered assumptions about variables X_i and the error term are considered relevant to achieve a valid interpretation of the estimated regression values. The most important assumptions are:

Linear regression model: The regression model is linear in the parameters, although may or may not be linear in the variables.

Fixed values of X, or independent X values of the error term: The values taken by the regressor X can be considered fixed in repeated samples (the case of the fixed regressor), or have been sampled with the dependent variable Y (the case of stochastic regressor).

The average value of the disturbance μ_i is zero: Given the value of X the mean or expected value of the random disturbance μ_i term is zero. Symbolically, we have $E(\mu_i|X_i) = 0$ if X is not stochastic, $E(\mu_i) = 0$

Homoscedasticity or constant variance of μ_i : The variance of the error term or disturbance, is the same no matter the value of X. Symbolically, we have: $\text{var}(\mu_i) = E[\mu_i - E(\mu_i|X_i)]^2$ where var means variance.

There is not autocorrelation between the disturbances: Given any two values of $X, X_i Y, X_i (i \neq j)$, and the correlation between any two μ_i and μ_j is zero. In few words, these observations are sampled independently. Symbolically, $(i \neq j) \text{cov}(\mu_i, \mu_j|X_i, X_j) = 0$ or $\text{cov}(\mu_i, \mu_j) = 0$, if it is not stochastic. Where i and j are two different observations and cov means cov aridity.

2.1.2 Box Jenkins

(Chatfield, 2003) The methodology proposed in this research is called "Box Jenkins," it is known as ARIMA models and was formalized by Box and Jenkins in 1976, so they are also called Box Jenkins models. This approach parts from the fact that the time series (monthly or quarterly and in some cases biannual and annual), that it is about predicting is generated by a stochastic process whose nature can be characterized by a model, this research on the rate retention, consider annual data of different variables.

(Chatfield, 2003) Points out "Basically the Box-Jenkins methodology consists in to find a mathematical model that represents the behavior of a range of data and allows to do forecasting only by entering the corresponding time period." The Box-Jenkins method provides predictions without the necessity of existence of any type of previous condition, besides being parsimonious with respect to the coefficients. Box Jenkins Methodology considers four steps:

Step 1. Identification. Also it is known as the specification, that is, to find the appropriate values of p, d and q.

Step 2. Estimation. After identifying the appropriate values of p and q, the next step is to estimate the parameters of the autoregressive terms and mobile average included in the model. Sometimes, this calculation is performed by simple least squares, but other we have to appeal to nonlinear estimation methods (in parameters). As this work is carried out now through routines in various statistical packages, in the practice it is not precise to worry about the mathematical developments of the estimation; the student interested in the topic can consult the references.

Step 3. Diagnostic test. Also it is known as checking, after selecting a particular ARIMA model and estimate their parameters, we try to see if the selected model fits to the data in a reasonably way, because there may be exists other ARIMA model which also do it. That is why the design of these models Box-Jenkins is an art rather than a science; great skill is required to select the correct model.

Step 4. Exploitation or prognosis. One reason for the popularity of the process of building ARIMA models is its success in the prognosis. In many cases, the prognosis obtained by this method are more reliable than those obtained from traditional econometric models, particularly in the case of short-term prognosis.

2.2 Results

The applying regression analysis trough OLS and using methodology Box Jenkins for the application of econometric model the same as already explained, consider four phases: specification, estimation, validation and exploitation so the modeling is presented below:

2.2.1 Specification

The model, in the initial phase will be constituted for the variable explained it would retention rate (RR); and the explanatory variables is selected variables as: Full time teachers, PhD, teachers who only have third level,

fourth level teachers, and the average number of students per full time teacher. The resulting model equation would be:

$$\text{RR} = B_0 + B_1 * (\text{FTT}) + B_2 * (\text{PhD}) + B_3 * (\text{FTT}_{\text{III}}) + B_4 * (\text{FTT}_{\text{IV}}) + B_5 * (\text{SFTT}) + \varepsilon_i \quad (1)$$

Where:

RR: retention rate; B0: intercept parameter; B1: Parameter Variable FTT; B2: Parameter Variable PhD; B3: Parameter Variable FTT_{III}; B4: FTT_{IV} variable parameter; B5: Parameter Variable SFTT; FTT: Number of full-time teachers at the institution; PhD: Number of teachers with PhD degree; FTT_{III}: Number of teachers who have only third level degree; FTT_{IV}: Number of teachers who have fourth level (Masters); SFTT: Average full-time students per teacher.

2.2.2 Estimation

ARMA processes, can become tedious and / or cumbersome because after selecting the particular model and to estimate its parameters, it is to see if the selected model fits to the data as reasonably good because there may be other ARIMA model as well do it, so if it fails the diagnostic test should begin the process again.

With the tool of econometrics package (Eviews) we have the estimation of the parameters and so it is presented in Table 1 below. (Medina Ibañez, 2018) It is important to show that there are other statistical and econometric packages as STATA, R, Gretl, among others that may be used for estimation, however it is considered that the Eviews software presents an interface that facilitates the Box Jenkins. The resulting equation will be the following:

$$\begin{aligned} \text{TR} = & -1,09454079 + 0,08326 * (\text{FTT}) + 0,01034 * (\text{PhD}) + 0,00038 * (\text{FTT}_{\text{III}}) - 0,00394 * (\text{FTT}_{\text{IV}}) \\ & + 0,00786 * (\text{SFTT}) + \varepsilon_i \end{aligned}$$

Table 1. Model estimation

Dependent Variable: TR				
Method: Least Squares				
Date: 02/20/17 Time: 9:12				
Sample: 2010 2016				
Included observations 7				
Variable	coefficient	Std. Error	t-Statistic	Prob.
C	-1.094541	1.679455	-0.651724	.6323
FTT	0.008326	0.003867	2.153328	.2768
PHD	0.010349	0.026915	0.384492	.7663
FTT _{III}	0.000388	0.005126	0.075721	.9519
FTT _{IV}	-0.003940	0.007553	-0.521672	.6939
SFTT	0.007859	0.016885	0.465448	.7227
R-squared	0.917602	Mean dependent var		0.745200
Adjusted R-squared	0.505612	SD dependent var		0.127798
SE of regression	0.089858	Akaike info criterion		-2.212794
Sum squared resid	0.008074	Schwarz criterion		-2.259157
log likelihood	13.74478	Hannan-Quinn criter.		-2.785829
F-statistic	2.227245	Durbin-Watson stat		2.362297
Prob (F-statistic)	0.467486			

It is important to show that R² is of 0.92; which means that the variables are related in 92%, which is good for the model.

2.2.3 Validation

As it can be seen the P-value are greater than 0.05; however, since it is a small population statistic "t" will be used to validate the parameters of the explanatory variables. With a probability of 95%, also known as confidence level, and with 4 degrees of freedom (n-1) (5-1 = 4) we have a statistical t of 3.7469.

Table 2. Validation parameters

<i>parameters</i>	<i>coefficients</i>
Interception	-1.09454079
Variable X 1	0.00832645
Variable X 2	0.01034869
Variable X 3	0.00038815
Variable X 4	-0.00394002
Variable X 5	0.00785924

<3.7469

As all coefficients, parameters are lower to the statistical "t" then they validate each of these, that is to say all coefficients are statistically significant for the model.

2.2.4 Exploitation

It is important to show that through this equation can predict the retention rate for future periods, however the objective of this research is to determine the variables that influence student retention rather than pronosing them.

Once modeled series until obtain the results, we proceed to interpret each parameter according to the explained variable "Retention Rate".

3. CONCLUSION

Criteria academy explained the retention rate, so that if all explanatory variables were zero (0), the rate of retention will be -119%, that is to say, the desertion rate increase by more than 100%. For every full-time teacher, the retention rate increases by 0.8%, so it is considered important to have full-time teachers. Each PhD. contracted by the institution the retention rate increases by 1%. The teachers who have third level degree contribute to the retention rate by 0.04%, that is to say its contribution is very marginal. For each teacher that have fourth level degree, the retention rate decreases by 0.04%, this can be seen in an unreasonable way, but teachers who have fourth level are usually more demanding which could provoke motivations to students to leave their college career. Finally, but not less important, we have that as increases the number of students per teacher full time by one unit, the retention rate increases by 0.8%.

REFERENCES

- Asamblea Nacional Ecuador, 2010. *Ley Orgánica de Educación Superior*, Quito: s.n.
- CEAACES, 2015. *Modelo de Recategorización de Universidades y Escuelas Politécnicas*, Quito: s.n.
- Chatfield, C., 2003. *The Analysis of Time Series: An Introduction*. Washington D. C, Taylor & Francis Group .
- Gujarati, D. & Porter , D., 2010. *Econometría*. Mc Graw Hill, México.
- Larrea, E., 2015. *Unidad Curricular de Titulación*. Consejo de Educación Superior CES, Quito.
- Medina Ibañez, J. M., 2018. *Aplicaciones Informáticas. Econometría*, 1(1), pp. 1-8.
- Novales Cinca, A., 2000. *Econometría*. Mc Graw Hill, España

EVIDENCE-BASED TEACHING FOR PERSONALIZED LEARNING

Boong Yeol Ryoo, Shannon H Degenhart, and Gary Bolts

Department of Construction Science, Texas A&M University, College Station, Texas, U.S.A.

ABSTRACT

The Construction management curriculum is an industry-based curriculum. For some courses, prior knowledge gained from prerequisite courses is critical for successful completion. In the scheduling class, higher-order thinking and problem-solving skills are necessary to develop a realistic project execution plan. The course assignments are often not prescriptive and may have multiple correct answers. Therefore, responding to the needs of individual learners becomes a great challenge for instructors. This paper focuses on the creation of a framework able to guide students based on their individual learning needs and behaviors. The framework will: 1) establish a baseline of student knowledge and skills prior to starting the course; and 2) categorize student learning styles by type of mistake to individually provide feedback. The framework will use a reverse logic approach, which works backward from the learning objective to a set of related core concepts to determine prior knowledge necessary to complete each assignment.

KEYWORDS

Online Platform, Personalized Guidance, Learning Behavior, Smart Tutoring

1. INTRODUCTION

Scheduling is an important tasks in project management. In construction, scheduling aims at developing a project operation plan known as the master schedule, or the contractor's construction schedule, for a project. In most construction management programs, the student has to complete several construction management courses before the scheduling class. Generally, the scheduling class has a series of prerequisite courses such as Plan Reading, Materials and Methods, and Estimating. In addition, learning scheduling software is essential for master scheduling.

Scheduling requires higher-order thinking and problem-solving skills, as scheduling is not prescriptive with only one correct answer. Individual students are asked to develop a realistic project operation plan that satisfies the owner's requirements and resource availability (PMI 1989). Many students lack the critical prior knowledge necessary to succeed, don't understand how formula are derived and must search for formula or equations to solve problems. Some students may have enough prior knowledge but are not able to successfully apply their knowledge to complete the multiple steps required to solve the real-world problems used in class. Additionally, without practical job experience, it is difficult to create master schedules which meet industry standards.

Student learning styles vary from visual learners to solitary learners. To effectively improve students' learning experience in online classes, it is necessary to diagnose students' learning styles (Davison et al. 2012). Özreçberoğlu and Çağanağa (2018) used two types of problems, routine and non-routine, to understand how students succeed in solving problems. The scheduling class is a suitable course for both types of problems.

The primary goal of this research is to provide students, faculty, and the university a self-paced, self-assessment framework based on students' performance data to minimize mistakes or omissions, for both face-to-face and online classes. This research provides an easily adaptable online platform for instructors to use in order to enhance student mastery of complex materials. An artificial intelligence model is used to diagnose students' learning behaviors for personalized learning.

The intended goals of this research are: 1) redesigning the target course based on students' learning styles and performance; 2) transforming course materials and assignments; 3) developing an artificial intelligence tutoring program which provides individualized feedback; 4) and creating guidelines for operation and data migration. The final product resulting from this project will be compatible with online teaching platforms.

2. FRAMEWORK OF THE MODEL

2.1 Nature of Mistakes and Approaches

The scheduling class requires knowledge from the plan reading, methods and materials, and estimating classes. Some students may have enough prior knowledge but are not able to successfully apply their knowledge to complete the multiple steps required to solve the real-world problems used in the class. The scheduling class contains two types of problems which require the basics of process time calculation and a special way of process planning in scheduling. The former is a Math-related problem and there is an order to reach the answer. The latter requires creative approaches and planning skills some of which do not require a particular order to reach an answer.

Diagnosis of the type, frequency, and trend of student mistakes is integral in identifying types of mistakes in learning. This project studied two aspects of computing and critical thinking abilities to understand students' problem-solving approaches. Math-related mistakes were categorized into four types of errors: 1) careless error, 2) computation error, 3) precision error, and 4) problem-solving error (Carter 2015). Mistakes were then categorized by common errors (misunderstanding the instructions), inattention errors (checking the final answer if it makes sense), and basics errors (understanding of basics). Mason et al. (2012) and Yerushalmi et al. (2012A, 2012B) proved self-diagnosis helped students correct their mistakes. Students' problem-solving processes can reveal the reasoning behind each step. A reverse logic approach provides students with personalized feedback, and by using self-assessment problems, an individual student can verify if he/she is meeting intended goals.

2.2 Connecting Academic Competence and Proficiency

A course content map was devised for students to measure their academic competence and performance. One content map was devised for problem calculation and one was devised for critical thinking problems. The maps illustrated the content necessary for a student to review in order to master a topic. Figure 1 shows how the project progress is calculated. There is one correct answer to the question. In order to correctly answer the question, students must correctly complete a progression of three separate questions (steps), the answer of which must be used in the question (or step) which follows it in order to correctly answer the original question. Therefore an early mistake can jeopardize the whole problem.

The types of mistakes and frequency the student makes the mistakes are used to assess students' knowledge and performance. The type of question illustrated in Figure 1 focuses on improving academic proficiency and performance ability. Carter's categories (2015) were used to find the general tendency of students' mistakes. Given specific feedback (academic and attentiveness), students can correct their mistakes. To categorize the types of mistakes, the problem is broken into sub-problems as follow:

- 1) determine the present project conditions (BAC, PV, EV, and AC)
- 2) calculate the project progress (PC)
- 3) calculate the current progress (SV, CV)
- 4) calculate the current progress ratios (SPI, CPI)
- 5) calculate the estimated cost for the remainder of the project (ETC)
- 6) determine the final cost (EAC), and
- 7) compare EAC to BAC

Figure 2 illustrates the typical master scheduling sequence. There are many potential correct answers for the problem. Personal academic competence and critical thinking skills ultimately result in the successful completion of the problem. Students' ability to apply critical prior knowledge to solve the real-world problem through this process of creating a master schedule is a major concern. In this problem, the focus should be on assisting students to develop individual competency and proficiency at each step. There should be focus on

individual ability in developing a roadmap (timed task) for the project. Given the drawing and specifications, the scope of the project is identified. To provide students personalized feedback, the program breaks the problem into sub-tasks as follow:

- 1) define the scope of work (WBS/WP)
- 2) choose crew tasks
- 3) select a construction method
- 4) Select a work crew (skills and equipment) for each task
- 5) determine task sequences
- 6) determine the crew cost and the daily output
- 7) perform quantity take-off, and
- 8) calculate the task duration

In the sub-tasks, the selection of a construction crew affects the cost, duration, and work sequences of the project tasks, and relies entirely on the selection of the construction method. The program contains general guidelines prepared for general problems in scheduling. Alternative answers are provided for students who may have different ideas about completing the project. Each student receives appropriate suggestions for each response while solving a problem. To categorize students' problem solving skills, the program identifies if a student:

- 1) understands the problem
- 2) follows the directions
- 3) lacks prior knowledge
- 4) knows all of the essential steps
- 5) makes careless errors
- 6) follows the arithmetic rules, and
- 7) completes all of the steps

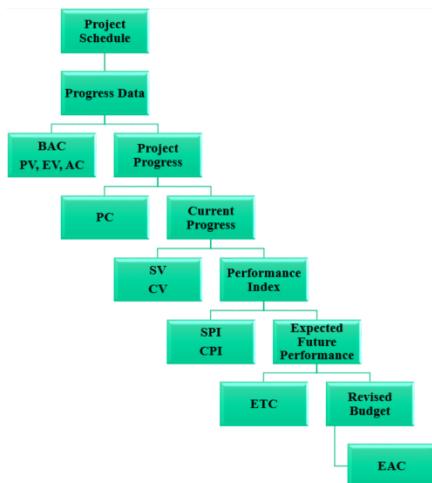


Figure 1. Example of calculation problem

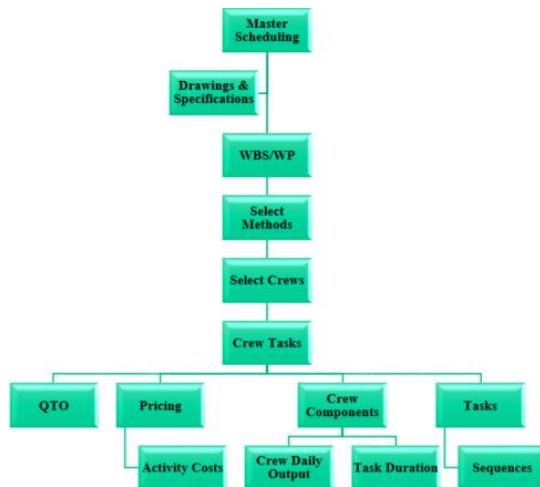


Figure 2. Example of critical thinking problem

2.3 Smart Tutoring Program

A rule-based artificial intelligence program is used to provide students with an individualized, interactive, and self-paced online learning experience. It identifies student strengths and weaknesses. It provides immediate supplemental and/or remedial instruction in weak areas. The AI engine maps out the course materials based on the course objectives. Then the program uses a reverse logic model which works backward from the learning objective to a set of related core concepts. Figure 3 shows an example process to illustrate how the project duration (or learning objective) will be determined using this method.

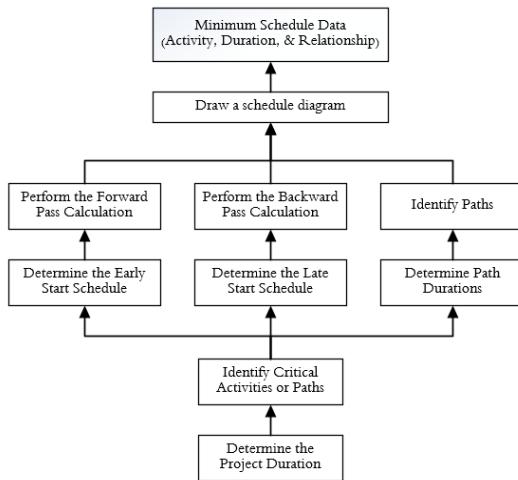


Figure 3. Reverse learning process for scheduling calculation

2.4 Data Collection and Analysis

Students' performance and proficiencies have been collected and analyzed in the past eight years. Data from a total of 24 semesters, including summer, were used to observe students' academic performance and proficiencies. At first, a few students received a grade of "A". Observation of longitudinal student data indicates, about a 10 percent increase in student grades over the eight-year observation period.

3. OBSERVATION AND FINDINGS

The result of the eight years of student performance data indicated student performance was associated with students' learning patterns, with no differences due to semester. Student performance data indicated students didn't have the adequate prior knowledge or enough proficiency to apply their knowledge in master scheduling. Low student performance in this course was attributed to the need for students to use knowledge gained from multiple courses to complete assignments and use higher-order thinking and problem-solving skills as assignments are not prescriptive with only one correct answer. Analysis indicates a positive improvement in students' performance after correcting their mistakes manually. An AI tutoring program is to illustrate individualize instruction. The proposed eight-step process is used to identify the types of mistakes. The AI program determines the types of instructions needed to address individual student needs.

ACKNOWLEDGEMENT

This research is being supported by the Texas A&M University. The author wishes to express their appreciation to all those who participated in this research.

REFERENCES

- Carter, S. 2015. "Error Analysis Sheet & Types of Errors Notebook Page," <https://mathequalslove.blogspot.com/2015/08/error-analysis-sheet-types-of-errors.html>, viewed on August 8.
- Davison, M. et al 2012. "The Learning Styles, Expectations, and Needs of Online Students," College Teaching, vol. 54, no. 1, pp. 185-189.
- Dawkins, P., 2006. "Common Math Errors," <http://tutorial.38.math.lamar.edu/terms.aspx>

- “The Most Common Errors In Undergraduate Mathematics,” <https://math.vanderbilt.edu/schectex/comerrs/>, viewed on March 1, 2018
- E. L. Prentis, E. L. 1989. “Master project planning: scope, time and cost,” Project Management Journal, vol. 20, no. 1, pp. 24–30.
- Gross, S. P. and Dinehart, D. W. 2012. “Identification Of Common Student Errors In Solving Fundamental Mechanics Problems,” 42nd Annual Frontiers in Education Conference, American Society for Engineering Education.
- Özreçberoglu, N. and Çağanağa, C. K. 2018. “Making It Count: Strategies for Improving Problem-Solving Skills in Mathematics for Students and Teachers’ Classroom Management,” Eurasia Journal of Mathematics, Science and Technology Education, vol. 14, no. 4, pp. 1253-1261.
- Yerushalmi, E. et al, 2012A. “What do students do when asked to diagnose their mistakes? Does it help them? I,” An atypical quiz context, Phys. Rev. ST PER, vol. 8. no. 2.
- Yerushalmi, E. et al, 2012B. “What do students do when asked to diagnose their mistakes? Does it help them? II,” An atypical quiz context, Phys. Rev. ST PER, vol. 8. no. 2.

EDUCATIONAL DATA ANALYSIS APPLYING A KDD METHODOLOGY

William Villegas-Ch¹, Diego Buenaño-Fernández¹ and Sergio Luján-Mora²

¹Universidad de Las Américas, Av. de los Granados E12-41y Colimes esq., Ecuador

²Universidad de Alicante, Carretera San Vicente del Raspeig s/n 03690, Spain

ABSTRACT

The educational institutions in their academic management make use of learning management systems. These systems contain a large volume of information about the activities and resources used by teachers and students. This information in most cases is not exploited to the maximum as if they do company through business intelligence. Business intelligence aims to identify patterns of behavior found in the data, applying methodologies, models and data mining techniques. The application of these tools to an educational model allows to discover in time factors that directly impact the performance of students. For this work, the application of the method of knowledge discovery in databases where patterns will be identified is considered, the information will be processed in such a way that is novel, useful and above all understandable. The application of this method of data mining and with the identification of patterns the institutions will be able to apply corrective in the academic management and it will be possible to improve the educational quality.

KEYWORDS

Data mining, e-learning Analytics, Moodle, Business Intelligence, KDD

1. INTRODUCTION

Learning management systems (LMS) store large volumes of information that are often not exploited by educational institutions. The use of this information encourages the continuous improvement of the content and structure of virtual courses contained in LMS. Under this scheme, the pedagogical model consists of web-based virtual courses supervised by a tutor (Dalsgaard 2006, 1-5). This work is developed with the use of online resources, activities, forums and various shared services. The data are gathered in the LMS and are based on the activities and didactic resources that are offered to the students (Lonn and Teasley 2009, 686-694). The data obtained when evaluated with the use of data mining leads to the possibility of choosing the most appropriate resources and adapting it to the characteristics and personal interests of each stakeholder in the educational context.

Educational data mining (EDM) models and tools create a bridge to the fields of traditional statistics through pattern recognition and learning (Siemens and Baker 2012, 252-254.). In this article we propose the analysis of the different models and tools that allow the exchange of experiences with regard to how engineering students learn in marketing. Section 2 details the processing required for data analysis using the main data mining techniques; Section 3 performs an analysis of the results obtained; finally, section 4 presents the conclusions and makes recommendations for future research.

2. METHOD

The discovery of knowledge in databases (KDD) It is a methodology proposed by Fayyad in this methodology are proposed five phases that are: selection, preprocessing, transformation, data mining, interpretation and evaluation (Fayyad, Piatetsky-Shapiro and Smyth 1996, 82-88). Below is a description of each of these phases within KDD:

Selection, is responsible for developing a process of understanding the domain of study and establishing the objectives of what is intended. In this phase, the most important data sources are recognized and over which control will be exercised. It is also important to include all related metadata, the size and the amount of data, and its formats. The relevant data for the application can be in relational databases, text files, photographs, customer transactional databases, logs or web logs. It is important to homogenize the formats so that they are easier to process and analyses.

Preprocessing, aims to improve the quality in the data of the original repositories that are incomplete or have missing attributes, have noise that is visualized as errors, or present inconsistency. These types of errors can confuse the data mining process and lead to unreliable results.

Transformation, consists of making syntactic modifications on the data that suppose changes in the mining of data. Modifications can be made through digitization, discretization, etc. The result is a set of rows and columns called minable view. The minable view integrates data from different sources, cleans, selects, transforms and types it in order to prepare them for modelling.

Data mining, data mining allows us to explore and delve into the immensity of data and discover the patterns and models present in the data. For the exploration it is used tools given by algorithms, which are a series of instructions or rules established in a computer program (Berry and Linoff 1997). An algorithm allows us to process a set of data to obtain new information about that same dataset.

Interpretation and evaluation, in this phase the visualization of extracted patterns and models is implied. It is important to understand the difference between the two terms where patterns are local structures that make statements only on a space restricted by variables. Models are global structures that make statements about any point in the measurement space; for example, models can predict the value of another variable. With this clarification, it is important that the results are presented in an understandable format so that the results are useful.

The information extraction process is a non-trivial process to discover potentially useful knowledge and information within the data that an educational institution possesses. If you follow step by step the process that has been indicated in very easy and decisive use. It is important to consider that this process does not end with the discovery of patterns or projections of students, it is important that each institution uses this knowledge to propose activities that focus on active learning for which decision-making is a fundamental pillar for educational improvement.

2.1 Field of Study and Selection of Data

The work is done with data from a university that has more than 7,000 students divided into four academic programs and 25 careers. This study has focused specifically on the distance academic program where the population is more than 2000 students. The distance study modality has already established the norms and policies that manage the educational model, which is based on the use of ICTs as a basis for their development (Vázquez-Cano, and García 2015, 62). The ICT in reference to the computer systems with which the university counts, allows the storage of information on activities that the students develop, on which a model will be applied that allows the analysis of data. The computer system for management is the Moodle e-learning platform used as the main tool for educational development. In Moodle a pyramidal structure has been deployed where the careers cover a determined number of virtual classrooms for each course, in the table 1 the study population that has been raised for this work is detailed. (Costa, Alvelos, and Teixeira 2012, 334-343).

Table 1. Distributive of the study population

Career	Number of academic periods	Number of subjects of the career
Business Administration	9	53

In the sample analyzed, it is considered that for each subject, there is a virtual classroom created in the Moodle platform to which a regular student is assigned once throughout his career. In data we have that in a 9-period career, in the 53 subjects there will be around 1325 students enrolled. This number of students has generated useful information that has been stored in the Moodle database, generating around 400 gigabytes of information that is not being used. In order for our analysis to be manageable, it has been decided to segment

the analysis only to the subject of office automation III belonging to the third semester. The analysis considered is done to the data within a period of two years (four academic periods). By regulation each course has a maximum number of students of 25 students, this consideration is detailed in table 2.

Table 2. Sample analyzed during four periods

Period	Chair	Number of studen
Period 2015	Office automation	23
Period 2015	Office automation	21
Period 2016	Office automation	18
Period 2016	Office automation	25

2.2 Application of Data Mining in a Case Study

As previously indicated, data downloaded from the Moodle platform is found in Excel files in xml and csv format. Once these data have been filtered and passed through a first transformation with respect to empty cells, they will be loaded into a repository (Goldstein and Katz 2005, 1-12). The data repository mining_moodle is a relational database created in the SQL server.

To load the data to the base mining_moodle was created a project of integration of services with the use of the tool Integration services of visual studio. What is sought with this project, besides loading the data, is to use the characteristics offered by this tool. Integration services allow the user to extract the information required by the system regardless of the form in which it is held: Excel, Access, plain text files, etc. This tool, besides having several import characteristics, facilitates the possibility of filtering and processing data. With these options we can debug the data we require and safely apply mining methods (Kimball and Caserta 2011).

In order to analyze the data, a project was generated with the use of analysis services. The aim of this step is to create dimensions from the mining_moodle repository. This allows us to generate a cube where you can view the different views and create a data mining structure. First we have to tell the tool which repository we are going to use, given that the data are already clean and have been transformed according to our needs.

Analysis services provide many advantages in relation to the analysis of data between them that we can include a structure of data mining (Ferrari, Russo, and Webb 2012). This structure will depend on the algorithm we want to use. The algorithms follow the concept already analyzed previously, depending on the application. For example, as the algorithm to use will involve grouping, when we submit our cube to this algorithm we will process and observe the results that it presents (Erl 2008, 1).

For the generation of the data mining structure, we simply right click on the tools. This will display an assistant which allows us to choose the desired algorithm. Once this process is done we simply execute, and the program will be responsible for the analysis and for the presentation of the data. As with the cube generation, its results can be downloaded to Excel for a better view of the analysis (Baker 2010).

3. ANALYSIS OF RESULTS

The application of data mining to data repositories has been used efficiently in companies that have focused on data intelligence to recognize patterns and trends in customers. As a first option we tried to apply the tools dedicated to the EDM as Weka or Gismo (Hall, Frank, et al., 2009, 10-18). The important thing in the process is to choose the tools of extract, transformation and load (ETL) as well as the data mining algorithms. In this case, we opted for Microsoft SQL tools. Its advantage is that it has several tools that allow the generation of an analytical project from the creation of a data repository to the final analysis. The advantage of using this package is that it allows the processing of data involving ETL. In this way it can perform data cleaning before the analysis. The creation of a project with the use of SQL and analysis services allows the updating of information without the need to create a new project, which is an advantage with respect to EDM tools. This advantage can be observed when updating one of the Moodle reports so that they are automatically transformed and loaded from the .xml or .csv files (Romero, Ventura, and García 2008, 368-384).

The SQL server data tools allow the creation of a data mining project where we can opt for the algorithm that is aligned with what we want to perform in our exercise. In the analysis of EDM, the method that each technique uses to generate information that is useful for decision making is analyzed. Based on this information, preventive or corrective measures can be taken in terms of the different activities in order to create a personalized educational environment. With the data obtained, decisions can be made regarding each group. For example, the creation of virtual tutorials to help students who are more likely to fail the subject, and changing the mechanisms of evaluation. Personalized education seeks to strengthen the student's skills and help them reduce his/her weaknesses by providing resources and activities that are aligned with his/her needs. It must be noted that the application of data mining has several formalities that must be taken into account with regard to the representation of data - probabilities, rules, trees and a series of statistical methods. These points allow stakeholders to save time on tasks such as finding one or a group of individuals with similar characteristics.

4. CONCLUSION

In the process of developing this work we were able to obtain several clarifications on the use of data mining in e-learning platforms. In the first place we can emphasize that the thrust that today is given to education has led to the creation of modalities that allow the student to manage the convergence between work and educational activities (Macfadyen and Dawson 2012). For this reason, the use of e-learning platforms should be emphasized as a component of virtual education or distance learning. In these education modalities, it is essential to look for methods that allow the generation of compliance control actions and alert systems. These actions will enable those involved to take immediate action to assist the student in meeting the expected goals. The actions that can be taken vary according to the approach of each institution, and the measurement parameters that government bodies dealing with academic control require. Among these actions we have the segmentation by groups that present the same patterns that allows us to offer a personalized education. Another type of action to take regarding desertion can be based on the data obtained from a data mining analysis that would allow us to apply correctives in order to lower these indices.

Technically, when applying the EDM models as shown in the case study examined in this paper we have to state that improvements could be made in terms of implementation. One of the improvements is that the extraction, transformation and loading could be done directly from the Moodle database found in MySQL. This would ensure that the data is always up-to-date and therefore so too would be the results of the analysis. Considering this proposal and taking into account that Moodle already has a transactional database, the efforts with regard to our analysis focused on the transformation and the data load means that we could even opt for a multidimensional basis. These advantages will help us focus directly on the perfecting of the data cubes and the data mining analysis. This process will be followed up in a future work.

With regard to the tool used, we can conclude that its use is very practical and that it has several advantages such as the application of ETL techniques which facilitate the obtaining of clean data that guarantees the results of the analysis. In relation to the other tools mentioned throughout the research (Weka and Gismo) we have the disadvantage that SQL and visual studio are licensed software. As for the manipulation part of the tool is very practical, and its programming is graphical which will help people with little experience in handling this tool.

With the results obtained from the analysis, we can implement measures that strengthen the processes by which students generate knowledge. Another measure that can be used is the identification of groups that have similar needs so that we can apply methods that help students to better understand the subjects which they are studying.

REFERENCES

- Dalsgaard C., 2006. Social software: E-learning beyond learning management systems. *European Journal of Open, Distance and e-learning*, Vol. 9, No. 2, pp. 1-5.
- Lonn S., and Teasley S. D., 2009. Saving time or innovating practice: Investigating perceptions and uses of Learning Management Systems. *Computers & Education*, Vol. 53. No. 3, pp. 686-694.

- Siemens G. and Baker R. S., 2012. Learning analytics and educational data mining: towards communication and collaboration. *In Proceedings of the 2nd international conference on learning analytics and knowledge*, pp. 252-254.
- Fayyad U. M., Piatetsky-Shapiro G., and Smyth P., 1996. Knowledge Discovery and Data Mining: Towards a Unifying Framework. *In KDD*. Vol. 96, pp. 82-88.
- Berry, M. J., and Linoff, G. 1997. *Data mining techniques: for marketing, sales, and customer support*. John Wiley & Sons, Inc..
- Vázquez-Cano E., and García M. L. S. 2015. Analysis of risks in a Learning Management System: A case study in the Spanish National University of Distance Education (UNED). *Journal of New Approaches in Educational Research*, Vol. 4, No. 1, pp. 62.
- Costa C., Alvelos, H., and Teixeira, L. 2012. The use of Moodle e-learning platform: a study in a Portuguese University. *Procedia Technology*, Vol. 5, pp. 334-343.
- Goldstein P. J., and Katz, R. N. 2005. Academic analytics: The uses of management information and technology in higher education. *Educause*. Vol. 8, pp. 1-12.
- Kimball, R., and Caserta, J. 2011. *The Data Warehouse? ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. John Wiley & Sons.
- Ferrari, A., Russo, M., and Webb, C. 2012. Microsoft SQL Server 2012 Analysis Services: The BISM Tabular Model. Pearson Education.
- Erl, T. 2008. *Soa: principles of service design*. Upper Saddle River: Prentice Hall, Vol. 1.
- Baker, R. S. J. D. 2010. *Data mining for education*. *International encyclopedia of education*, Vol. 7. N0. 3, pp. 112-118.
- Hall M., Frank E., et al., 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, Vol. 11 No. 1, pp. 10-18.
- Macfadyen, L. P., and Dawson, S. 2012. Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *Journal of Educational Technology & Society*, Vol. 15, No. 3, pp. 149.
- Romero C., Ventura S., and García E. 2008. Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, Vol. 51 No. 1, pp. 368-384.

SERIOUS GAMES AND BOARD GAMES IN THE COMPUTER SCIENCE CLASS

Dawid B. Jordaan

*School of Computer Science and Information Systems, North-West University, South Africa
TELIT-SA, North-West University, South Africa*

ABSTRACT

Various studies reveal that students develop a negative mind set towards computer programming because they experience programming as a difficult subject. This is alarming when taken in consideration that the software development industry is one of the fastest growing areas in the labour market currently. Enrollments figures at tertiary institutions in computer science courses are low and dropout rates are high. In addition to the perception that programming is difficult, outdated teaching and learning strategies and irrelevant course material could be to blame for this situation. Young people engage with modern technology and this is often not what they find in the classroom. Failure to implement appropriate and interactive teaching methods in the learning environment often lead to wasted resources, learner apathy and discouragement. Effective teaching methods and approaches should inspire, empower and encourage students and this is especially true in complex, technical subjects such as computer science.

Students find the competitive, fast-paced and interactive environment which serious games provide appealing. Studies show that games could be a powerful and effective tool to create such a learning environment. Although progress has been made to incorporate digital educational games into the learning environment, research on understanding the value that serious games can add to learning in computer science courses at higher education institutions is limited.

The value of board games has been recognized in corporate settings but to a limited extend. Most teachers do not recognize the impact and true value of board games designed for learning. The purpose of this “work in progress” study is to investigate the possibility to use board games to enhance the learning experience of students in the computer science class. Insights gained from this study may assist educators in the designing and planning the implementation of board games as part of the learning experience in the computer science class.

KEYWORDS

Games, Board Games, Serious Games, Programming, Computer Science

1. INTRODUCTION

Peters and Pears (2012) summarise the perceptions of various authors on reasons for low enrolment and high dropout rates in computers science courses. Besides from misconceptions such as a “nerdy” image and anti-social lifestyle which is associated with computer-related occupations, low levels of motivation to study computer science prevail due to the perception that computer programming is a difficult to master. High dropout rates and low throughput figures contribute to this perception. Furthermore, students often experience the discipline of learning of computer programming as boring (Ali and Smith, 2014).

Students are familiar with advanced digital means of learning and communicating which includes playing games and across the internet, with social media and through other various technologies (Nag, Katz, et al., 2013). They are often discouraged and disappointed by the absence of technology in the learning environment (Prensky, 2001).

The value of serious games and board games has been recognized in commercial environments but to a limited extend. Most teachers do not recognize the impact and true value of serious games and board games designed for learning.

2. BACKGROUND

2.1 Computer Science

According to the US Bureau of Labour Statistics, careers in software development are in demand with a predicted 45.3 per cent growth in software design careers up to 2018 (Jackson and Moore, 2012). Contrary to this several researchers (Gomes and Mendes, 2010; Heersink and Moskal, 2010; Muratet, Torguet, et al., 2011) report a decline trend in the number of enrolments for computer science at HEI all over the world. In South Africa the enrollment of computer science students is steady at a rate of 3.4 percent, but this is disappointing low when compared to the strong increase of 27.4 percent in most other subjects in the period from 2005 to 2010 (Kirlidog, van der Vyver, et al., 2016). The ongoing enrolment crisis in computer science can be blamed on unrelated course material and learning content. This is a serious concern as we live in a technology driven society.

Khaleel, Ashaari, et al. (2015) report that students find it difficult to learn technical subjects, such as programming, as they lose interest and eventually develop a negative mind set for the subject. This creates a perception that programming, which is an integral part of almost all computer science curricula, is a difficult skill to master and is reported as a major problem when reasons for the decline in interest in studying computer science are addressed (Peters and Pears, 2012; Khaleel *et al.*, 2015). Higher order thinking skills are required in the programming environment (Tan and Rahaman, 2009), especially the ability to solve problems (Kurkovsky, 2013). Nag *et al.* (2013) are of opinion that the ability and fundamental skills to solve problems are skills that today's digital environment requires from students. Motivation and interactive involvement in what they are doing, and when programming problems are presented in a context that students can relate to (Shabanah and Chen, 2009; Tan and Rahaman, 2009), improve students' determination to attempt to solve problems (Stanescu, Stefan, et al., 2011).

2.2 Learning Environment

Students are digitally oriented with different needs and therefore a change in teaching strategy is required to meet their needs (Prensky, 2001). Different learning requirements of technology savvy students of the 21st century (Bidarra, Figueiredo, et al., 2015) require different learning environments, but according to Conneely, Girvan, et al. (2012) the education systems have primarily remained the same. Students must be active participants rather than passive observers in the classroom (Prensky, 2001) and learning must be relevant to the requirements of the changing world (Kaiser and Wisniewski, 2012). The traditional teaching and learning environment has proven to be less effective to digitally connected students, especially in a resource constrained learning environment (Rooney, 2012; Cunningham, 2015). Their constantly engagement with technology means that participatory, sensory-rich environments and experiential or discovery-based learning activities will appeal to students (Annetta, Minogue, et al., 2009; Heersink and Moskal, 2010; Husain, 2011). Zeeman (2014) calls for a change in teaching strategy for the new generation of digitally oriented students.

2.3 Technology in the Learning Environment

While Schwabe and Göth (2005) state that students are better motivated when technology is used in the learning environment and Mbogo, Blake, et al. (2014) concur that students can benefit from technology. According to Coughlin and Lemke (1999) technology fast-tracks and enriches basic skills, is extremely motivational since it provides ease to digitally oriented students, facilitates new fields through games, simulations, three-dimensional models, etc., and prepares students for industry at a young age. Mobile technology allow students to learn anywhere, anytime and at their own pace (Bartel and Hagel, 2014).

2.3.1 Games

The idea of adopting a playful approach in education for motivational purposes is a well estalished research concept (de Villiers and Blignaut, 2016). Progress has been made to incorporate digital games into the learning environment ((Heintz and Law, 2012) because it contributes to an effective and powerful learning

environment (Wrzesien and Alcañiz Raya, 2010). Video games have increasingly become popular (Su and Cheng, 2015) and according to Zyda (2005) researchers have considered the serious games concept in hopes of enticing learners because of the exceptional motivation and user involvement that games offer.

As students easily accept new technology they are motivated by the playful element and other attributes like rewards, interactivity, scores and challenges (Vogel, Greenwood-Ericksen, et al., 2006) in serious games (Krassmann, Paschoal, et al., 2015) which in turn may yield positive results in terms of learning (Wang and Tseng, 2014). According to de Villiers and Blignaut (2016), there seems to be a seamless integration between learning and entertainment which lead to a unique user experience when serious games are introduced in the learning environment.

The literature reveals many definitions of serious games. Djaouti, Alvarez, et al. (2011) define serious games as a game created with a primary objective other than just entertainment. Zyda (2005) defines a serious game as “a mental contest, played with a computer in accordance with specific rules that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives.” Solving problems and elements of learning should be the focus of serious games while the fun part is also important, according to Susi, Johannesson, et al. (2007).

2.3.2 Board Games

Board games have been around for quite some time, dating back to 5000 BC and were probably “dice” games as a series of 49 small carved painted stones were found at the 5,000-year-old Başur Höyük burial mound in southeast Turkey (Attia, 2016). According to the Merriam Webster dictionary (Merriam-Webster) a board game is “a game of strategy (such as checkers, chess, or backgammon) played by moving pieces on a board.” Wikipedia (2018) define a board game as “a tabletop game that involves counters or pieces moved or placed on a pre-marked surface or “board”, according to a set of rules.” Tabletop games are games that are normally played on a table or other flat surface. To achieve the goal of the game some board games needs pure strategy while others contain an element of chance and some has no element of skill and is played on purely chance. Board games have an educational property with the possibility of providing players some type of knowledge through game play (Joseph and Diack, 2014).

2.4 Primary Objective

Analysis of the literature reveals that most studies that focus on serious games or board games in education targeted younger age groups (school children) (Papastergiou, 2009). These studies report largely positive attitudes towards serious games or board games in class, but students in tertiary education may have different profiles, requirements, prospects and perceptions. Assumptions that students in tertiary education will embrace the idea of serious games or board games in the computer science class are not well documented.

The primary objective of this study is to investigate the use of serious games and board games in the computer science class at HEI. To achieve the primary objective, the following secondary objectives and research questions were formulated.

2.4.1 Secondary Objectives

- Identify existing serious games and board games used for learning computer programming.
- Evaluate the suitability of the identified serious games and board games to assist students in learning computer programming.

2.4.2 Research Question

Referring to students in the computer programming class at HEI:

- Do students embrace serious games and board games?
- Do students prefer serious games over board games or vice versa?
- Are existing serious games and board games appropriate for learning computer programming?

3. CONCLUSION

The findings of this *study in progress* have the potential to be extended into a larger project where a variety of serious games and board games may be identified and/or developed to assist in learning computer programming.

ACKNOWLEDGEMENT

I would like to thank TELIT-SA for in part supporting this work. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and therefore TELIT-SA does not accept any liability in regard thereto.

REFERENCES

- Ali, A. and Smith, D., 2014. Teaching an introductory programming language in a general education course. *Journal of Information Technology Education: Innovations in Practice*, 13, 57-66.
- Annetta, L. A., et al., 2009. Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education*, 53,(1), 74-85.
- Attia, P., 2016. "The Full History of Board Games." <https://medium.com/swlh/the-full-history-of-board-games-5e622811ce89> Accessed on: 28 May 2017.
- Bartel, A. and Hagel, G., 2014, *Engaging students with a mobile game-based learning system in university education*. Global Engineering Education Conference (EDUCON), 2014 IEEE, IEEE.
- Bidarra, J., et al., 2015. Interactive design and gamification of eBooks for mobile and contextual learning. *International Journal of Interactive Mobile Technologies (iJIM)*, 24-32.
- Conneely, C., et al., 2012. Case Study Report for the NCCA.
- Coughlin, E. C. and Lemke, C., 1999. Professional competency continuum: Professional skills for the digital age classroom.
- Cunningham, M., 2015, *Factors impacting on adoption of Technology-enhanced Learning techniques by universities in Nairobi, Kenya*. Technology and Society (ISTAS), 2015 IEEE International Symposium on, IEEE.
- de Villiers, P. T. and Blignaut, A. S., 2016. Design Evaluation of a Forklift Serious Game.
- Djaouti, D., et al., 2011. Classifying serious games: the G/P/S model. *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches*, 2, 118-136.
- Gomes, A. J. and Mendes, A. J., 2010. A study on student performance in first year CS courses. *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education: 2010. ACM*, 113-117.
- Heersink, D. and Moskal, B. M., 2010. Measuring high school students' attitudes toward computing. *SIGCSE*, 10, 10-13.
- Heintz, S. and Law, E. L.-C., 2012, *Evaluating design elements for digital educational games on programming: A pilot study*. Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers, British Computer Society.
- Husain, L., 2011. Getting serious about math: serious game design framework & an example of educational game. 1-34.
- Jackson, J. and Moore, L., 2012. Engaging students in research: benefits, lessons learned and recommendations. *Proceedings of the 50th Annual Southeast Regional Conference: 2012. ACM*, 339-340.
- Joseph, S. and Diack, L., 2014. Playing interprofessional games: reflections on using the Interprofessional Education Game (iPEG). *Journal of Interprofessional Care, Informa UK Ltd*, 29,(3), 2.
- Kaiser, C. M. and Wisniewski, M. A., 2012. Enhancing student learning and engagement using student response systems. *Social Studies Research and Practice*, 7,(2), 137-149.
- Khaleel, F. L., et al., 2015, *The study of gamification application architecture for programming language course*. Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication, ACM.
- Kirlidog, M., et al., 2016. Unfulfilled need: reasons for insufficient ICT skills In South Africa. *Information Development*, 026666916671984.
- Krassmann, A. L., et al., 2015, *Evaluation of Game-Based Learning Approaches through Digital Serious Games in Computer Science Higher Education: A Systematic Mapping*. Computer Games and Digital Entertainment (SBGames), 2015 14th Brazilian Symposium on, IEEE.

- Kurkovsky, S., 2013. Mobile game development: improving student engagement and motivation in introductory computing courses. *Computer Science Education*, 23(2):138-157. *Computer Science Education*, 23,(2), 138-157.
- Mbogo, C., et al., 2014. Initial Evaluation of a Mobile Scaffolding Application That Seeks to Support Novice Learners of Programming. *International Association for Development of the Information Society*.
- Merriam-Webster, *Merriam Webster Dictionary*.
- Muratet, M., et al., 2011. Experimental Feedback on Prog & Play: A Serious Game for Programming Practice. *COMPUTER GRAPHICS forum*, 30,(1), 61-73.
- Nag, S., et al., 2013. Collaborative gaming and competition for CS-STEM education using SPHERES Zero Robotics. *Acta Astronautica*, 83, 145-174.
- Papastergiou, M., 2009. Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education*, 52,(1), 1-12.
- Peters, A. and Pears, A., 2012, *Students' experiences and attitudes towards learning computer science*. Frontiers in Education Conference (FIE), 2012, IEEE.
- Prensky, M., 2001. The games generations: How learners have changed. *Digital game-based learning*, 1-26.
- Rooney, P., 2012. A Theoretical Framework for Serious Game Design: Exploring Pedagogy, Play and Fidelity and their Implications for the Design Process. *International Journal of Game-Based Learning*, 2,(4), 41-60.
- Schwabe, G. and Göth, C., 2005. Mobile learning with a mobile game: design and motivational effects. *Journal of computer assisted learning*, 21,(3), 204-216.
- Shabanah, S. and Chen, J. X., 2009, *Simplifying algorithm learning using serious games*. Proceedings of the 14th Western Canadian Conference on Computing Education, ACM.
- Stanescu, I. A., et al., 2011. Interoperability in Serious Games. *eLearning & Software for Education*.
- Su, C. H. and Cheng, C. H., 2015. A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning*, 31,(3), 268-286.
- Susi, T., et al., 2007. *Serious games: An overview*, Institutionen für kommunikation och information.
- Tan, B.-K. and Rahaman, H., 2009, *Virtual heritage: Reality and criticism*. CAAD futures.
- Vogel, J. J., et al., 2006. Using virtual reality with and without gaming attributes for academic achievement. *Journal of Research on Technology in Education*, 39,(1), 105-118.
- Wang, T. and Tseng, Y., 2014. An empirical study: Develop and evaluation a mobile serious game on environmental education. In *Computer Science & Education (ICCSE), 2014 9th International Conference on*, 4.
- Wikipedia. (2018). "Board Games." Retrieved 18 March, 2018, from https://en.wikipedia.org/wiki/Board_game.
- Wrzesien, M. and Alcañiz Raya, M., 2010. Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the e-junior project. *Computers & Education*, 55,(1), 178-187.
- Zeeman, M. J., 2014. Modelling the factors that influence computer science students' attitude towards serious games in class. *Unpublished dissertation, NWU*.
- Zyda, M., 2005. From visual simulation to virtual reality to games. *Computer*, 38,(9), 25-32.

PROTOTYPING A PERSONAL LEARNING ENVIRONMENT

Timothy Arndt

Department of Information Systems, Cleveland State University, Cleveland, OH, USA

ABSTRACT

The rise of e-Learning as a primary platform for higher education promises to open up higher education to a wider range of learners than ever before. In order to best cater to this ever more diverse group of students, a personal learning system, which reflects the individual student's learning style and needs, would be valuable. Such a system would successfully integrate a user's learner profile, as well as his or her social networks, and big data sources, as well as time and location information in order to support ubiquitous learning. In this paper, we give an overview of the initial prototyping of a personal learning environment based on the PLAN formalism.

KEYWORDS

E-learning, Personalized Learning, Recommendation Systems

1. INTRODUCTION

In (Arndt and Guercio, 2016) we presented a formalism for our personal learning assistant called PLAN (Personal Learning AssistaNt). The formalism integrates Big Data (learning and social network) analytics, a finite state transducer which acts as a recommender system for the learner, the learner's calendars, location, a learner profile, etc. into a complete system for assisting the student to achieve his or her personalized goals. The aggressiveness of the recommendations can be controlled by the user. We laid out several scenarios of the system in action which help to explain how it will be used in practice. The cloud-based architecture of the system with a mobile app user interface was also described. The formalism is currently being implemented as a prototype system for experimentation purposes. The prototype being developed is described in this paper.

2. PLAN

In this section we will very briefly describe our personal learning assistant PLAN (Personal Learning AssistaNt) (Arndt and Guercio, 2016). A personal learning assistant uses multiple sources of information to provide appropriate recommendations and alerts when active. Some data are collected through the direct interaction with the user, while other data are mined by searching across available data sources. The first source of information used by the personal learning assistant comes from the user's learning profile. An additional source of information is provided by the user's social network from which desired data are extracted. Finally, additional data are mined from the multiple data sources available across the networks.

Mining significant data from data sources means to be able to identify data that can be of interest to the user. So, while an event can be considered interesting for achieving a learning goal, that event could be irrelevant if the user is unable to participate in it. This means that there are constraints that should be taken into account in order to select significant data and provide useful recommendations.

The two primary constraints come from the spatial and temporal information associated with the user. The temporal information is derived from a calendar. The spatial information is derived from the user's geolocation which, from the implementation point of view, is associated with the GPS (Global Positioning System) coordinates extracted from the sensor of the device where the personal assistant has been launched from. The amount of information produced by the personal learning assistant could be very large and

overwhelm the user. A level of aggressiveness must be used to customize the personal learning assistant to a user-desired level. An aggressiveness level of 0.0 will result in no recommendations being generated by the personal learning assistant, while 1.0 will result in a maximum number of recommendations being generated.

The activity of the personal learning assistant PLAN is formally described by a learning finite state transducer (LFST). The LFST moves from state to state in order to reach the desired learning goals and in each transition produces zero or more outputs.

When the system is initialized for a user, the PLAN generates the user's learner profile based on an interactive process with the user, as well as on the basis of a default profile. The learner profile may be refined as the learning process advances. The system also interrogates the user to determine the user's set of learning goals. The learner profile and learning goals are then used by the system to generate the LFST.

Data is mined from key-value data stores, as well as from the user's social networks, using traditional data mining processes. This results in knowledge items being discovered as data mining proceeds. Each time a knowledge item is discovered, the state transition function may result in a transition to a new state in the LFST. Calendar events from the user's calendar set may also be generated as time passes. These calendar events are treated as knowledge items by the LFST.

A state transition generally results in zero or more recommendations (i.e. the output function of the LFST) being made to the learner (e.g. to take a section of a particular source). On the other hand, some action of the learner may result in a state transition (e.g. the user successfully completing a course).

In addition, the data mining processes interact with the student's social networks – both online social networks such as Facebook and Twitter, as well as more informal social networks such as those identified by email and text message communication as well as those which are deduced by examining the course rosters of the courses the students are enrolled in. These social networks form another important asset in the student's learning process, being sources of expertise in course topics, various university processes, job markets and so on. The data mining processes can interact with online social networks through the APIs (Application Programming Interface) that they provide and through the more informal social networks through custom processes which may be developed.

3. INITIAL PROTOTYPE

Initial prototyping of the PLAN system has begun with the development of several components. These will be described in this section.

3.1 Data Cleanup/Transformation with OpenRefine for Cloud hosting on AWS

In our previous work, we noted that one of the difficulties of having a recommendation system controlled by the student and not by the institution is lack of access to institutional data. This component of our prototype handles this problem. Data (on class schedules in this case) is collected via screen scraping from a university website and then cleaning up the data using the OpenRefine (OpenRefine 2017) tool for data clean up and transformation. Figure 1 shows the OpenRefine interface.

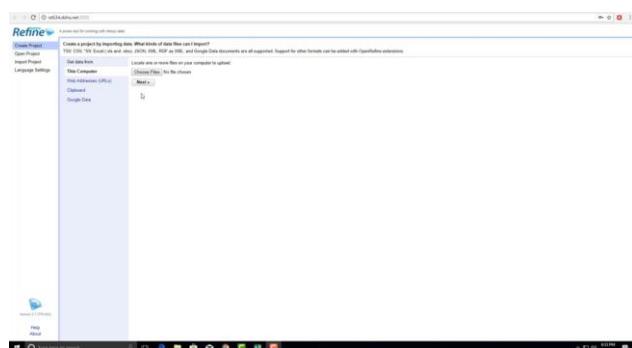


Figure 1. OpenRefine interface

In order to import this data into a relational database, much clean up needs to be done. For example, as seen in figure 2, there is misalignment of course data. This and other problems are cleaned up with OpenRefine, resulting in the data as shown in figure 3. Subsequently, the data was imported into a MySQL database hosted in the cloud on Amazon Web Services (AWS 2017) via Amazon RDS (Relational Database Service), as in the PLAN architecture. From there, it is queried to generate recommendations for the learner.

All	Enrl.	Date	ClassNr	Sect.	Secs	Begin Date	End Date	Time	Room	Instructor	Comp.	Column	Stat.	EnrlTot
1. ACT 301 Financial Accounting										Roman Enciso S	LEC	O	10140	
2.	1872	50	1	05/13/18 - 05/15/18	Thu	6:00 PM-8:50 PM								
3.														
4.	3249	501	1	05/13/18 - 05/15/18	TBA	TBA	WEB	STAT(CS)	LEC	O	17140			
5.														
6. ACT 321 Intermediate Accounting I														
7.	4957	50	1	05/13/18 - 05/15/18	Tue,Thu	6:00 PM-7:15 PM	BU 202	Parmentier Peter J	LEC	O	4110			
8.														
9.	4958	501	1	05/13/18 - 05/15/18	TBA	TBA	WEB	Parmentier Peter J	LEC	C	10110			
10.														
11. ACT 322 Intermediate Accounting II														
12.	5746	1	1	05/13/18 - 05/15/18	Tue,Thu	10:00 AM-11:15 AM	BU 101	Holsatuk Mark A	LEC	O	315			
13.														
14.	4960	2	1	05/13/18 - 05/15/18	Tue,Thu	2:00 PM-3:15 PM	BU 101	Molina Jo Richard	LEC	O	4110			
15.														
16.	4961	50	1	05/13/18 - 05/15/18	Tue,Thu	6:00 PM-7:15 PM	BU 101	Molina Jo Richard	LEC	C	10110			
17.														
18. ACT 323 Information Systems Audit														
19.	5215	501	1	05/13/18 - 05/15/18	TBA	TBA	WEB	Ridell Laura K	LEC	C	510			
20.														

Figure 2. Imported screen scrape data

1	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	Thu	6:00 PM-8:50 PM	BU 202	Roman Enciso S	LEC	O	10140				
2.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	TBA	TBA	WEB	STAT(CS)	LEC	O	17140				
3.															
4.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	Thu	6:00 PM-8:50 PM	BU 202	Roman Enciso S	LEC	O	10140				
5.															
6. ACT 321 Intermediate Accounting I															
7.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	Thu	6:00 PM-8:50 PM	BU 202	Roman Enciso S	LEC	O	10140				
8.															
9.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	TBA	TBA	WEB	Parmentier Peter J	LEC	C	10110				
10.															
11. ACT 322 Intermediate Accounting II															
12.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	Thu	6:00 PM-8:50 PM	BU 202	Roman Enciso S	LEC	O	10140				
13.															
14.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	TBA	TBA	WEB	Parmentier Peter J	LEC	C	10110				
15.															
16.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	Thu	6:00 PM-8:50 PM	BU 202	Roman Enciso S	LEC	O	10140				
17.															
18. ACT 323 Information Systems Audit															
19.	YCL 201 Intermediate Accounting I	501	1	05/13/18 - 05/15/18	TBA	TBA	WEB	Ridell Laura K	LEC	C	510				
20.															

Figure 3. The cleaned data

3.2 Recommendation interface through Facebook Messenger Bot

The data captured in the process described in the previous section, along with institution provided data which is not in need of as much cleanup is stored in a database system hosted in the cloud (MySQL relational database hosted on AWS cloud, in this case). An innovative interface for interrogating the user in order to generate recommendations has been prototyped using the Facebook Messenger Bot API. The initial message from the Bot is shown in figure 4. An interrogation gives the results shown in figure 5.

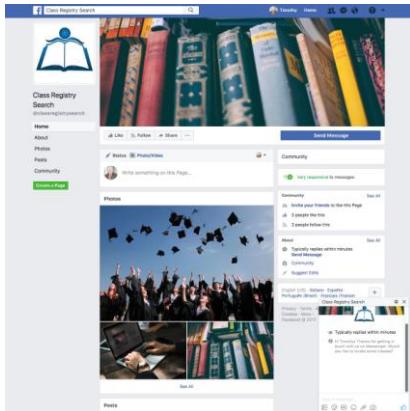


Figure 4. Messenger Bot UI



Figure 5. Messenger Bot class search results

The backend of the prototype is written in PHP which connects to the cloud-hosted MySQL database. Part of the PHP code is shown in figure 6.

```

40    session_start();
41    var_dump($session);
42
43    if ((int)$session['query_state'] === 0)
44    {
45        DatabaseSetState($sender, '');
46        return FacebookSendResponse($sender, "Please enter your field (abbreviated) of study.");
47    }
48
49    $schoolChoices = "Case Western Reserve University (CUW), Cleveland State University (CSU), Kent State University (KSU), University of Akron (UA)";
50
51    if ((int)$session['query_state'] === 1)
52    {
53        Studies = DatabaseGetDistinctStudies($conn);
54
55        if (in_array($thisInput, $Studies))
56        {
57            return FacebookSendResponse($sender, "I don't recognize that field. Try something else.");
58        }
59
60        DatabaseSetState($sender, '');
61        DatabaseSetState($sender, 'subject', str_replace(' ', '', $message));
62        return FacebookSendResponse($sender, "Please enter your preferred school. Choices: (" . $schoolChoices ")");
63
64    }
65
66    if ((int)$session['query_state'] === 2)
67    {
68        $schools = ["CUW", "CSU", "KSU", "UA"];
69
70        if (in_array($thisInput, $schools))
71        {
72            DatabaseSetState($sender, '');
73            return FacebookSendResponse($sender, "I don't recognize that school. Choices: (" . $schoolChoices ")");
74        }
75
76    }
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
559
560
561
562
563
564
565
566
567
568
569
569
570
571
572
573
574
575
576
577
578
579
579
580
581
582
583
584
585
586
587
587
588
589
589
590
591
592
593
594
595
596
597
597
598
599
599
600
601
602
603
604
605
606
607
607
608
609
609
610
611
612
613
614
615
616
617
617
618
619
619
620
621
622
623
624
625
626
627
627
628
629
629
630
631
632
633
634
635
635
636
637
637
638
639
639
640
641
642
643
643
644
645
645
646
647
647
648
649
649
650
651
652
653
654
655
656
657
658
659
659
660
661
662
663
664
665
666
667
667
668
669
669
670
671
672
673
674
675
676
677
677
678
679
679
680
681
682
683
684
685
685
686
687
687
688
689
689
690
691
692
693
694
695
695
696
697
697
698
699
699
700
701
702
703
704
705
706
707
708
709
709
710
711
712
712
713
714
714
715
716
716
717
718
718
719
719
720
721
722
723
723
724
725
725
726
726
727
728
728
729
729
730
731
732
732
733
733
734
734
735
735
736
736
737
737
738
738
739
739
740
740
741
741
742
742
743
743
744
744
745
745
746
746
747
747
748
748
749
749
750
750
751
751
752
752
753
753
754
754
755
755
756
756
757
757
758
758
759
759
760
760
761
761
762
762
763
763
764
764
765
765
766
766
767
767
768
768
769
769
770
770
771
771
772
772
773
773
774
774
775
775
776
776
777
777
778
778
779
779
780
780
781
781
782
782
783
783
784
784
785
785
786
786
787
787
788
788
789
789
790
790
791
791
792
792
793
793
794
794
795
795
796
796
797
797
798
798
799
799
800
800
801
801
802
802
803
803
804
804
805
805
806
806
807
807
808
808
809
809
810
810
811
811
812
812
813
813
814
814
815
815
816
816
817
817
818
818
819
819
820
820
821
821
822
822
823
823
824
824
825
825
826
826
827
827
828
828
829
829
830
830
831
831
832
832
833
833
834
834
835
835
836
836
837
837
838
838
839
839
840
840
841
841
842
842
843
843
844
844
845
845
846
846
847
847
848
848
849
849
850
850
851
851
852
852
853
853
854
854
855
855
856
856
857
857
858
858
859
859
860
860
861
861
862
862
863
863
864
864
865
865
866
866
867
867
868
868
869
869
870
870
871
871
872
872
873
873
874
874
875
875
876
876
877
877
878
878
879
879
880
880
881
881
882
882
883
883
884
884
885
885
886
886
887
887
888
888
889
889
890
890
891
891
892
892
893
893
894
894
895
895
896
896
897
897
898
898
899
899
900
900
901
901
902
902
903
903
904
904
905
905
906
906
907
907
908
908
909
909
910
910
911
911
912
912
913
913
914
914
915
915
916
916
917
917
918
918
919
919
920
920
921
921
922
922
923
923
924
924
925
925
926
926
927
927
928
928
929
929
930
930
931
931
932
932
933
933
934
934
935
935
936
936
937
937
938
938
939
939
940
940
941
941
942
942
943
943
944
944
945
945
946
946
947
947
948
948
949
949
950
950
951
951
952
952
953
953
954
954
955
955
956
956
957
957
958
958
959
959
960
960
961
961
962
962
963
963
964
964
965
965
966
966
967
967
968
968
969
969
970
970
971
971
972
972
973
973
974
974
975
975
976
976
977
977
978
978
979
979
980
980
981
981
982
982
983
983
984
984
985
985
986
986
987
987
988
988
989
989
990
990
991
991
992
992
993
993
994
994
995
995
996
996
997
997
998
998
999
999
1000
1000
1001
1001
1002
1002
1003
1003
1004
1004
1005
1005
1006
1006
1007
1007
1008
1008
1009
1009
1010
1010
1011
1011
1012
1012
1013
1013
1014
1014
1015
1015
1016
1016
1017
1017
1018
1018
1019
1019
1020
1020
1021
1021
1022
1022
1023
1023
1024
1024
1025
1025
1026
1026
1027
1027
1028
1028
1029
1029
1030
1030
1031
1031
1032
1032
1033
1033
1034
1034
1035
1035
1036
1036
1037
1037
1038
1038
1039
1039
1040
1040
1041
1041
1042
1042
1043
1043
1044
1044
1045
1045
1046
1046
1047
1047
1048
1048
1049
1049
1050
1050
1051
1051
1052
1052
1053
1053
105
```

3.3 NoSQL Databases

The PLAN architecture foresees data mining of data from NoSQL databases in order to provide best-in-class recommendations to learners. Spatial location data is also to be used to enable a ubiquitous learning experience. In the prototype system, NoSQL has been incorporated in the form of MongoDB (MongoDB 2017), whose geospatial query facilities are used to provide a facility for students to find the best parking location on their campus. This part of the prototype is shown in figure 7.

```

    "Health Sciences"
    ],
    "Availability" : 200
  }
  db.operations.find("Building": "Rhodes West").pretty();
  {
    "_id" : ObjectId("5a1b8891329e639d2ea6e9db"),
    "Garage" : "Central Garage",
    "Capacity" : 750,
    "Permit" : "Green Permit",
    "Visitors" :
  },
  "Building" : [
    "Music & Communication",
    "Rhodes West"
  ],
  "Availability" : 250
}
{
  "_id" : ObjectId("5a1b8891329e639d2ea6e9da"),
  "Garage" : "Lot 22 / Student Center",
  "Capacity" : 200,
  "Permit" : "Green Permit",
  "Visitors" :
},
  "Building" : [
    "Rhodes West",
    "Rhodes Tower",
    "Main Classroom"
  ],
  "Availability" : 5
}
{
  "_id" : ObjectId("5a1b8891329e639d2ea6e9db"),
  "Garage" : "Main Classroom Building",
  "Capacity" : 600,
  "Permit" : [
    "Green Permit",
    "Visitors"
  ],
  "Building" : [
    "Main Classroom",
    "Chester Building",
    "Chester Building Annex"
  ],
  "Availability" : 100
}

```

Figure 7. UI and MongoDB collections

4. DISCUSSION AND FUTURE WORK

The prototype components implemented demonstrate many of the facets of the PLAN architecture previously described: NoSQL database; cloud-based; incorporation of learners' social networks; location information; structured and semi-structured data. The components are currently stand-alone. In the future, we will integrate them into a single system. To do this, we need to integrate a system which implements the LFST and serves the high-level controller for the recommendation system. Other components need to be implemented as well – such as a user interrogation scheme which can make use of the same Messenger Bot technology seen in section 3. We will then proceed to experiment with the system using volunteer learners to determine its efficiency.

ACKNOWLEDGEMENT

The help of students in the Master's of Information Systems course IST 634 Enterprise Databases is gratefully acknowledged.

REFERENCES

- Arndt T. and Guercio A., 2016. A Formalism for PLAN – a Big Data Personal Learning Assistant for University Students. *In Journal of e-Learning and Knowledge Society*, Vol. 12, No. 2.
- <http://openrefine.org>. Accessed December 8, 2017.
- <https://aws.amazon.com>. Accessed December 8, 2017.
- <https://developers.facebook.com/docs/messenger-platform>. Accessed December 8, 2017.
- <https://www.mongodb.com>. Accessed December 8, 2017.

"NOBODY KNOWS WHAT THE BODY CAN DO": A COMPLEX APPROACH OF AUTISTIC SPECTRUM DISORDERS

Nize Maria Campos Pellanda and Maria de Fátima de Lima das Chagas
University of Santa Cruz do Sul – UNISC, Brazil

ABSTRACT

This paper attempts to describe a complex theoretical framework to support a research that addresses the emergence of cognition / subjectification as technological coupling, in subjects diagnosed with ASD. Our concept-axis coined by our research group - Ontoepistemogenesis- deals exactly with the inseparability knowing / subjectivation. With this, these researchers are betting on the self-organization and neuroplasticity, fundamental principles of the Paradigm of complexity and breaking with the behavioristic, hegemonic and simplifying approaches of reality by separating cognition of subjectivation causing suffering to these subjects.

KEYWORDS

Self-organization, Ontoepistemogenesis, Neuroplasticity, Complexity, Cognition-Subjectivity

1. INTRODUCTION

"We did not bathe twice in the same river" Heraclitus (frag.101)

To address Autism Spectrum Disorders (ASD), we made a theoretical background option: complexity. This refers to a paradigm that comes to break with a tradition of 2,500 years: the metaphysical inheritance, which emerged and consolidated in classical Greece and became hegemonic from modern science in the seventeenth century.

What does metaphysical inheritance mean? Why complexity?

After a long philosophical, scientific, ethical and cultural tradition based on stability, linearity, and a demarcated separation between knowledge and wisdom, a new way of thinking, living and doing science emerges from the nineteenth century, but the former never completely disappeared. It is the paradigm of complexity whose organizing epistemic axis is the junction of what was separated (Morin, 1991).

With this theoretical background, we decided to make a complex approach to TEA (Autism Spectrum Disorders) to understand the emergence of cognition / subjectivation in the subjects diagnosed with this syndrome. So far, from a vast literature review, what we have seen about TEA approaches is the presence of linear, reductionist, fragmentary approaches to different human dimensions and simplifying reality. They are hegemonic in the approach to autism.

The implications of these approaches are disastrous in that they do not address the biological conditions of humans, thus bringing much suffering to the autistic subjects. Furthermore, these researches and treatments are not considering seminal discoveries of neurosciences such as the principle of self-organization that has an important unfolding in the dynamic principle of neuroplasticity and which had its genesis in the cybernetic movement of the 40s / 50s of the twentieth century. These are approaches that still bear the mark of the modern paradigm, fragmentation. By focusing directly on the objective behavior of children as communication and stereotypical they fragment the reality because they leave aside the work with self-consciousness that would lead to autonomy. This is fundamental in autopoietic beings that are living beings according to the theory of Biology of Cognition by Humberto Maturana and Francisco Varela. Autopoiesis is the central concept of this theory that emphasizes that living beings create themselves by living (Maturana; Varela, 1980).

Brain science has changed a lot in recent times thanks to a complex approach. As mentioned, the approaches to Autism Spectrum Disorder - ASD have not followed this research in practice. For a long time, neuroscientists considered that the frontal region of the brain, the frontal cortex, was responsible for the "higher" faculties of human beings such as reasoning, abstract thinking, decision making, imagining the thinking of others, etc. And here are the approaches forgetting that the functioning of the nervous system is very sophisticated and distributed throughout the body. There is the subcortical brain that lies beneath the thin layer of the cortex. This part of the brain, being more hidden, was not valued properly in its organic functions.

Five years ago, we began to study what was done in terms of research and treatment of ASD and we were perplexed by what we found: repetition attitudes, management treaties, reinforcements and other mechanical actions that, in our complex eyes, only reinforced the stereotypes characteristic of these subjects denying to the neuronal work the chance of connections and reconfigurations. The claim of the professionals who thus act is that they cannot "disturb" these subjects because they are already in a situation of great anxiety and difficulty of breaking of routines. They end up convincing families to do the same. Our approach is in complete different ways.

Our services did not have in their beginnings, nor we have today the therapeutic objective because we are not therapists, but researchers of epistemology, interested in understanding how the process of cognition / subjectivation emerges in an inseparable way in the flow of living. It was in fact, from this complex theorem that we developed in our research group the operative concept that we coined of Ontoepistemogenesis.

That is where we have been in recent years what we consider a complexification of the group over a period of 18 years of research in the perspective of complexity. What marks our research from an epistemological-ontological turn is the bet on central assumptions of neuroscience in terms of affirming the self-organizing capacity of living beings. In this sense, we support the neoplastic attitude of neuroscientist Oliver Sacks:

In this perspective, deficiencies, disorders, and diseases can play a paradoxical role, revealing latent powers, developments, evolutions, life forms that might never have been seen, or even imagined, in the absence of these evils. (Sacks, 2008, p. 13)

What follows in these writings are justifications for our epistemological-ontological options in terms of complexity as well as the summary narrative of an empirical work with autistic children in a methodology that does not deal with representations of an external, objective, pre-dated, independent world of the knowing subject but with an internal work of the system that configures its reality when operating.

2. INTEGRATING DIFFERENT DIMENSIONS OF HUMAN

The opposition between culture and technique, between man and machine is false and without foundation; it denotes only ignorance or resentment; hides in an easy humanism some reality rich in human efforts and natural forces and constituting the world of Technical Objects, mediators between culture and man. (Simondon 1958).

In the complex perspective in which we are working, nothing can be left out. In this sense, the research we develop with the autistic children includes the coupling with a technical object - the iPad - with the objective of creating a coupling environment in which the subjects are challenged (noises) to respond with important reconfigurations that congregate cognitive aspects, affective and social interaction. It is, therefore, the technical dimension of human life. From what we know of neuroplasticity, we believe that there may be in these cases neurophysiological activities that have the potential to leverage important neuronal work to be expressed by the emergence of synapses. It seems that the creation of an environment rich in human and technological interactions can lead people with damaged brain to modifications with plastic alterations that give rise to learning (ROTTA, 2006). It has been observed that autistic children have special predilection for technical devices. This is very important evidence that has to be considered. Therefore, we have decided to offer them a technology-rich environment that could lead to meaningful learning. To that end, we offer the iPad to them with challenging games from the complex epistemological point of view in which we can find epistemic-ontological assumptions such as self-overcoming, narrative potential, and noises (disturbances).

According to the background theories that support us, these assumptions can lead to the reconfiguration of the subjects in terms of self-affirmation / self-constitution (autopoiesis), the possibility of thinking about oneself (the autistic person has difficulty in reflective thinking and saying "I", or "myself") and, finally, the situations of provocative noises that can trigger important reconfigurations in the said subjects.

The technology-rich environment, therefore, can mobilize the neuroplasticity condition of these subjects, which is very different from the mechanical, adaptive, representational actions with external, repetitive references and reinforcements that, in our evaluation, for not mobilizing the brain, end up fixing the stereotypes characteristic of these disorders.

But why does technology carry all this epistemic / ontological potential? To answer this question, we again clarify that we do not work on the issue of learning as an adaptation of the organism to an environment but as a dynamic coupling in which emotions are involved and in which technology enters creatively (BERGSON 1979, SIMONDON, 1958). In this form of dynamic and genetic relation (in the sense that things are generated in the present action), the technique has the function of equipping the human beings of resources such that they potentiate them.

Our team of researchers, supported by Simondon (1958), believes that the approach to learning understood as adaptation leads to stable systems which, in our view, would impede cognitive flexibility. For Simondon, therefore, this would be an evolution without vitality. In this sense, a coupled system would be full of vital potentials. That is why this philosopher thinks of technology as a process of individuation with ontogenetic power helping the subject to think for himself. Thus, the difficulties of social communication, speech and symbolization can be circumvented with a significant technological environment. Simondon defends the principle of individuation (singularization) process in which human beings relate to technical objects in a genetic becoming such that they are constituted in these relations with the materiality of technical objects together with the functioning of them. Simondon, therefore, does not believe in a substantial subject, already formed, "already individuated", but in constant constitution and, for this, depends on technical objects. He refuses, therefore, the denial of becoming, alerting to the need for an ontogenesis that is inseparable from technique, in the flow of human life (Simondon, 1958).

Our approach to technique, therefore, neither simplifies nor separates it from the constituent becoming of the human. In this sense, we use a technical object, the iPad, as well as opportunities for living with technological environments, such as the visit to the museum, to offer these children with difficulties of coupling with reality possibilities of becoming full of virtualities (power). When an autistic child plays on the computer, it is not simple information what it finds, but it experiences an environment full of processes of symbolization in which it is subjecting what outside the virtual world is extremely difficult for it.

2.1 A Brief Empirical Report of the Research

The subjects of this research described and currently under development are 4: a 11-year-old girl (L), an 8-year-old girl (J), one 12-year-old boy (M) and a 4-year-old boy (P). They participate in the project with very different entries in time.

These subjects have been attended once a week by a team member in a specialized room of the University, separated with a unilateral mirror of another room where the other researchers and scholars of the project observe the empirical process. In the calls, the children work with the iPad from the applications that we make available to them. The researcher in contact with the child makes few interventions but tries to be affective and attentive. Whenever necessary, it acts as a challenge. The selected applications involve epistemological assumptions that are in accordance with the needs of the project such as the opportunity to make relationships between very different elements, flexibility to break established patterns in game architecture, subjectivation strategies, etc. Our main requirement regarding the type of application is that they have the function of breakdown, that is, disrupters challenging the subjects in the direction of a

2.2 The Project Unfolds in Four Stages

1. Children are attended at the university in 12 weekly sessions
2. The children spend two months at home and receive an iPad to use during this period under the observation of families

3. Subjects return to university for a period of 12 more sessions. In this return, the first activity is the parents' report

4. Treatment of research emergencies with a view to answering the central problem and writing the final report. This study material includes the subjective production of the researchers giving account of their own affections.

Researchers make a logbook where they are included as part of the reality observed showing how they are being affected by the research and reporting the process of technological engagement of children with the iPad. At the end of each meeting the team meets to discuss the session by crossing data from the observations with the perceptions of each researcher. Sessions are recorded on video.

The emergencies of each session are analyzed in the light of theoretical assumptions. In order to do this, we do not use categories that embody a reality in becoming, leaving out subtleties of singularization, but rather markers derived from the theoretical presuppositions we have used and which are intended to account for a complex reality in constitution. It is not an abstract reality, but the living form in which each subject singles out his life. The markers used are: autopoietic process, structural coupling and complexification process. These markers serve to signal transformations in the subjects in the becoming of the lived processes.

3. CONCLUSION

As stated at the outset, the purpose of this article is not to describe empirical research with all its ramifications but rather to present the complex theoretical assumptions to justify a new approach to autism. Therefore, we have done here only to make a clipping of empirical emergencies of some subjects with the purpose of illustrating what we theoretically defend. The attitude of challenging the child to mobilize in all its dimensions and the creation of autopoietic instruments of self-affirmation and narrativity (I can, I have done, this or that is me, narrative attitude) are central strategies in this methodological process.

The experience of this research has shown us, in a more profound and sensitive way, the most significant demands of these subjects. Therefore, we are developing the prototype of a digital platform for the use of these people with functionalities that respond to even more subjective demands of self-awareness, social interaction and others.

REFERENCES

- Bergson, H. (1979). *The creative evolution*, Zahar. Rio de Janeiro.
Maturana, H. Varela, F. (1980). *Autopoiesis and cognition*. D. Reidel, Dordrecht
Sacks, O (2008). *An anthropologist on Mars*. Companhia das Letras, São Paulo.
Simondon, G. (1958). *Du mode d'existence des objets techniques*, Aubier, Alençon.
Google translator for companies: Google Toolkit de tradition for apps Translator of sites

EXTENDING THE UTAUT TO MEASURE THE ADOPTION OF WEB-BASED KNOWLEDGE SHARING SYSTEMS IN SAUDI UNIVERSITIES

Nouf Almujally¹ and Mike Joy²

¹*University of Warwick & Princess Nourah bint Abdulrahman University*

Department of Computer Science, Coventry CV4 7AL, UK

²*University of Warwick, Department of Computer Science, Coventry CV4 7AL, UK*

ABSTRACT

Knowledge management has emerged as an important practice for achieving competitive advantage among global corporations, including higher education institutions. Knowledge sharing is a significant source of success in knowledge management, however, in Saudi Arabian universities, knowledge management is often lacking when it comes to web-based knowledge sharing amongst academics. To ensure that a knowledge sharing system can be well implemented and used when communicating internally in an academic context, there is a need to know why academics accept or reject the use of web-based knowledge sharing systems. Therefore, the aim of this paper is to investigate the important factors that would influence academics' behaviour towards accepting and adopting the use of web-based knowledge sharing systems in Saudi Arabian universities. A conceptual model is proposed based on the Unified Theory of Acceptance and Use of Technology (UTAUT) and Task-technology Fit (TTF) models, as well as other factors which are explored in knowledge sharing literature reviews to enrich the proposed model. Then, the model will be edited and refined using a mixed method approach. Future work will expand the model and evaluate it to ensure that it fits the academics' needs.

KEYWORDS

Knowledge Management, Knowledge Sharing, Model, Technology Acceptance, Technology Adoption, Academics, Higher Education, UTAUT

1. INTRODUCTION

In today's dynamic and competitive era, a knowledge management system is becoming an important asset for organisations. Therefore, in order to be successful and gain competitive advantage, organisations, including higher education institutions, depend heavily on knowledge and the implementation of a knowledge management system which has become a crucial success factor (Fullwood, Rowley and Delbridge, 2013). A review of the literature reveals that knowledge sharing is one of the essential roles of higher education institutions where knowledge is created through research, disseminated through publication, shared via teaching, or presenting at a seminar (Kim and Ju, 2008; Jolae et al., 2014). However, knowledge sharing amongst academic staff is often inadequate in Saudi Arabian universities. Consequently, tacit knowledge, which is preserved in the mind of expert academics, would be more difficult to share and express in tangible form, as it involves personal intellectual skills and problem-solving capabilities, and is gained through teaching experience (Kim and Ju, 2008). Uncodified tacit knowledge can affect academics' performance and may result in lower levels of achievement.

The use of web technology plays an important role as a knowledge enabling tool which can support the knowledge sharing process, encourage academics' collaboration, and enhancing the internal knowledge management environment (Panahi, Watson and Partridge, 2013; Usman, Ishaq and Oyefolahan, 2014). Sharing knowledge via web technology can be defined as the process of transferring different types of knowledge to others using information and communication technology (ICT) tools such as emails, blogs, video conferencing, or internet applications (Merrill and Ravi, Vijayan; 2001). Using web-based knowledge sharing can help academics stay internally connected, and builds relationships by establishing a knowledge community (Panahi, Watson and Partridge, 2012).

This paper is structured as follows. Section 2 describes the research problem, while section 3 provides additional background on knowledge sharing and technology adoption, as well as the theoretical bases of the study. Following this, the conceptual model is constructed to understand web-based knowledge sharing adoption. Finally, the methods which will be used for validating the model are explained.

2. PROBLEM STATEMENT

This research considers the application of knowledge sharing in Saudi Arabian universities. Most Saudi universities are not geographically co-located, therefore, experts need to travel between campuses to share knowledge with others who have common interests. Additionally, many years of teaching experience could be lost due to the retirement of academics without their expertise being recorded in a proper knowledge sharing system. Consequently, face-to-face communication is no longer an effective way of sharing knowledge. Thus, universities need to implement a knowledge sharing system that could facilitate the flow of knowledge amongst academics and could overcome the challenges mentioned above. In order to ensure that a knowledge sharing system can be well implemented and used in the academic context, there is a need to investigate technology acceptance among academics (Al-Gahtania, Hubona and Wang, 2007). Therefore, a conceptual model aimed at addressing the factors that affect an academic's behaviour toward accepting the use of a web-based knowledge system is developed based on the literature review.

This research seeks to add value to its research area, as there is a clear demand to improve knowledge sharing activities and practices in Saudi Arabian universities, by answering the following research question: what are the factors that influence the adoption of a web-based knowledge sharing system among academic staff in in Saudi Arabian universities?

3. BACKGROUND

3.1 Knowledge Sharing

Jolaee et al. (2014) suggested that knowledge sharing is crucial activity in knowledge-based organizations, like universities, since the majority of the academics are knowledge workers who engage in teaching and research activities. Their view is supported by Seonghee and Boryung (2008) who stated that faculty members share the same common missions where academics work as the disseminators of knowledge through teaching and as the producers of knowledge through conducting research. Fullwood, Rowley and Delbridge (2013) found that knowledge sharing depends on individual behaviour toward accepting or rejecting the use of web technology for disseminating knowledge. Behavioural intention is an indicator of an individual's willingness to share knowledge and can be affected by various factors (Alammari and Chandran, 2017). In this research, behavioural intention refers to the academic's likelihood to share knowledge using web technology.

3.2 Theoretical Background

The literature on IT adoption supports a large number of theoretical models (Goodhue and Thompson, 1995; Venkatesh and Brown, 2013) that have been applied to different contexts. However, a review of the literature reveals that there is no well-defined adoption theory in terms of web-based knowledge sharing adoption. Therefore, to investigate the factors that could affect academics in terms of the degree of their acceptance of the use of web technology for knowledge sharing, theories that place emphasis on the technology acceptance area should be taken into consideration.

3.2.1 The Unified Theory of Acceptance and Use of Technology (UTAUT)

It was found that the Unified Theory of Acceptance and Use of Technology (UTAUT) has been widely used to determine behavioural intention and usage (Venkatesh et al., 2003). The aim of UTAUT is to explain a user's intention to adopt and use an information system. Since the UTAUT model has resulted from testing and combining different technology adoption models, it is considered as a bench-mark for constructing the web-based knowledge sharing adoption model.

3.2.2 Task-Technology Fit (TTF)

The Task-Technology Fit (TTF) model measures the capabilities of technology to support the function of a task and meet the user's requirements with the available technology functionality. Individuals are more likely to adopt a new technology if the functions of that technology meet the users' requirements (Goodhue and Thompson, 1995).

As previous studies have determined a strong relationship between what the technology can do (task-technology fit) and what the benefits of using it are in terms of enhancing task performance (performance expectancy) (El Said, 2015), this research adopts both models to explore academics' attitudes towards accepting the use of technology for knowledge sharing.

The next section outlines the significant factors and combines them into a web-based knowledge sharing model.

4. CONCEPTUAL MODEL

Based on the literature review, the proposed model, which can be seen in figure 1, is constructed to explore academics' behaviours towards sharing knowledge via web technology. The presented model intends to blend factors from the Unified Theory of Acceptance and Use of Technology (UTAUT) model and the Task-Technology Fit (TTF) model, as well as combining external factors that have been explored from knowledge sharing literature reviews (Paroutis and Al Saleh, 2009; Usman and Oyefolahan, 2014).

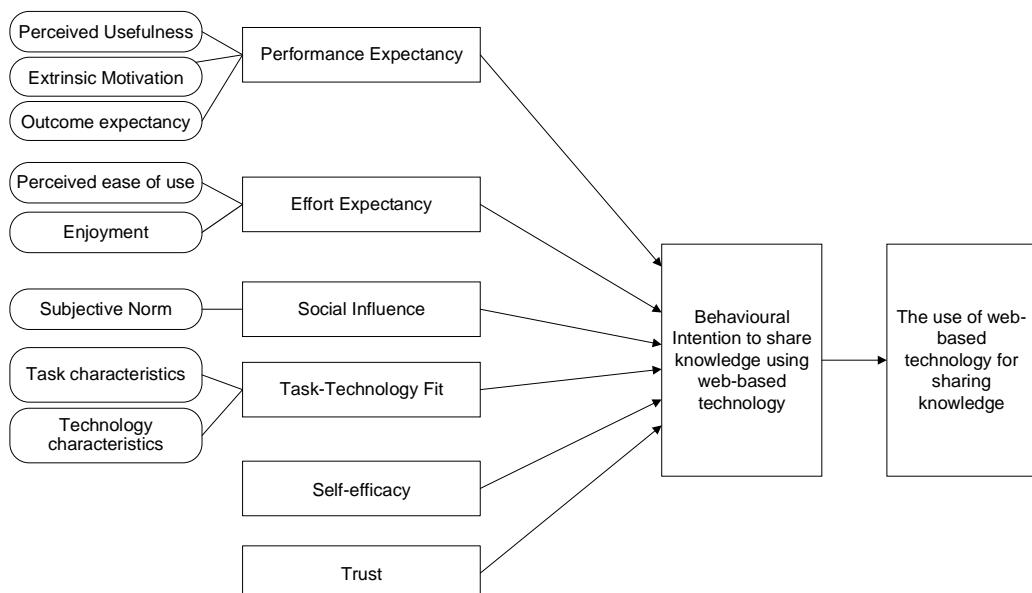


Figure 1. Web-based knowledge sharing adoption model

4.1 Performance Expectancy

Performance expectancy is defined as the degree to which an individual believes that using the system will help them to improve job performance. Three measurements were used to measure this construct, namely outcome expectancy, extrinsic motivations, and perceived usefulness. Outcome expectancy is strongly correlated with attitude toward using web technology for sharing knowledge. According to the Saudi study conducted by (Al-Gahtania, Hubona and Wang, 2007), outcome expectancy has a positive influence on behavioural intention to use new technologies. When employees are unaware of the potential outcomes, it is

more likely that they will perceive the cost of using new tools to be higher than the benefits which consequently could inhibit them from sharing knowledge using web tools.

Previous research shows that one of the most common barriers to sharing knowledge via web technology is the absence of extrinsic motivations. Paroutis and Al Saleh (2009) noted that some employees avoid using Web 2.0 technologies such as blogs and wikis as online knowledge sharing tools due to the absence of recognition programs.

Perceived usefulness has a significant impact on adopting the use of a web-based knowledge sharing system. If users perceive that using the online systems can improve their job performance, they become more motivated to use the systems for knowledge sharing purposes (Esmaeilzadeh et al., 2013).

4.2 Effort Expectancy

Effort expectancy is defined as the degree of ease associated with use of the system. Two constructs capture the concept of effort expectancy. Perceived ease of use can be defined as the person's belief that using a web technology for knowledge sharing is free of effort (Hsu and Lin, 2008). Many faculty members, especially older professors, might face difficulties when using a knowledge sharing system such as an e-learning system. Therefore, designing an ease of use knowledge sharing system which meets users' requirements is likely to enhance the behaviour of knowledge sharing (Riege, 2005; Kim and Lee, 2006; Alhalhouli, Hassan and Der, 2014).

Enjoyment in helping others appears to be a key factor that could affect knowledge sharing using web technology. It is defined as the degree of personal perception that the system is fun and enjoyable (Hsu and Lin, 2008). It has been found that knowledge workers who feel pleasure when helping others are likely to be more motivated to participate in knowledge sharing activities (Alias, Abbas and Nordin, 2016).

4.3 Social Influence

Social influence is defined as the degree to which an individual perceives that it is important that others believe they should use the new system. Social influence is a direct determinant of behavioural intention and is represented as subjective norm, and is considered as one of the factors that might affect the intention to accept using web tools for sharing knowledge amongst academics. This has been defined by Jolaee et al. (2014) as the degree of the person's perceptions regarding social pressure from their managers and colleagues which could impact the adoption of a new technology for knowledge sharing. Alammari and Chandran (2017) found that subjective norms, such as family and friends, have a positive effect on an individual's attitude due to the high individualism-collectivistic characteristic of Saudi Arabian culture.

4.4 Task-technology Fit

Task-technology fit measures the capabilities of a technology to support the function of a task to meet user's requirements with the available technology's functionality. Individuals will adopt a new technology if the functions of that technology meet their requirements (Goodhue and Thompson, 1995). Task-technology fit suggests that the appropriate fit between knowledge sharing systems' features and task requirements is a key determinant of accepting the use of technology for sharing knowledge (El, 2015).

4.5 Self-efficacy

Self-efficacy could also affect knowledge sharing using web technology. Previous studies suggested that individuals with high self-efficacy are more likely to have positive attitudes toward using web tools for knowledge sharing (Fullwood, Rowley and Delbridge, 2013; Skaik and Othman, 2015). According to Mahmoud et al. (2014), some academics use technology to share their knowledge because of their belief in their abilities and skills, and the belief that using technology can lead to work improvements.

4.6 Trust

Another factor that has been discussed in the context of knowledge sharing amongst academics is trust, which includes trusting the quality of knowledge being shared and trusting that others will also share their valuable knowledge. In an online setting, Hassandoust and Perumal (2011) noted that a high level of interpersonal trust among team members has a positive effect on the adoption of e-collaboration tools for knowledge sharing.

5. RESEARCH METHODS

To reach the objectives of the research, semi-structured face-to-face interviews, which include open and closed questions, will be conducted with 18 experienced or novice Saudi academics to explore other unidentified factors, derived from academics' perspectives, which influence academics towards sharing knowledge using web tools in Saudi Arabian higher education institutions. Then, a questionnaire will be designed to confirm all the components (from previous studies and academics' perspectives) of the designed model. Finally, in order to investigate the relationships between the model's components and validate the research hypothesis, the questionnaire will be conducted with Saudi academics to evaluate the resulting model.

6. DISCUSSION AND CONCLUSION

The research objective is to enhance web-based knowledge sharing practices among academics in Saudi universities. Therefore, this study will explore the factors that affect academics' behaviours toward using web technology to share knowledge in Saudi Arabian higher education institutions. The conceptual model has been constructed by combining factors investigated through reviewing the existing knowledge sharing literature, as well as factors from the UTAUT and TTF models.

From the author's perspective, to ensure a successful implementation of knowledge management systems in Saudi universities, all the model's factors that will be explored in the empirical study should be taken into consideration. Future work will expand the model and evaluate it to ensure that it fits the academics' needs.

REFERENCES

- Al-Gahtania, S. S., Hubona, G. S. and Wang, J. (2007). "Information technology (IT) in Saudi Arabia: Culture and the acceptance and use of IT", *Information & Management*, 44(8), pp. 681–691.
- Alammari, A. and Chandran, D. (2017). "Impact of Knowledge Sharing Adoption on Universities' Virtual Learning Communities", pp. 4485–4494.
- Alhalhouli, Z. T., Hassan, Z. B. and Der, C. S. (2014). "Factors affecting knowledge sharing behavior among stakeholders in Jordanian hospitals using social networks", *International Journal of Computer and Information Technology*, 3(5), pp. 919–928.
- Alias, D. F., Abbas, W. F. and Nordin, A. (2016). "Knowledge Sharing Scenario Capture: A Case Study in IT Department of Public Higher Education Institution", *Journal of Advanced Management Science*, 4(6), pp. 507–510.
- El Said, G. R. (2015). "Understanding Knowledge Management System antecedents of performance impact : Extending the Task-technology Fit Model with intention to share knowledge construct", *Future Business Journal*. Elsevier, 1 (1–2), pp. 75–87. doi: 10.1016/j.fbj.2015.11.003.
- Esmailzadeh, P., Sambasivan, M., Kumar, N. and Nezakati, H. (2013). "The effect of knowledge sharing on technology acceptance among physicians", *Global Advanced Research Journal of Engineering, Technology and Innovation*, 2(2), pp. 48–57.
- Fullwood, R., Rowley, J. and Delbridge, R. (2013). "Knowledge sharing amongst academics in UK universities", *Journal of Knowledge Management*, 17(1), pp. 123–136. doi: <http://dx.doi.org.ezproxy.apollolibrary.com/10.1108/13673271311300831>.

- Goodhue, D. L. and Thompson, R. L. (1995). "Task-Technology Fit and Individual Performance", *Mis Quarterly*, 19(2), pp. 213–236. doi: 10.2307/249689.
- Hassandoust, F. and Perumal, V. (2011). "Online Knowledge Sharing in Institutes of Higher Learning: A Malaysian Perspective", *Journal of Knowledge Management Practice*, 12(1).
- Hsu, C. L. and Lin, J. C. C. (2008). "Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation", *Information and Management*, 45(1), pp. 65–74. doi: 10.1016/j.im.2007.11.001.
- Jolaee, A., Md Nor, K., Khani, N., Md Yusoff, R., Nor, K. M., Khani, N. and Yusoff, R. M. (2014). "Factors affecting knowledge sharing intention among academic staff", *International Journal of Educational Management*, 28(August 2016), pp. 413–431. doi: 10.1108/IJEM-03-2013-0041.
- Kim, S. and Ju, B. (2008). "An analysis of faculty perceptions: Attitudes toward knowledge sharing and collaboration in an academic institution", *Library & Information Science Research*, 30(4), pp. 282–290.
- Kim, S. and Lee, H. (2006). "The impact of organizational context and information technology on employee knowledge-sharing capabilities", *Public Administration Review*, 66(3), pp. 370–385.
- Mahmoud, M., Rasli, A. M., bin Othman, M. F. and Abdulahad, B. M. (2014). "The effect of organizational culture on knowledge sharing among academic staff holding an administrative position in university", *Journal Of Management Info*, 3(1), pp. 67–83.
- Merrill, A., Ravi, B., Vijayan, S., (2001). "E-knowledge networks for inter-organizational collaborative e-business". *Logistics Information Management* 14, 1/2: 149-163.
- Panahi, S., Watson, J. and Partridge, H. (2012). "Social media and tacit knowledge sharing: developing a conceptual model", *World academy of science, engineering and technology*, (64), pp. 1095–1102.
- Panahi, S., Watson, J. and Partridge, H. (2013). "Towards tacit knowledge sharing over social web tools", *Journal of Knowledge Management*, 17(3), pp. 379–397.
- Paroutis, S. and Al Saleh, A. (2009). "Determinants of knowledge sharing using Web 2.0 technologies", *Journal of Knowledge Management*, 13(4), pp. 52–63.
- Riege, A. (2005). "Three-dozen knowledge-sharing barriers managers must consider", *Journal of Knowledge Management*, 9(3), pp. 18–35.
- Usman, S. H., Ishaq and Oyefolahan, O. (2014). "Determinants of Knowledge Sharing Using Web Technologies among Students in Higher Education", *Journal of knowledge management, Economics and information technology*, IV(2), pp. 1–22.
- Usman, S. H. and Oyefolahan, I. O. (2014). "Encouraging Knowledge Sharing Using Web 2.0 Technologies in Higher Education: a Survey", *International Journal of Managing Information Technology (IJMIT)*, 6(2), pp. 19–28. doi: 10.5121/ijmit.2014.6202.
- Venkatesh, V. and Brown, S. A. (2013). 'Research article ridging the qualitative – quantitative divide : guidelines for conducting mixedm ethods", 37(1), pp. 21–54.
- Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. (2003). "User acceptance of information technology: Toward a unified view", *MIS Quarterly*, 27(3), pp. 425–478. doi: 10.2307/30036540.

ONLINE DATA ABOUT BRAZILIAN JUDICIARY: COMPARATIVE ANALYSIS OF JUSTICE IN NUMBERS 2004-2017

Lahis Pasquali Kurtz¹, Amanda Nunes Lopes Espiñeira Lemos² and Alexandra Krastins Lopes Souto Maior³

¹*UFMG, Avenida João Pinheiro, Nº 100, Bairro Centro, 11º Andar, Edifício Villas-Boas Belo Horizonte - Minas Gerais, Brazil, Zip Code 30130-180*

²*UNB, Campus Universitário Darcy Ribeiro Brasília-DF - Brazil. Zip Code 70919-970*

³*Lawyer*

ABSTRACT

Brazilian Judiciary Power promoted access to information culture with the implementation of new management policies involving information and communication technologies (ICT). One of these initiatives resulted in the publication of Justice in Numbers, an annual report published by National Council of Justice (CNJ). This report's goal is to demonstrate productivity metrics and data online about Judiciary entities to any citizen. This paper analyzes qualitatively report's development and how it attends to access to information culture, based on models of government-citizen interaction established by Organization for Economic Cooperation and Development (OECD). That initiative is seen as disruptive to the bureaucratic and hermetic tradition of the Brazilian Judiciary. However, when the communicative capacity of ICT is observed, Justice in Numbers does not achieve its full potential and leaves gaps of effectiveness regarding the dissemination of data in virtual environment.

KEYWORDS

Access to information, Internet, Communication Technologies, Judiciary Power, Justice in Numbers

1. INTRODUCTION

In order to cope with technological and social demand, public policies in Brazil purpose the use of information and communication technologies to bring government and citizens closer. In both executive and legislative powers these policies are generally decentralized. In this context, judiciary power stands out by its insertion in virtual environment, with centralization of data showing productivity of all judicial entities, since 2003, which resulted in Justice in Numbers report, nowadays published by National Council of Justice (CNJ). Systematized to combine the collected data and present it to citizens and judicial managers, Justice in Numbers innovated by gathering administrative information of all judicial entities and investing in institutional cooperation to feed the report. Thus it is relevant to evaluate in which measure the report comply with a better understanding of activities and performance from courts and administrative transparency. The question put into discussion is: From data formats of Justice in Numbers report, what elements of access culture remain outside its strategies? This paper is discussed in two sections, by hypothetical-deductive approach, from a great context to comprehend a specific matter. The first section presents documental research results in order to demonstrate how the report can include judiciary power into access culture introduced by ICT as well as obstacles those institutions may encounter in the process. The second one presents results from direct observation of Justice in Numbers report and also consists in a synthetic comment with a preliminary analysis about advances and deficiencies of the published models.

2. ACESS TO INFORMATION CULTURE: HOW JUSTICE IN NUMBERS REPORTS INSERT IN THIS CONTEXT

Internet transforms the obstacles represented by space and time, creating a new pattern of interactions that inaugurates information society (CASTELLS, 2011). By keeping high communicative potential, internet can be a key tool for democracy advancement. Permeability between government - here understood as the Executive, Legislative and Judicial Branches - and citizen can enhance if inherent characteristics of virtual environment are properly used. Civil society has a chance to step out of its role as a mere consumer of information, being empowered to debate, produce and transmit information about and to the government, as recognized by Gomes (2005, p.218) and Felice (2008, p.56). OECD identifies three different models for this communication structure, which are: information (one-way relation from government to citizen), consultation (two-way dialog with feedback from citizens) and active participation, with engagement from citizens in policy-making (OECD, 2003, p. 30). As a reaction to the cultural transformation promoted by ICT interaction with society, Brazilian government promoted various actions in order to insert its entities in virtual environment. A legal landmark is Federal Statute number 12.527 of 2011, known as Access to Information Act (Lei de Acesso à Informação - LAI), which establishes that every citizen has the right to access a wide range of governmental data, as a legit owner of it, and that the constitutionally protected right to access shall be limited only in exceptional.

Electronic government strategies on Legislative, Executive and Judiciary reached the three OECD communication models described above, although not evenly among them, as the projects are still in process of improvement and overcoming of barriers, such as bureaucracy, governmental culture and Brazilian public administration model. Although the expression "electronic government" is generally applied only to Executive branch, Judiciary has a similar initiative within its scope. CNJ, internal controller of Judiciary, has published several regulations to informatize - and, consequently, increase productivity and visibility - of judicial performance (SILVA et al, 2013). If, in one hand, it can be said that CNJ formally joined the new communication structures by the information OECD model, on the other hand, the doubt whether it would be at an advanced stage remains. Although electronic government strategies in Brazil are normatively oriented (in the sense of written regulations) to the standardization of both visual strategies and the organization and accessibility of public information, in a practical application some problems are verified (MARTINS, KURTZ, 2013; OLIVEIRA, DINARTE, 2012). Two deficiencies standout: i) adherence of the entities to the guidelines is not unanimous, which leads to a diversity of implementations not necessarily convergent with culture of access to information or with potential of government-society contact, as already raised in quantitative and qualitative studies (SANTOS et al, 2012; HOCH et al, 2012); and ii) Brazilian public administration's own strategy of keeping information separate in several websites, one for each organization, only reflects what exists on physical environment, instead of proposing a model that centralizes data and services, taking advantage of the possibilities of network interconnection and conveniences that it would bring to the. However, there is a differentiated proposition for the use of ICT by Judiciary, which is the focus of this paper: it intends to centralize courts' information in a single annual online report, through CNJ, which publishes the performance of Brazilian judiciary institutions: Justice in Numbers.

3. JUSTICE IN NUMBERS REPORTS: METHODOLOGY AND ANALYSIS OF EDITIONS 2004 TO 2017

This section will focus on the general innovations, peculiarities and continuities between versions of Justice in Numbers report from the point of view of interface and information presented in them. The report already has fourteen editions, and between the first edition, having as base year 2003 and published in 2004, and the most recent, base year 2016 and publication in 2017, several changes were made, both on the indicator and the way to present and process the report. In order to point out what has changed, general aspects of observation were chosen and analyzed, checking if the following items were present. The first period of Justice in Numbers collected data from Federal, State and Labor Brazilian courts. The report originated in 2003, based on a research requested by Federal Supreme Court to the Center for Public Opinion Research of the University of Brasília (DATAUnB), in order to build an integrated information system of the Judiciary

Power, mainly focused on indicators statistics (CNJ, 2008a). The focus given to the study, therefore, was for internal management of Judiciary. The initial version was a raw version, published in 2004, with a quantitative approach, composed only of a sequence of tables and graphs in a pdf file, with some explanations about the terms considered for the calculations (CNJ, 2004). Although this version did not reflect the priority of access to information, it was the first step to implement a management policy in Judicial Branch, which is the effective context of the debates about the Constitutional Amendment nº 45/2004 that created the CNJ. Therefore, Justice in Numbers report has been created under a proposal for greater democratization of the Judiciary, which should have systematization of its management and greater possibility of control over it. The form adopted to the reissued publications for the base years 2004 to 2009, was digital books, structured to contain identification of the responsible team, a section with an introduction to the report and a summary. At this point indicators were standardized, by four categories: (i) income, allocation and degrees of use, (ii) litigation, (iii) access to justice, and (iv) profile of judicial demands, as well as concepts used in research, as established in CNJ Resolution 15 of, repealed by Resolution number 76 of 2009 (CNJ, 2004; 2005; 2006; 2007).

The model used for that period reflects the lack of contact from Judicial Branch with civil society, or even lack of interest in such opening, given the strong technicality of the data presented on the report. Although the concepts insertion and systematization have represented advances if compared to the first version, the charts presented without any context would hardly transmit to a lay person the message of how is the Court's performance. In an analysis of this first period remains a great gap between what the technologies made available at the time and their use to deliver data to civil society, that should qualify citizens to debate about a more efficient and effective management of Judiciary. Few was invested in turning data friendly to civil society - an approach that began to change from the 2008 report on. In the 2008 report (CNJ, 2008a), base-year 2007, changes in the way of releasing data are perceived, mainly by the publication of another file, with the graphics and concepts, called "Brief Analysis", which presented to citizens interpretations regarding the graphics. Throughout the publication, there are explanations about the data that go beyond the concepts of the previous versions and have a commentary character, which enriches the information therein. (CNJ, 2008b, p.11). In 2009, the traditional report, in a graphics and concepts format is followed by two additional documents: historical series (State Justice, Labor Justice and Federal Justice, from 2004 to 2008) and commented summary of the series. The series' summary comes with commented graphics reinforcing the idea of taking the data and report for debate beyond internal scope of Judiciary (CNJ, 2008). In 2010, the publication gains an executive summary that counts on the predominance of comments about graphics. It also innovates by presenting new indicators, defined by CNJ Resolution nº 76 of 2009, maintaining in the summary two of the categories, which become more detailed: i) incomes, allocation and degrees of use and ii) litigation. It presents a report for each branch, in addition to a General Judicial Overview document (covering joint analysis of all branches), published in English and Spanish versions (CNJ, 2009). The 2010 base-year version follows the pace of innovation, with an Interactive Report which united figures, links and numerical summaries of indicators, as well as an Overview, with commented data on the report itself and an executive summary composed only of comments. (CNJ, 2011).

This proposal for the information coverage permeates the Interactive Report, which seeks to give a more ludic form – links embedded in a map image - to the data that appear in a raw stage in the main report, because there is a much greater potential of ICT to give content plasticity (LÉVY, 2010, p. 122). It is visible the attempt to reformulate the reports presentation in the following years, but this limitation persists. It can be said that Justice in Numbers report has gone through a change of perspective in 2011. In its previous versions it aimed to publish a data collection showing Judiciary's concern with internal planning and management, whereas the 2011 version would focus on the possibility of the citizen to understand the functioning and the limitations faced by Judiciary. This year was of beginning of validity for Access to Information Act, during the execution of the report published in 2012. Given the new guidelines made under the Access to Information Act, the 2011 report sought to meet elements of access culture (CNJ, 2011). On this version of Justice in Numbers in addition to the State, Federal and Labor Courts data, there are those of Electoral, State Military and Superior Courts. As execution proceedings was shown to represent great costs to the Judiciary, a special category for those data was created, so the dimensions of this problem can be acknowledged. Since the 2012 version of the report, there was also improvement of statistics calculation about data presented, with implementation of Compared Productivity Indicator (IPC-Jus). (CNJ, 2012) And in the version published on 2014 (base-year 2013), the report shows more direction to access culture. This report also presented a new edition of interactive reports. Despite investments in visual strategies to match

infographics and map in order to make data comprehension more ludic, the interactivity character remained weak, restricted to the choice, through clicking, about what data to access (CNJ, 2013).

Justice in Numbers between 2015 and 2017 is characterized by the broadening of the report's scope. Procedure time, more demanded issues, concern with criminal justice and with digitalization of Judiciary are highlights of this report's new approach. Over the last three years, "profile of demands" category was rescued, which was also included on publications from 2004 to 2009. However, if in former reports the data purpose was the accounting of cases in which government figured as litigant, new versions present data about the main subjects of the demands. "Digital justice" was addressed on a topic on 2015 report, recognizing benefits that the virtualization of the process brings to gathering information and diagnosis about Judiciary. (CNJ, 2015). Data about the procedures length and number of proceedings in all judicial branches also began to be presented, with a separate measuring of execution and legal proceedings. The concern with the enforcement of execution process continued to be shown, with an item on "execution proceedings bottlenecks", which, however, is restricted to knowing the length and productivity of the execution proceedings, as well as a comparison between cases in which there was conciliation and those where there was not. (CNJ, 2016; 2017) The report points out that although this is a judicial branch with fewer cases, it is startling that there is a continuous increase of criminal demands, because of its character as the last coercive measure, as well as the fact that dealing with freedom restriction requires a slower process to ensure proper processing (CNJ, 2016; 2017). Contrary to the initial assumption, Justice in Numbers did not always cover all branches of Judiciary, and was restricted to data collected in State, Federal and Labor Justice for a long period. Its scope expansion for other Courts was only established on the last six versions, with the establishment in 2011 of statistical indicators adequate to the specifics of each branch. The presence of conclusions together with data also shows greater concern with those who access the report. The reports also began to show extra indicator categories, such as IPC-Jus and impact of execution proceedings, in order to offer proper management parameters. Those were based on academic contributions, which show certain permeability of the report to contributors external to Judiciary. The presentation of sociological data was also included, with profile of legal demands, criminal justice and electronic process, allowing to acknowledge social problems that end up in Judiciary, besides the observation of modernization advances in which it is invested to make legal cases resolve more rapidly. The report profile is also more about diagnosing external problems or fault observation. It is necessary to go beyond establishing that there is a crisis, in order to study it deeply in order to allow the construction of solutions and the evaluation of these by the citizen.

4. CONCLUSION

This paper has presented an analysis of Justice in Numbers report based on what data is presented to citizens by it and how it complies with culture access. Although the general nature of Justice in Numbers as a compiler of the Judiciary data in all its branches is recent, advances are in line with the idea of improving the report's interface with citizen. However, as far as the interactive character of ICT is concerned, data shows that initiatives of Justice in Numbers remains innocuous, since the format presented is very similar to what could be done in a physical platform. In an overview, which deserves further reflection and study, it is clear that there is no specific space in the report to indicate measures to be taken within the scope of Judiciary management - that the citizen doesn't get access to what measures will be taken to improve indicators, and, therefore, cannot act over it. It is necessary to further compare this management tool with others in order to make Justice in Numbers information go beyond a one-way street, where the public is treated as a mere recipient of the data, and seek for measures for it to be an enabler of participation.

REFERENCES

- Castells, M., 2011. *A sociedade em rede*. Trad. Roneide Venancio Majer, Paz e Terra, São Paulo.
- _____, 2004. *Justiça em números* 2004. <http://www.cnj.jus.br/images/stories/docs_cnj/relatorios/justica_numeros_2003.pdf> Access: 19 th oct. 2017.
- _____, 2005. *Justiça em números 2005*. <http://www.cnj.jus.br/images/programas/justica-em-numeros/justica_numeros_2004.pdf> Access: 20th oct. 2017.

- _____, 2006. *Justiça em números 2006.* < http://www.cnj.jus.br/images/programas/justica-em-numeros/justica_numeros_2005.pdf > Access: 20th oct. 2017.
- _____, 2007. *Justiça em números 2007.* < http://www.cnj.jus.br/images/programas/justica-em-numeros/justica_numeros_2006.pdf > Access: 20th oct. 2017.
- _____, 2008a. *Breve Análise do Poder Judiciário.* <http://www.cnj.jus.br/images/stories/docs_cnj/relatorios/justica_em_numeros_volume_2.pdf > Access: 21th oct. 2017.
- _____, 2008b. *Justiça em números 2008.* <http://www.cnj.jus.br/images/stories/docs_cnj/relatorios/justica_em_numeros_2007.pdf > Access: 21th oct. 2017.
- _____, 2009. *Panorama do Judiciário 2009.* <http://www.cnj.jus.br/images/stories/docs_cnj/relatorios/justica_em_numeros_2008.pdf> Acess: 21th oct. 2017.
- _____, 2010. *Panorama do Judiciário 2010.* <http://www.cnj.jus.br/images/programas/justica-em-numeros/rel_sintetico_jn2009.pdf > Access: 21th oct. 2017.
- _____, 2011. *Justiça em números 2011.* http://www.cnj.jus.br/images/programas/justica-em-numeros/2010/rel_justica_numeros_2010.pdf > Access: 21th oct. 2017.
- _____, 2012. *Justiça em números 2012.* <<http://www.cnj.jus.br/programas-e-acoes/pj-justica-em-numeros>> Access: 21th oct. 2017.
- _____, 2013. *Justiça em números 2013.* <http://www.cnj.jus.br/images/pesquisas-judiciais/Publicacoes/relatorio_jn2013.pdf> Access: 21th oct. 2017.
- _____, 2014. *Justiça em números 2014.* <ftp://ftp.cnj.jus.br/Justica_em_Numeros/relatorio_jn2014.pdf> Access: 22th oct. 2017.
- _____, 2015. *Justiça em números 2015.* em: <ftp://ftp.cnj.jus.br/Justica_em_Numeros/relatorio_jn2015.zip> Access: 22th oct. 2017.
- _____, 2016. *Justiça em números 2016.* <<http://www.cnj.jus.br/files/conteudo/arquivo/2016/10/b8f46be3dbbf344931a933579915488.pdf>> Access: 23th oct. 2017.
- _____, 2017. *Justiça em números 2017.* <<http://www.cnj.jus.br/files/conteudo/arquivo/2017/11/100d387b0339d6d8544a29e30a3b2150.pdf>> Access: 23th oct. 2017.
- Felice, M., 2008. Das tecnologias da democracia para as tecnologias da colaboração. In: *Do público para as redes: a comunicação digital e as novas formas de participação social.* FELICE, M. D. (org.), Difusão, São Caetano do Sul.
- Gomes, W., 2005. A democracia digital e o problema da participação civil na decisão política. *Revista Fronteiras – estudos midiáticos.* Unisinos, set./dez.
- Hoch, P. A.; Righi, L. M.; Silva, R. L., 2012. Desafios à concretização da transparência ativa na internet, à luz da Lei de Acesso à Informação Pública: análise dos portais dos Tribunais Regionais Federais. *Revista Direitos Emergentes na Sociedade Global.* Vol. 1, No. 2, pp. 257-286.
- Kurtz, L. P.; Silva, R. L., 2014. Da tradição de hermetismo à cultura da transparência: o Poder Judiciário na era do acesso à informação. In: Rover, A. J.; Cella, J. R. G; Ayuda, F. G. (Org.). *Direito e Novas Tecnologias.* Florianópolis: FUNJAB, pp. 214-241. <<http://www.publicadireito.com.br/artigos/?cod=33de4c785fc85dfc>> Access: 30th oct. 2017.
- Lévy, P., 2010. *As Tecnologias da Inteligência: o futuro do pensamento na era da informática.* Trad. de Carlos Irineu da Costa, 34, Rio de Janeiro.
- Martins, A. C. L.; Kurtz, L. P, 2013. Novos Horizontes Democráticos Abertos pela Internet: dados e análise sobre a comunicabilidade dos sites do governo eletrônico brasileiro. In: Rover, A.; Simão Filho, A.; Pinheiro, R. F. (Org.). *Direito e novas tecnologias.* 1. ed. Florianópolis: FUNJAB, pp. 33-62. <<http://www.publicadireito.com.br/artigos/?cod=636efd4f9ae5781>>. Access: 15th oct. 2017.
- Oliveira, G. F.; Dinarte, P. V., 2012. Governo Eletrônico e a disponibilização do serviço “Fale Conosco” em seus portais: a potencialização da democracia e efetivação do exercício dos direitos fundamentais dos cidadãos. *IX Seminário Internacional de Demandas Sociais e Políticas Públicas na Sociedade Contemporânea.* Santa Cruz: Unisc <<http://nudiufsm.files.wordpress.com/2012/06/36.pdf>>. Access: 16th oct. 2017.
- OECD - Organization for Economic Cooperation and Development, 2003. *Promise and Problems of E-Democracy: challenges of online citizen engagement.* Paris: OECD PUBLICATIONS. <<http://www.oecd.org/internet/public-innovation/35176328.pdf>>. Access: 17th oct. 2017.
- Santos, P. M., Bernardes, M. B., Rover, A. J, 2012. *Teoria e Prática de Governo Aberto: lei de acesso à informação nos executivos municipais da região Sul.* Fundação Boiteux, Florianópolis.
- Silva, R. L.; Hoch, P. A.; Righi, L.M, 2013. Transparência pública e a atuação normativa do CNJ. *Revista Direito GV,* São Paulo, ed. jul./dez. pp. 489-514. <http://direitosp.fgv.br/sites/direitosp.fgv.br/files/artigo-Edicao-revista/06-rev18_489-514_-_rosane_leal_da_silva.pdf>. Access: 20th oct. 2017.

E-READINESS COMPOSITE INDICATORS MEASUREMENT METHODOLOGIES: LITERATURE REVIEW

Jaouad Chennouf and Najib Belkhayat
LAMAI, Cadi Ayyad University, Marrakech, Morocco

ABSTRACT

The number of IT indexes has become very important in the last few years; IT (Information technology) indexes serve countries to evaluate their performance over time and make comparison with others. Every year, new indexes are created and others are cancelled due to different reasons: fusions, repetitions, those which are no longer useful... but the most important of which are difficulties in collecting data. Nowadays, organizations tend to update the conception of their composite indicators to reduce the number of included indicators without losing the indexes' relevancy. In this context, we will report in the present paper the composite indicator's measurement methodologies and represent a comparative analysis of them.

KEYWORDS

Comparison, Readiness, Composite, Indicators, Measurement, Methodology

1. INTRODUCTION

In the recent years, several attempts to set up e-readiness indexes to express how deep information technology has been well integrated in countries, because of its important role in monitoring countries development.

The main objective of composite indicators is to provide a real picture on the development of countries, but this is limited by many difficulties. The main limit is the difficulty of data collecting (very expensive, time consuming, annually/biannually indexes, repetition...), in addition too few existing solutions to help countries in monitoring their performances.

Firstly, this paper aims to compare ICT (Information and Communications Technology) composite indicators methodologies: IDI (ICT Development index), NRI (Network Readiness Index) and EGDI (E-government Development Index) following the main recommendations of (OECD, 2008) to construct composite indicators. Secondly, this paper presents their limitations and finally to discuss solutions to the limitations.

The rest of this paper is structured as follows: Section II will deal with these composite indicators and compare their conceptual frameworks. Section III will compare the methodologies used in data manipulation and will be ended by comparing the countries outcomes of the year of 2016 according to these composite indicators. Finally, we will end with by discussions and conclusions.

2. CONCEPTUAL FRAMEWORK COMPARISON

The three composite indicators are chosen as they cover a great number of countries, also because they are very active and maintained by international well known organizations. Indeed, they cover more than 138 countries. NRI was created by WEF (The World Economic Forum), ICT was created by ITU (International Telecommunication Union) and EGDI was created by UNPAP (United Nations Public Administration Programme). All these composite indicators are annually updated except EGDI which is biannually updated.

2.1 Design and Indicators Selection

The objective of the NRI is to measure the propensity of countries in order to exploit the opportunities offered by ICT. Figure 1 shows three main aspects of ICT that are represented by NRI: readiness, usage and impact indicators. The IDI is used to monitor and compare developments in ICT between countries and over time. Figure 2. shows three main aspects of ICT represented by IDI: readiness, usage and impact indicators.

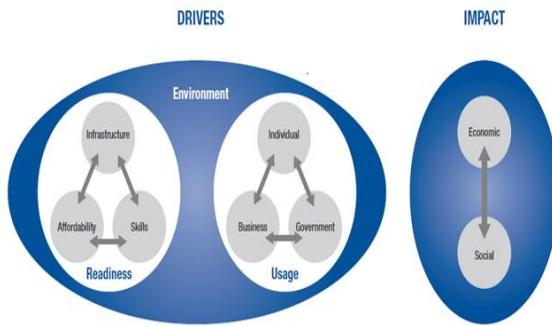


Figure 1. NRI Conceptual framework (WEF, 2016)

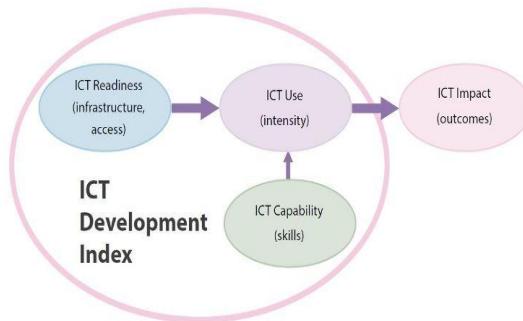


Figure 2. IDI Conceptual framework (ITU, 2016)

The IDI selection of indicators and the allocation of sub-index weights are based on an inadequate model. (Ahmadi, 2015).

The objective of EGDI is to rate the performance of national governments (UN, 2016). The EGDI framework is presented as follows:

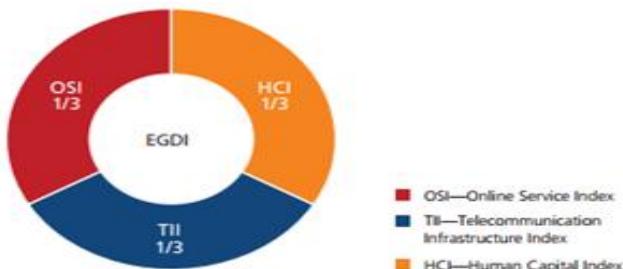


Figure 3. EGDI Conceptual framework (UN, 2016)

Figure 3. shows that the EGDI framework is based also on three ICT aspects: usage indicators represented by online service index, Readiness indicators represented by telecommunications infrastructure and impact indicators represented by human capital.

The methodological framework of EGDI has remained consistent and adjusted every year to respond to the trends of governments' strategies (UN, 2016).

2.2 Composite Indicators Relevancy Analysis

After defining indicators to be used in each composite indicator, its objective is checked to verify whether it is well expressed and does respond to the objective using the indicators selected.

There are two ways to analyze relevancy: by indicators or by countries (OECD, 2008):

- The PCA (Principal Component Analysis) method, FA (Factor analysis) or Cronbach coefficient alpha (c-alpha).
- The CLA method (Cluster Analysis).

The IDI uses PCA method to reduce the dimensionality in data and to convert a large number of variables into a smaller data set without losing too much information. PCA is used to reduce the cost of data collection.

Thanks to PCA, IDI has reduced the number of the selected indicators from 20 to 11 (ITU, 2009) (Belkhayat, 2015) has proposed a predictive e-readiness layered model (ELM) to reduce composite indicators measures from 20 to 6.

2.3 Sensitivity Analysis

In this section, the composite indicators are tested on how their individual inputs affect the output of a composite index, for example the selection of indicators, data normalization, weights and aggregation methods

The IDI states that it is extremely robust in regard to different methodologies (ITU, 2016), but (Giuseppe, 2011) proved using PCA that by using only four components of the index, we can get 90% of the results.

This step has been well received by organizations and it was used in IDI which prove good results of robustness. Regarding the two other composite indicators, we haven't found any robustness tests in the literature. So, these tests deserve to be established, in order to guarantee the robustness of the composite indicators components to represent clearly their fields.

There is no ideal methodology for the selection of indicators. Each organization proposes a set of indicators according to its experts in the field and analyzes if the chosen indicators represents actually the phenomenon concerned. In the case of ICT, the selection of indicators is very volatile due to the quick evolution in the use of technologies. The ICT indexes frameworks are based on the same aspects to represent ICT indexes: Readiness, Usage and Impact indicators. Indeed, according to the results obtained and critics, organizations can replace, add or reduce indicators. But any change in the composite indicators frameworks can affect the possibility of data comparison over time. All these make decisions related to changing composite indicators frameworks very difficult.

3. DATA MANIPULATION

3.1 Design Data Collection and Imputation of Missing Data

3.1.1 Data Collection

Organizations when collecting data for the input of the composite indicators ensure the use of existing indicators or if necessary seek new sources of data, with the principle that if we have poor data in the input, we will certainly have bad results in the output (OECD, 2008). The techniques used by organizations to collect data are: Interviews, Questionnaires and surveys: using a list of questions, observations, focus group, history and case studies, documents and records.

All the composite indicators studied rely on data from different organizations: ITU, UNESCO (The United Nations Educational, Scientific and Cultural Organization), UN (United Nations), WORLD BANK....

The ITU collects data through questionnaire, statistical tables from UNESCO and directly from operators (ITU, 2016). Statistical surveys are most complex and expensive for data collection (UN, 2014)

Many developing countries does not collect data on household statistics, which impact the quality of the indicators (examples: Percentage of households with a computer, Percentage of households with Internet access, Percentage of individuals using the Internet) (ITU, 2016).

The censuses are the most costly data collection activity that national statistics system undertakes (World Bank, 2004).

3.1.2 Imputation of Missing Data

Organizations make very attention to impute missing data because it can influence the entire exercise when data are not imputed correctly.

The UIS (UNESCO Institute for Statistics) analyze the data collected in 4 steps (UNESCO, 2016): Review the data with a series of automatic checking, request the sender to justify why data are partial or missing, generate a complete data, the UIS estimates and imputes missing data and send the data to national authorities to validate, then the data are published.

The Eurostat Approach performs a lot of checks to validate data collected (logical, coherence, comparison with external indicators...) complete data with enquiries by phone mails... Also the past experience has shown that processing errors can be minimized using electronic, web, and online web questionnaires checks (EUROSTAT, 2016).

(OECD, 2008) Different kinds of missing data exist (MAR (Missing at Random), MCAR (Missing Completely at Random), NMAR (Not Missing At Random)) and only three approaches to impute missing data: case deletion (delete all missing records), single imputation and multiple imputation. Methods to impute missing data for example the mean, median, mode substitution, regression imputation, hot-and cold-deck imputation, expectation-maximization imputation, or multiple imputations. The purpose of data imputation is the minimization of biased data and the use of expensive data to collect that would otherwise be eliminated by deleting a case.

The IDI uses the hot-deck imputation, while the NRI use PPP estimate (Purchasing Power Parity) method.

The hot deck imputation is a method for handling missing data using a similar set of data and it is no longer recommended to use it because it gives bias results while used in clinical studies (Frank J. Molnar MSc MDCM, 2008), while the PPP estimate rests on comparing countries currency using the cost of a basket bag in a similar country.

The advantage of PPP is its stability over time and its drawback is mostly difficult to estimate even in a group in the same country, as the measures depends on the nature of the group: rural or urban center (Samuel Moon, June 2010)

Based on the findings, the data collection is the most delicate step, if it is not well done, we surely will come up with bad results. The "Digital divide" between countries requires that the data collection must be done according to the nature of each country and each region, in order to have a very high quality of results, which will multiply data collection costs. To help reducing the data collection costs, organizations use missing data imputation methods to estimate the results, for example using Gdp by capita or PPP estimate between similar countries. Each organization chooses one or more methods for imputing missing data, according to those that have given better estimates over time, except EGOV which relies on reading all countries laws by locally experts without having missing data.

3.2 Method of Standardization

There are many methods of standardization: Ranking, standardization, min-max, distance to a reference, categorical scale, indicators above or below the mean, Balance of opinions, percentage of annual differences over consecutive years (OECD, 2008).

The IDI uses the distance to reference and NRI uses the min max transformation.

The normalization scale differs from a composite indicators to other, the NRI uses a scale from 1-7 and the IDI uses a scale from 1-10.

3.3 Calculating: Weighting and Aggregating

A number of weighting techniques exist: some are derived from statistical models, such as FA, or from participatory methods like equal weighting (EW). PCA and FA are used to group individual indicators according to their degree of correlation (OECD, 2008).

The IDI uses a standard weighting for all indicators of all countries: 40% access, 40% uses, 20% skills. The IDI has verified this weighting using PCA method. The NRI use EW for all indicators.

Equal weights represent a lack of knowledge about a correct weighting scheme (Ahmadi, 2015)

Due to different number of indicators in the sub-index access compared to the number of indicators in the sub-index use and the number of indicators in sub-index skills (respectively five, three, three), the ITU is implicitly allocating unequal weights to the individual indicators in the sub-index access without justifying (Ahmadi, 2015)

All these composite indicators use a static weighting method: EW for NRI and EGDI, IDI with fixed pillars weights. However it is preferable to invest in testing a dynamic method that takes into consideration the "Digital Divide" to see results, rather than giving static and unjustified weights.

3.4 Rank Correlation of the Composite Indicators Countries Results

After comparing these composite indicators, we took the results of the latest version of the reports published in this year 2017 of all 108 common countries covered by the three composite indicators, and we have calculated their correlation using Spearman and Pearson techniques, Table 1 and Table 2 show the results:

Table 1. Spearman correlation MATRIX

Composite Index	IDI	NRI	EGDI
IDI	1	0,95	0,96
NRI	0,95	1	0,94
EGDI	0,96	0,94	1

Table 2. Pearson correlation MATRIX

Composite Index	IDI	NRI	EGDI
IDI	1	0,93	0,96
NRI	0,93	1	0,93
EGDI	0,96	0,93	1

Table 1 and Table 2 show a very high Spearman and Pearson correlation between all the ICT indexes.

Results show a strong correlation between IDI, NRI and EGDI, respectively using Spearman correlation 0.95 and 0.96, which means that a country with a high score in IDI in 2016 has got a high score in NRI and EGDI. Also, we can observe that the different used methods give correlated results and consequently we decide to push our research in order to find the minimum set of the most relevant indicators which would give correlated results to those observed.

4. CONCLUSION

The comparison established on the composite indicators has confirmed the difficulties encountered by the organizations during the elaboration of the indicators, notably the collection of data: limits of the data collected, the high cost of the data collection, dealing with missing data. This affects the quality of the published results, their frequency and their coverage; pushing organizations to maintain a reasonable balance between the numbers of indicators, the number of countries covered by composite indicators and the relevancy of the composite indicators to represent the phenomenon concerned.

Finally, the very strong correlation between ICT indexes NRI, IDI and EGDI offers us new axes of analysis to identify the reasons behind this strong correlation and to take advantage on it. Consequently, we can achieve good results by aggregating these composite indicators and minimizing the number of indicators included taking into account the fact that NRI is composed of 53 indicators whereas IDI is composed of only 11 indicators and they have extremely correlated results.

REFERENCES

- Ahmadi, T. et al, 2015. Composite indices for the evaluation of a country's information technology development level: Extensions of the IDI of the ITU. *Elsevier*. pp. 174-185
- Belkhayat, N. et al, 2015. Dynamic Systems State Model For E-Readiness Estimation. ICDS, pp. 1-5.
- Eurostat, 2016. *Prodcom data quality*, Eurostat.
- Frank, J. et al, 2008. Does analysis using "last observation carried forward" introduce bias in dementia research?, *CMAJ*, pp 751–753.
- Giuseppe, B. et al, 2011. A critical analysis of digital divide measurement. *The Information Society*, Vol 27, pp. 16-28.
- ITU, 2016. *Measuring the information society report*.
- OECD, 2008. *handbook on constructing composite indicators: methodology and user guide*.
- Samuel, M. et al, 2010. Purchasing power parity measures: advantages and limitations. *Overseas Development Institute*.
- The world bank, 2004. *Better data for better results*.
- UN, 2016. *United nations e-government survey*.
- UNESCO, December 2016. *Uis frequently asked questions*.
- WEF, 2016. *The global information technology report*.
- WeiWen, W et al, 2011. Exploring the critical pillars and causal relations within the NRI: An innovative approach. *Elsevier*. Hsinchu, Taiwan, pp. 230–238.

PROTECTION OF INTELLECTUAL PROPERTY RIGHTS IN POLAND FROM THE PERSPECTIVE OF POLISH YOUTH

Lukasz Tomczyk

Pedagogical University of Cracow, Ingardena 4 str., 30-060, Kraków, Poland

ABSTRACT

The paper outlines the phenomenon of digital piracy in Poland with particular focus on young Internet users. In the introduction, characteristics of piracy are presented by showing the scale and mechanisms of downloading files from illegal sources. Empirical part presents solutions to reduce this type of risk behavior, based on opinions shared by adolescents. The author tries to present different answers to the question: What activities could, in your opinion, eliminate digital piracy? Presented results of qualitative research are part of the project "Computer piracy as a challenge for media pedagogy BS-494 / P / 2017" and are in line with the risk paradigm of media pedagogy.

KEYWORDS

Digital piracy, Youth, Poland, Prevention, Intellectual Property

1. DIGITAL PIRACY AMONG ADOLESCENTS — INTRODUCTION TO TOPIC

There are many publications presenting determinants of digital piracy among youth. Quantitative research imply that self-control, time spent on using new media and other problematic online behaviours contribute to increased illegal file download by young people (Kim, Kim, 2015). Level of digital competencies and social susceptibility to take over negative habits should be added to these factors (Zhengchuan, Qing, Chenghong, 2013). It is also interesting that young people with strong ethics and ability to control their own behaviours in order to comply with certain norms, download illegal files much less often (Chan, Ma, Wong, 2013). According to other researchers, the fact that adolescents see file downloading as unethical is not enough to prevent them from pirating because this behaviour is determined by other factors such as: prices of files or availability of legal alternatives (Vermeir, 2009). Analysis of key determinants suggests that level of piracy, when it comes to entertainment content, is also connected with: gender, experience and ways of using the Internet, membership in groups promoting similar behaviour and type of school (Malin, Fowers, 2009).

Subject matter publications also include attempts to show piracy in the perspective of measuring how much school students are aware of consequences resulting from piracy (Moores, Nill, Rothenberger, 2009), and qualitative interpretations of piracy determinants among young people (Konstantakis, Palaigeorgiou, Siozos, Tsoukalas, 2010). However, analysis of risky behaviours in the Internet requires a deeper study into the multi-factorial character of this phenomenon (Pyżalski, 2012). That is why this paper is a rare attempt to present piracy in the context of solutions to limit downloading activity, suggested by young Internet users.

2. METHODS

The research objective was to present opinions of young Internet users about solutions to minimise digital piracy. The objective was met by means of anonymous quantitative study (diagnostic survey) with an open question included. The study was conducted among 1220 students of lower- and upper-secondary schools in

Poland (Silesian voivodeship). The survey took place in September and October 2017. Collected answers were categorised into 18 different interpretations of how piracy should be minimised.

Each answer that was an exemplification of a given category of interpreting piracy-reducing mechanisms, was coded. With each quote, the following characteristics were marked in brackets: gender (F — Female, M — Male), type of school (L — lower secondary, G — general upper secondary, T — secondary technical school) and respondent's age.

3. RESEARCH RESULTS - WAYS TO MINIMISE DIGITAL PIRACY IN OPINIONS OF ADOLESCENTS

Interpretations of mechanisms that would - according to young people - limit digital piracy are presented below. This type of phenomenological approach facilitates even more precise understanding of the phenomenon that is not yet sufficiently investigated in the literature of the subject, e.g. in media pedagogy publications (Read, 2007).

According to the adolescents, blocking services that offer access to pirated files is a primary alternative to illegal file download. Lack of access can be executed on national level (e.g. by dedicated organizations) or by installing specialist software that decides on behalf of users.

Block such websites (F, L, 13)

Blocking pirate websites (M, T, 18)

Adolescents believe that to block access to files also means to minimise use of programs that provide access to P2P networks and mass downloads from data repositories (Chang & Walter, 2015). These applications enable easy, multiple downloading files from various sources. Some of them also share files during download, and that exposes unaware users to legal liability for distributing copyrighted materials.

Access to these programs should be more difficult (F, G, 17)

Disable websites and apps with illegal materials (F, G, 16)

Higher security of torrent network (M, G, 16)

According to youths, filtering websites and blocking access to them is one of the key factors. Removal of illegal files reduces the scale of the phenomenon easily because it removes opportunities for easy file downloads. Adolescents think that services which store files previously classified as pirate should be controlled in the first place. However, in the age of advanced methods of sharing files through network devices, this idea does not seem feasible from a technical point of view (Phau, Teah, Liang, 2016). Every removed website or protocol will be replaced by other, more advanced solutions.

Better security measures on websites (F, L, 15)

More strict terms of use of community portals, and regular and thorough control of sites like chomikuj.pl (M, T, 16)

Other solution-postulate is to reduce prices of digital files. This suggestion surely results from respondents' financial standing. Many young users point out to the forms of distribution of files and their prices. Because of lower online distribution costs, films, books and software should be cheaper — young people say. This would significantly reduced the scale of piracy (Huang, Lin, Fang, 2017).

Cheaper software is the only way (M, L, 13)

Maybe if everything was cheaper. Not everyone can afford to spend money all the time, everywhere (F, L, 15)

Introduce commonly available music and video files, for much lower price so everyone can use them online (F, G, 16)

Lower prices of CD or movies are greater motivations to buy (F, L, 13)

Free access to films, music and other programs (F, T, 17)

Nothing would eliminate digital piracy because it is impossible. One can only reduce digital piracy by lowering prices (M, T, 17)

According to young new media users, this situation is the result of general economic condition of our society. Considering the fact that prices of copyrighted works in Poland are at similar level to those in richer countries, this situation forces illegal use of digital materials. Financial aspect is one of the typical arguments given in the survey to justify pirating behaviours (Sundararajan, 2004).

Change the way a “typical Pole” thinks. We are famous for pirating, because we earn little and we can’t afford on legal software. When a statistical Pole would hear they have to pay 400 złoty for an operating system, they would laugh. For people earning in euro, 100 euro is like 100 złoty for us — it is some expense but if you tell them it is necessary, they will accept it. Not mentioning that computer games, movies or music are really minor costs abroad (M, G, 16)

Behaviours in the virtual world have their consequences and causes in the real world too. The respondents point out to the fact that prices in the real and virtual world are related. Too high prices of entertainment result in young people moving their activities from the real world into the virtual space. It is much easier to access free sources of entertainment (movies, music, magazines and books) over the Internet than in the real world. According to young people, change in prices and marketing offer in the real world may lead to reduction of piracy ratio in social media (KHOJJA, PARK, 2007).

Lower prices of DVDs may motivate people to buy them (F, G, 17)

Cheaper CDs, books and movies (F, G, 17)

Promotions, special deals, encouraging to legal purchases (M, T, 16)

Young Internet users emphasise that piracy is not highly detectable. In a sense, this situation leads to social acceptance of this kind of activities (Rybina, 2011). Adolescents point out that successful reduction of piracy may be also achieved through more intense engagement of services that protect intellectual property.

Control by certain services (F, G, 16)

Increased control by appointed services (F, T, 19)

Authorities could control the Internet. If I’m able to find various files, surely others can do it too (M, G, 16)

Traceability is also connected with introduction of system of penalties for certain behaviour. Many research results confirm that adolescents and adults have no proper knowledge about legal consequences of piracy (Hashim, Kannan, Maximiano, Ulmer, 2014). There is a recurring theme in the analysed responses, of legal liability for distributing files without copyright holder’s consent. The respondents say that severe penalties may contribute to reduction of this type of activities.

Introduction of penalties for piracy (F, T, 18)

High penalties imposed on pirates, e.g. a fine of several thousand (M, L, 14)

Penalty fine (F, T, 19)

Heavier penalties for digital pirates (F, L, 15)

A special regulation that every site should be checked before being published online (M, L, 13)

More severe law regarding crimes connected with digital piracy (M, T, 17)

In their statements, the adolescents point out that change of negative behaviour results from social awareness. Social norms change with increasing awareness, what may translate into classification of certain behaviour as negative. Increased awareness about legal and economic results of the phenomenon is one of the key factors reducing frequency of a given behavior (van Rooij, Fine, Zhang, Wu, 2017). One of the primary awareness raising methods are didactic solutions - assigned to formal or non-formal education category. However, there are also individuals in this group, who point out that illegal file downloading methods are too popular for these solutions to be feasible. Some respondents mentioned also the fact of modelling behaviours that support new standards of usage of digital products.

Society is still little aware how harmful illegal downloading is (M, T, 19)

Increase social awareness about the consequences of piracy (F, T, 18)

Increase awareness about threats resulting from pirating (F, T, 18)

I think it is impossible. Too many people download things illegally from the Internet to detect it all (M, T, 18)

It should be told everywhere that piracy is punishable. Every file that is not original, should be reported e.g. to moderator. Buy original files and do not take free stuff, to show people that this is right. And finally, stay away from pirated products (M, T, 18)

Technical aspects of file security are - in adolescents’ opinion — important element minimising piracy. Improved security measures in computer games or monitoring file legality by operating systems are, according to some, effective solutions. Online verification of legality or software keys are already known mechanisms used by some software producers (Brandes, 2012). These solutions were recognised by the respondents as fulfilling their role.

Better security measures (M, T, 17)

Online verification before it is possible to use program illegally (F, L, 15)

Securing files with proper keys (M, T, 18)

Special programs registering used data, built in computer systems before purchase (F, L, 13)

Anti-virus controlling what we download and if it's legal or not. If it is illegal, it blocks automatic download of this file (M, T, 16)

Perspective of parental control as piracy reducing factor also appears in some statements provided by the young users. However, this answer was not given too often. The fact of knowledge and setting limits by adults is important in minimising risky behaviours in new media space — not only when it comes to digital piracy but also in the area of general development of digital competencies in family environment (Reiss, 2010; Walotek-Ściańska, Szyszka, Wąsiński, Smołucha, 2014). When a computer is used, how much time is spent using new media, control of downloaded and installed files — these are one of many educational activities that facilitate increase in the level of digital safety.

Checking computer once in a while — by parents or some software doing is automatically (F, L, 14)

Services that offer legal access to entertainment content become an alternative to pirated files. VOD channels, music platforms allowing online use of music files are — in the eyes of the respondents — means to reduce piracy (Bhattacharjee, Gopal, Sanders, 2003). Faster and faster Internet connections, opportunity to finance this type of websites and remunerations in audiovisual industry through advertisements or subscription fees become a realistic solution. In this context, it seems interesting to investigate to what extent individuals who use pirated files fulfil their needs to listen to music or watch movies through such distribution channels.

Launching legal websites where you can download music and films (F, L, 13)

More legal sources (F, T, 17)

Buy a subscription package on Spotify (F, L, 13)

Free websites with movies, that work like cda.pl (F, G, 16)

I don't know if there is a perfect solution to the problem but greater popularity of platforms like Soundcloud or Spotify would be helpful (F, G, 16)

We need to add that, according to the young people, piracy results also from users' laziness. Buying entertainment materials requires time to visit a store, whereas online one can easily download every file (Koh, Murthi, Raghunathan, 2014). Archivisation and high availability of all files on dedicated commercial servers is one of the solutions to the analysed problem.

Increased availability, so you don't have to go to store every time you want to watch a movie at home. Truth is, people will watch films online because less and less people want to go to cinema and pay (F, L, 14)

As for technical aspect, eliminating piracy is connected with blocking URLs that offer access to illegal files (Duchêne, Waelbroeck, 2006). Adolescents notice that search engines show links to warez services or similar sites as top search results.

Often piracy is a fault of search engines that show addresses of websites with illegal programs, games and music in the first place (M, G, 16)

No delays in distribution of files through classic distribution channels is another factor that would reduce piracy, respondents say. Quick access to files — the moment they are launched — allows to develop channels other than illegal forms of distribution (Petrović, 2015). Young people also stress out the significance of other, reliable distribution channels based on rental systems.

Files published faster in stores, frequent television premieres (F, T, 17)

Make it easier to buy or rent CDs/movies in nearby stores (F, T, 16)

According to the respondents, educational aspect connected with reducing piracy involves informing society about negative consequences of pirating activities (Janssens, Vandaele, Beken, 2009). In the subject matter literature, young people provide answers that are in line with suggestion of researchers who study online risky behaviours.

Advertisements showing that piracy is wrong (F, T, 19)

There is also a group emphasizing that piracy is an advanced process and it is impossible to remove it completely. It is due to established habits and the ways access to digital files is interpreted. Sceptical approach is typical for older group of interviewed adolescents. Complete elimination of piracy must, in their opinion, involve advanced control of digital resources, also private ones. This postulate is not easy to implement due to lack of social permission and restriction of civic freedoms it entails.

People will always find a way to get something for free or cheaper, so it is difficult to eliminate piracy (M, T, 18)

It is impossible, because every time new piracy website will appear. Internet is too developed to block it all (M, G, 16)

Nothing can limit piracy, there will always be a way to pirate (M, T, 17)

Digital piracy is a problem that is in every country where most of the society have access to the Internet. It will be impossible to reduce it in 100%, but with more effort, this old problem can be significantly limited. Internet users will always find a way to download files on their mobiles free of charge. Everyone breaks copyright somehow, but total security from this kind of law infringement is just impossible (M, G, 16)

There are no actions that could eliminate digital piracy without surveillance of citizens (M, G, 17)

Some young people declare that we should not consider measures to eliminate piracy because, in their opinion, it is not a negative behaviour. This group says that attempts to minimise piracy is an unjustified activity.

None, why eliminate something good (M, T, 17)

None. What is so wrong with that? (F, L, 13)

4. CONCLUSION

The vast majority of the respondents did not provide any answer or declared they did not know the ways to minimise piracy. This may be due to several reasons: lack of knowledge about piracy and its mechanisms, lack of self-reflection in this area, lack of interest in the subject of piracy. Despite common opinion, piracy is not a phenomenon every adolescent is aware of. However, collected answers can be grouped into eighteen categories, partially in line with suggestions regarding reduction of piracy given by professionals from audio-video industry.

Adolescents mention the following means to reduce piracy: blocking websites, removing software that supports downloading and sharing files, deleting copyrighted files, more thorough compliance with terms of use of hosting services, price reduction, change in mentality and attempt to transform social awareness, more cost-effective entertainment offers in the real world, increased controls by institutions that protect copyright, more severe penalties for pirates, more effective technical security measures, increased parental control (younger Internet users), alternative sources of access to files, change of search algorithms in the most popular search engines, faster distribution of entertainment materials in the real world. We should also add to this list opinions that this phenomenon is impossible to eliminate completely, as well as positive evaluations of illegal file download. These answers complement our knowledge about piracy and are valuable in terms of shaping modern digital competencies that include not only technical operation of applications, but also social contexts of new media (Potyrała, 2017).

ACKNOWLEDGEMENT

Text was created as part of the statutory research founded by the Pedagogical University of Cracow, titled “Digital piracy as a challenge for media pedagogy” BS-494/P/2017.

REFERENCES

- Bhattacharjee, S., Gopal, R. D., & Sanders, G. L. 2003. Digital Music and Online Sharing: Software Piracy 2.0?. *Communications of the ACM*, 46(7), 107-111. doi:10.1145/792704.792707
- Brandes, K. 2012. Design Protection in the United States and European Union: Piracy's Detrimental Effects in the Digital World. *Brooklyn Journal Of International Law*, 37(3), 1115-1137.
- Chan, R. K., Ma, K. Y., & Wong, Y. H. 2013. The Software Piracy Decision-Making Process of Chinese Computer Users. *Information Society*, 29(4), 203-218. doi:10.1080/01972243.2013.792302
- Chang, Y., & Walter, J. 2015. Digital piracy: Price-quality competition between legal firms and P2P network hosts. *Information Economics & Policy*, 31/22-32. doi:10.1016/j.infoecopol.2015.04.002

- Duchêne, A., & Waelbroeck, P. 2006. The legal and technological battle in the music industry: Information-push versus information-pull technologies. *International Review Of Law & Economics*, 26(4), 565-580. doi:10.1016/j.irle.2007.01.008
- Hashim, M. J., Kannan, K. N., Maximiano, S., & Ulmer, J. R. 2014. Digital Piracy, Teens, and the Source of Advice: An Experimental Study. *Journal Of Management Information Systems*, 31(2), 211-244. doi:10.2753/MIS0742-1222310208
- Huang, Y., Lin, S., & Fang, C. 2017. Pricing and coordination with consideration of piracy for digital goods in supply chains. *Journal Of Business Research*, 7730-40. doi:10.1016/j.jbusres.2017.03.023
- Janssens, J., Vandaele, S., & Beken, T. V. 2009. The Music Industry on (the) Line? Surviving Music Piracy in a Digital Era. *European Journal Of Crime, Criminal Law & Criminal Justice*, 17(2), 77-96. doi:10.1163/157181709X429105
- Khouja, M., & Park, S. 2007. Optimal Pricing of Digital Experience Goods Under Piracy. *Journal of Management Information Systems*, 24(3), 109-141.
- Kim, J. E., & Kim, J. 2015. Software Piracy among Korean Adolescents: Lessons from Panel Data. *Deviant Behavior*, 36(9), 705-724. doi:10.1080/01639625.2014.977111
- Koh, B., Murthi, B. S., & Raghunathan, S. 2014. Shifting Demand: Online Music Piracy, Physical Music Sales, and Digital Music Sales. *Journal Of Organizational Computing & Electronic Commerce*, 24(4), 366-387. doi:10.1080/10919392.2014.956592
- Konstantakis, N. I., Palaigeorgiou, G. E., Siozos, P. D., & Tsoukalas, I. A. 2010. What Do Computer Science Students Think about Software Piracy?. *Behaviour & Information Technology*, 29(3), 277-285.
- Malin, J., & Fowers, B. J. 2009. Adolescent self-control and music and movie piracy. *Computers In Human Behavior*, 25(3), 718-722. doi:10.1016/j.chb.2008.12.029
- Moores, T. T., Nill, A., & Rothenberger, M. A. 2009. Knowledge of software piracy as an antecedent to reducing pirating behavior. *Journal Of Computer Information Systems*, 50(1), 82-89.
- Petrović, V. 2015. PIRATERIJA FILMSKIH DELA U ERI DIGITALNE TEHNOLOGIJE. *Megatrend Review*, 12(2), 219-236.
- Phau, I., Teah, M., & Liang, J. 2016. Investigating the Factors Influencing Digital Movie Piracy. *Journal Of Promotion Management*, 22(5), 637-664. doi:10.1080/10496491.2016.1185491
- Potyrała, K. 2017. *iEdukacja. Synergia nowych mediów i dydaktyki (Synergy of New Media and Didactics)*. Kraków: Wydawnictwo Uniwersytetu Pedagogicznego.
- Pyżalski, J. 2012. *Agresja elektroniczna i cyberbullying jako nowe ryzykowne zachowania młodzieży*. Oficyna Wydawnicza Impuls, Kraków.
- Read, B. 2007. The First Close Look at Colleges' Digital Pirates. *Chronicle Of Higher Education*, 54(2), A1.
- Reiss, J. 2010, January 1. Student Digital Piracy in the Florida State University System: An Exploratory Study on Its Infrastructural Effects. *ProQuest LLC*
- Rybina, L. 2011. Music Piracy in Transitional Post-Soviet Economies: Ethics, Legislation, and Expertise. *Eurasian Business Review*, 1(1), 3-17.
- Sundararajan, A. 2004. Managing Digital Piracy: Pricing and Protection. *Information Systems Research*, 15(3), 287-308. doi:10.1287/isre.1040.0030
- van Rooij, B., Fine, A., Zhang, Y., & Wu, Y. 2017. Comparative Compliance: Digital Piracy, Deterrence, Social Norms, and Duty in China and the United States. *Law & Policy*, 39(1), 73-93. doi:10.1111/lapo.12071
- Vermeir, I. 2009. The Consumer Who Knew Too Much: Online Movie Piracy By Young Adults. *Advances in Consumer Research*, 36, 929-931.
- Walotek-Ściańska, K., Szyszka, M., Wąsiński, A., Smołucha, D. 2014. *New media in the social spaces. Strategies of influence*. Verbum: Prague.
- Zhengchuan, X., Qing, H., & Chenghong, Z. 2013. Why Computer Talents Become Computer Hackers. *Communications Of The ACM*, 56(4), 64-74. doi:10.1145/2436256.2436272

DATA IN THE CLOUD: WHOSE CONTROL AND WHOSE CONSENT?

George Bouchagiar¹ and Maria Bottis²

¹Attorney-at-law, Tutor/Research Fellow on Data Protection/IP, Ionian University, Kalokeretou 14, Corfu, Greece, 49100

²Associate Professor on Information Law, Ionian University, Ioannou Theotoki 72, Corfu, Greece, 49100

ABSTRACT

Present-day firms process a huge volume of personal data under conditions that they impose. Such data seems to be the new “diamond” that is, stored in the clouds or, kept at companies’ storage facilities. In this paper, scenarios are posed to detect uncertainties that might, under existing laws and norms, occur in the future and it is examined whether personal data could be treated as content of databases and, thus, protected by Intellectual Property laws. Finding that such rules should not govern such data, attention is drawn to the key role that the data subject has to play with regard to the processing of her information.

KEYWORDS

Personal Data, Database, Cloud Computing, Privacy

1. INTRODUCTION

Today the protection of personal data is probably the hottest topic in the European Union. Indeed, a pandemonium of legal scholarship has already been published and is being published every single day. Some authors argue that free circulation and processing of such data is needed, so as to achieve several economically and socially useful purposes. Indeed, data processing is important for crucial purposes, such as health care, education, commerce, crime detection or terrorism prevention (Manovich 2011; Prins Corien 2006; Knoppers & Thorogood 2017; Broesa 2017). Others claim that people have lost control over their data (Solove 2013). More and more academics see data as the fuel of the information economy (Pasquale 2015), while some propose particular classifications of data (Cradock, Stalla-Bourdillon & Millar 2017) or worry about the data of the deceased (Edwards & Harbinja 2013). Academic concerns are numerous. This is fully justified in view of the need to achieve the right balance between the free movement of personal data and the protection of natural persons with regard to the processing of such data. Indeed, this right balance is something that the European lawmaker recently tried to strike, with the new General Data Protection Regulation.

Private enterprises behave as owners of information that they possess (Yeh Chih-Liang, 2017). Personal data (under the Regulation, personal data is any information relating to an identified or identifiable natural person) is shared and circulated under conditions that firms impose with their “standard online contracts” (Guibault 2006), meaning the terms of use and privacy policies (Solove 2013; Yeh Chih-Liang 2017), or their “digital agreements” (Radin 1996) that have a standard form and are ready to be accepted and confirmed by each user (Guibault 2002).

Firms not only treat data as a commodity and as an asset (Almunia 2012), but also invest heavily in data processing (Prins 2006). At the same time, as some authors have observed, any company may access, collect, process or even exchange users’ personal data and, thus, make sizeable profits (Pariser 2011). For instance, in recent years, some firms have chosen to sell their clients’ personal data as a means of generating cash flow and as a way of silencing, and meeting claims of creditors (Enos 2001; Prins 2006). In other occasions, databases that contained personal data changed hands and ownership during merger acquisitions or other strategic company movements (Gauthronet 2001). Moreover, some companies decided to offer higher quality services to consumers, on condition that they would accept further processing of their data. Google, as detailed in their privacy policy, took the initiative to offer bigger storage space for its e-mail service (Gmail)

in exchange for projecting more personalized advertisements that are targeted after having scanned keywords in users' e-mails.

Given these practices, this paper looks into new "cloud technologies" that enable service providers to collect a huge volume of personal data and poses an "awkward" scenario to detect some uncertainties that could, under current laws, occur in the near future. Moreover, as such data is stored in massive databases, we discuss whether personal data could be protected by laws that govern Intellectual Property. Concluding that such intellectual property rules should, however, not apply to personal data, we highlight the role that the data subject needs to play with regard to the processing of her information.

2. ALL DATA IN THE CLOUDS

Cloud Computing technologies enable users to store multiple items of information in the "clouds". Firms may collect a huge volume of data not only on consumers, but also on third parties. Indeed, by using services, such as Google Drive, DropBox or Mediafire, a consumer may upload "in the clouds" her personal files, e.g. a photo, or third parties' items of information, e.g. an image downloaded from the Web. Such activities constitute an ordinary practice to ensure preservation and successful management of people's files. Indeed, many consumers use cloud computing services to avoid potential deletion of their files, due to e.g. malware/virus attacks, or wastage of their personal computers' storage space or use of external devices (usb sticks etc). For instance, lawyers use Cloud, where they can upload and further process e.g. all documents/applications, which they have produced since the beginning of their career. Moreover, physicians may use such technologies so as to store their files that may include their patients' sensitive data. Furthermore, minors are able to, and do, use the Cloud to store their favorite songs or movies or personal images and videos.

These storage technologies are controlled not by the user but, by the company/provider (Reed 2010). Information constitutes a major source of income for firms, and so, all items of information, stored in the Cloud, need to be kept at a secure place. Companies have "built" huge data centers ("server centers" or "storage facilities"), the function and the preservation of which may not only cost millions of euro but also damage the environment (Morozov 2011). This means that Cloud Computing technologies refer not only to applications provided as services, but also to hardware or software, which is located in datacenters and which enables consumers to use such technologies and companies to process data. Moreover, "Cloud" refers only to this datacenter's hardware and software (Armbrust 2009).

However, in addition to data storage, Cloud Computing providers may offer a wide variety of other services. For instance, Google provides not only Google Drive, i.e. data-storage-service, but also Gmail, i.e. e-mail-service, or Google Docs, i.e. word-processing-service (Yoo 2011; Lametti 2012). Hence, it is in these datacenters, and not in "some cloud", where a huge volume of data is stored. The latter includes any item of information that has been uploaded while using e.g. DropBox or Google Drive, or all files that users share in multiple platforms, such as YouTube videos, or data contained in e-mails or, generally, all information that anyone may upload on the Internet while using Cloud Computing services. Thus, as some have put it, Cloud Computing has turned our work, finances, health and relationships into invisible data, centralized in datacenters (Cook, Van Horne 2011).

In other words, firms possess innumerable important files and data, which may originate from the private messages that a minor sent to her boyfriend, or the location where they met on one particular day or at one certain time/minute/second, images or videos from a family's summer holidays, claims of a person accused of rape or his ID data, or all evidence that he is about to provide to the court, or a medical clinic's database that contains personal items of information relating to patients (Powles, Hodson 2017). All this data has independent value, it is collected "*en masse*" and on a daily basis, and may cover people's full personal, ordinary or professional life. So, digital firms have probably found their way to safeguard their financial interests by processing information that anybody produces just by living her life (Powles, Hodson 2017).

This processing is legal: Consumers "generously tick" boxes, and, thus, validly give their consent to the processing of their personal data (GDPR, recital 32). But could personal data, under current laws, be considered as an "asset" enclosed in firms' repositories?

3. AN EXAMPLE

Let us assume that some “parents of the past” were spending their days with their newborn child in a medical clinic. Let us further assume that an “itinerant trader” approached these parents so as to promote a company’s “clothing accessories for babies”. The trader would offer some clothes and accessories for free and would promise the parents that every year, on their child’s birthday, the company would send, free of charge, clothing and accessories. All these would be offered for free, on condition that the parents would complete the relevant form with some basic data (such as their names, address etc). In the offline environment, such an offer would be accepted not only as the trader’s and the company’s kind deed, but also as a welcome proposal. There would be no reasonable grounds for suspecting that the company had further objectives. Namely, in the worst case scenario, this offer might be understood as an advertising campaign, which, however, would not necessarily negate the welcoming of the proposal.

Coming to digital environment, parents of near future might be enjoying their special moments with their newborn kid in a medical clinic. They might also receive an e-mail at their smart phones. This e-mail would have been sent by a firm as an expression of its best wishes. The firm would have been aware of the exact date and time of birth after having processed data that these parents would have generously provided in e.g. social media or after having accepted multiple services. Namely, the father could have published a post declaring he is on his way to the clinic and the mother could have searched for price comparison sites with regard to baby clothing and accessories at several engines, or could have ordered such items at specific dates, or could have arranged for an appointment at a digital company for the newborn baby’s first health tests. Besides, companies, whose websites parents could have visited, would very likely exchange parents’ data with the medical clinic of above mentioned example, where child’s birth would have taken place. Continuing the hypothetical “near future example”, the company that would have sent the greeting e-mail might offer the parents some accessories, free of charge, and might promise baby’s annual birthday gift, on the condition that the parents would accept “terms of use and privacy policy”. Let us also assume that three years would have passed, during which parents would have been receiving the “annual birthday gift” against payment with their data. Finally, let us assume that the baby would most unfortunately suffer from an incurable illness and dies at the early age of 3.5.

No one could guarantee that these “near future parents” would not continue to receive each year and for an indeterminate period of time that “annual birthday gift” that would not let them forget the tragic event. This could not be guaranteed, because the parents would have unreservedly accepted terms of use, which they would have never, under present-day practices, really read. Thus, no one could know what kind of commitments they would have made. Moreover, no one could ensure that memorial photographs of parents and their -previously living- baby, which might have been published in social media, would be deleted and would not remain “forever-stored” in a datacenter of a firm that could perpetually exploit them by further processing them, together with all other deceased people’s data, and, thus, make perhaps very high profits.

This example indicates that any “mouse-click” on terms of use may not harm a consumer “today or tomorrow” but in the long term and in conjunction with all other “mouse-clicks” she has undertaken. The photograph post might seem unimportant or harmless, but the Internet, contrary to humans, seems to be immortal. Hence, the photograph may accompany the user’s profile forever. Even if relatives and friends apply, so as to delete the profile, all digital personal data of the deceased will be “inherited” by the one and only “digital-heir”, i.e. the firm that collected data and stored it in its datacenter.

4. NOT ANOTHER *SUI GENERIS* RIGHT

Companies behave as owners of personal data. This has been condemned by some authors, who speak of theft of “Humanistic Property” (Mann 2000). Personal data is intangible, the right to the protection of personal data is fundamental (CFREU art. 8), so one would assume that these items of information could not be treated as a “physical asset”, nor could they be subjected to rules that govern tangible property (Prins 2006; Laudon 1996). But could they be protected as contents of *sui generis* protected databases (Samuelson 2000)?

Under Article 1(2) of Directive 96/9/EC¹, a ‘database’ shall mean a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means. In accordance with Directive 96/9/EC both original and unoriginal databases shall be protected.

Databases which, by reason of the selection or arrangement of their contents, constitute the author’s own intellectual creation shall be protected as such by copyright (Article 3(1) of Directive 96/9/EC). In such case, copyright protection shall not extend to their contents (Article 3(2) of Directive 96/9/EC). On the other hand, with regard to unoriginal databases, the maker of a database, where there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents, has the right to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database (Article 7(1) of Directive 96/9/EC). Hence, on the basis of qualitative or quantitative investment unoriginal databases and their contents may be protected with a *sui generis* right that expires fifteen years from the year following the date of completion (Article 10(1) of Directive 96/9/EC). Furthermore, any substantial change to the contents of a database shall qualify the database for its own term of protection (Article 10(3) of Directive 96/9/EC).

In other words, under Directive 96/9/EC, protection was extended to contents of databases, such as simple facts, data or elements that used to belong in the public domain (Boyle 2003; Frosio 2012; Davison 2006). *Sui generis* protection may only last for a fifteen-year-period, albeit, given the potential of renewal of this period after any substantial change to the content, which is indeed constantly being updated, this protection is for 15 years or eternity, whichever is longer (Davison 2006).

Firms do invest heavily in data processing and, hence, one could argue that the standard of “qualitative or quantitative investment” could be met. Thus, it could be claimed that unoriginal “personal data databases” could be protected by *sui generis* rights. However, data stored in datacenters may constitute a “collection”, albeit, given opaque procedures of its collection, storage and processing it is not possible to know whether such data is arranged in datacenters “*in a systematic or methodical way*”, so as to fulfill the above database definition. Besides, even if personal data were arranged in such a systematic or methodical way, it is arguably not individually accessible by electronic or other means, as each firm owns its own “secret repository of personal information”. Companies may provide access or exchange data, not with individuals but, with other enterprises, but one perhaps cannot really speak of these databases as collections of data “individually accessible”.

Moreover, Directive 96/9/EC speaks of a “right for the maker of a database” to prevent extraction or re-utilization of the whole or of a substantial part of the contents of the database (Article 7(1) of Directive 96/9/EC). While the objective of this Directive is to protect the rights of the maker of the database and to secure the maker’s remuneration (Recital (48) of Directive 96/9/EC), if “property” rights were granted with regard to personal data, the rightholders in this case would be the data-subjects and not the makers of the databases. Furthermore, in case of a hypothetical extension of *sui generis* rights to personal data, one would demand the protection of each item of information and not the protection of such data as “*the whole or a substantial part*” of the content of the database (Huse 2005).

In any case, the probable inability to apply Intellectual Property rules, if “*sui generis*” rules are considered as such (Bottis 2014), to personal data is obvious for one more reason. Laws that govern Intellectual Property aim, amongst others, to offer incentives to create works and to promote arts and sciences. If such rights were established, with regard to personal data, we would construct an environment, where incentives to produce or collect new personal data would be offered, so that the latter would be exploited as database contents (Lemley 2000). In such an environment, a dynamic competition would take place among firms as to which would collect and manage “more and of better quality personal data”. However, such a competition would be a race to the bottom, because the purpose of this protection would be to promote processing practices (Mayer-Shönberger 2010; Prins 2006), not for the benefit of data subjects but, in favor of database makers’ interests. Besides, the GDPR promotes the principle of data minimization, under which personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (Article 5(1)(c) of GDPR).

¹ See Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, *infra* referred to as Directive 96/9/EC.

Contrary to “works”, which are products of human mind and demand some kind of “effort/action”, personal data are created unconsciously (Samuelson 2000) and without any effort (Mann 2000). Also, the right to the protection of personal data should not be treated as an intellectual property right, because the right to data protection, unlike IP rights, is not an absolute right (Recital (4) of GDPR).

5. CONCLUSION

In this new environment of massive collection of personal data, any user may provide multiple sensitive items of information, while accepting “free” digital services (Pariser 2011). Subscription to a platform, such as a hotel-booking one, may demand scanning user’s physical id, registering phone number or any other data that the company wishes to collect during either subscription, when user provides her data, or use of a service, when user does not provide her data, but the latter is automatically provided².

This “proprietary”, at least in practice, status of personal data leaves the consumer exposed to the companies’ wishes. Although the subjection of personal data to Intellectual Property rules does not seem possible, firms do treat such data as some kind of property. Indeed, under EU regime and Directive 96/9/EC, database makers may do so, since, as some authors have aptly put it, for the first time in intellectual property history rights were established on ideas and facts, such as our names or our streets’ names.

If the only thing that matters today is making profits out of commercial exploitation of any kind of tangible or intangible good, let it be noted that the protection of, consumer’s rights³ and interests (Whish & Bailey 2015) constitutes the fundamental principle that governs all stages of economic exploitation of any commercialized good. The parents of our example (under 3) would never have accepted those “free annual gifts”, if someone had made known what “future-parents” would not have read in terms of use that they would have clicked, i.e. that accepting such an offer would mean that one or multiple firms could forever make sums out of commercial exploitation of all data that could be filled at the relevant “physical form” and that would correspond to digital data that a present-day user may accept to share. Such data would equal to all offline family photographs, all offline personal discussions, all traditional personal letters that these parents would have sent to each other whilst being away, all addresses where they would have lived during their lives, all professional activities they would have pursued, all spatial and temporal information with regard to any “offline” ride they would have taken, all transactions they would have made, all illnesses from which they would have suffered, or all names of people to whom they would have spoken. If all these items of information, in their digital form, are a commodity, then the one to decide whether they shall be processed is their subject, not their collector.

REFERENCES

- Almunia Joaquín, 2012. Speech (Nov. 26, 2012) “Competition and personal data protection”.
- Armbrust Michael et al., 2009. “*Above the Clouds: A Berkeley View of Cloud Computing*”. Technical Report No. UCB/EECS-2009-28. Electrical Engineering and Computer Sciences University of California at Berkeley.
- Bottis Maria, 2014. Law and information: a “love-hate” relationship, in *The History of Information, from Papyrus to the Electronic Document*. Maria Bottis (Ed.). Nomiki Vivliothiki. Greece. pp. 141-152.
- Boyle James, 2003, Foreword: The Opposite of Property?. *Law and Contemporary Problems*. Volume 66 Winter/Spring, Numbers 1 & 2, Duke University School of Law.
- Broes Stefanie et al., 2017. Sharing human samples and patient data: Opening Pandora’s box, *Journal of Cancer Policy*, Vol. 13, pp. 65–69. Available at [http://www.journalcancerpolicy.net/article/S2213-5383\(17\)30034-6/fulltext](http://www.journalcancerpolicy.net/article/S2213-5383(17)30034-6/fulltext).
- Cook Gary & Jodie Van Horne, 2011. *How dirty is your Data? A look at the energy choices that power cloud computing*, Greenpeace International, Amsterdam, The Netherlands.

² Persistent cookies remain on the device after the user closes her browser and may be used again the next time she accesses the platform. Moreover, “web beacons” technologies let the company know whether the user has opened a certain message or accessed a certain link. See e.g. Airbnb’s policies: https://www.airbnb.gr/terms/privacy_policy - https://www.airbnb.gr/terms/cookie_policy.

³ See modern EU consumer policy: https://europa.eu/european-union/topics/consumers_en.

- Cradock Emma, Stalla-Bourdillon Sophie, Millard David, 2017. Nobody puts data in a corner? Why a new approach to categorising personal data is required for the obligation to inform, *Computer Law & Security Review*, vol. 33, issue 2, pp. 142–158.
- Davison Mark, 2006. Database Protection: The Commodification of Information, in *The Future of the Public Domain: Identifying the Commons in Information Law*, L. Guibault and P.B. Hugenholtz (eds), Kluwer Law International, The Netherlands, pp. 167–189.
- Edwards Lilian & Harbinja Edina, 2013. Protecting Post-Mortem Privacy: Reconsidering The Privacy Interests Of The Deceased In a Digital World, *Cardozo Arts & Entertainment Law Journal*, Vol. 32, No. 1. Available at SSRN: <https://ssrn.com/abstract=2267388> or <http://dx.doi.org/10.2139/ssrn.2267388>.
- Enos L., 2001. Deal Afoot to Destroy Toysmart Database, *E-commerce Times*, January 10, 2001. Available at <https://www.ecommerce-times.com/story/6607.html>.
- Frosio Giancarlo, 2012. Communia and the European Public Domain Project: A Politics of the Public Domain, in *The Digital Public Domain: Foundations for an Open Culture*. Edited by Melanie Dulong de Rosnay & Juan Carlos De Martin. Open Book Publishers CIC Ltd. United Kingdom.
- Gauthronet S., 2001. The Future of Personal Data in the Framework of Company Reorganisations, *23rd International Conference of Data Protection Commissioners*, Paris, September 2001.
- Guibault L., 2006. Wrapping Information in Contract: How does it affect the Public Domain, in *The Future of the Public Domain: Identifying the Commons in Information Law*, L. Guibault and P.B. Hugenholtz (eds). Kluwer Law International. The Netherlands. pp. 87-104.
- Guibault L., 2002. *Copyright Limitations and Contracts: An Analysis of the Contractual Over-ridability of Limitations on Copyright*, The Hague, London, Boston, Kluwer Law International.
- Huse C. Charles, 2005. Database Protection in Theory and in Practice: Three recent cases, *Berkeley Technology Law Journal*, Annual Review 2005, Vol. 20, Issue 1, pp. 23-45.
- Knoppers Bartha Maria & Thorogood Adrian Mark, 2017. Ethics and Big Data in Health, Big Data Acquisition and Analysis, *Current Opinion in Systems Biology*, vol. 4, pp. 53-57.
- Lametti David, 2012. The Cloud: Boundless Digital Potential or Enclosure 3.0? Available at SSRN: <https://ssrn.com/abstract=2077742> or <http://dx.doi.org/10.2139/ssrn.2077742>.
- Laudon Kenneth, 1996. Markets and Privacy. *Communications of the ACM*, Volume 39 Issue 9, pp. 92-104.
- Lemley M.A., 2000. Private Property, *Stanford Law Review*, vol. 52, pp. 1545-1557.
- Mann Steve, 2000. Computer Architectures for Protection of Personal Informatic Property: Putting Pirates, Pigs, and Rapists in Perspective, *First Monday*, vol. 5, no 7.
- Manovich L., 2011. *Trending: The Promises and the Challenges of Big Social Data, Debates in the Digital Humanities*, M. K. Gold (ed.), The University of Minnesota Press.
- Mayer-Shönberger Viktor, 2010. Beyond Privacy Beyond Rights – Toward a “System” Theory of Information Governance, *CALIFORNIA LAW REVIEW*, Vol. 98, Issue 6, pp. 1853-1885.
- Morozov Evgeny, 2011. *The Net Delusion, The Dark Side of Internet Freedom*, Public Affairs, USA.
- Pariser Eli, 2011. *The Filter Bubble, How the New Personalized Web Is Changing What We Read and How We Think*, Penguin Group, USA.
- Pasquale Frank, 2015. *The Black Box Society, The Secret Algorithms that Control Money and Information*, Harvard University Press, 2015.
- Powles Julia & Hodson Hal, 2017. Google DeepMind and healthcare in an age of algorithms. *Health and Technology*, Vol. 7, Issue 4, pp. 351–367.
- Prins Corien, 2006, Property and Privacy: European Perspectives and the Commodification of our Identity, in *The Future of the Public Domain: Identifying the Commons in Information Law*, L. Guibault and P.B. Hugenholtz (eds), Kluwer Law International, The Netherlands, pp. 223-258.
- Radin M.J., 1996, *Contested Commodities*, Cambridge, Harvard University Press, p. 18.
- Reed Chris, 2010, Information “Ownership” in the Cloud, Queen Mary University of London, School of Law, *Legal Studies Research Paper*, No. 45/2010.
- Samuelson Pamela, 2000, Privacy As Intellectual Property? *Stanford Law Review*, Vol. 52, pp. 1125-1173.
- Solove J. Daniel, 2013, Privacy Self-Management and the Consent Dilemma, *Harvard Law Review*, vol. 126, p. 1880-1903.
- Whish Richard & Bailey David, 2015, *Competition Law*, 8th Edition, Oxford University Press. United Kingdom.
- Yeh Chih-Liang, 2017, Pursuing consumer empowerment in the age of big data: A comprehensive regulatory framework for data brokers, *Telecommunications Policy*, <https://doi.org/10.1016/j.telpol.2017.12.001>.
- Yoo Christopher, 2011, Cloud Computing: Architectural and Policy Implications, *Rev Ind Organ*, 38: 405. <https://doi.org/10.1007/s11151-011-9295-7>.

SEMANTIC WEB APPROACHES FOR BIG DATA INTEGRATION

Danilo Hasse¹, Fernando A. Ostuni Gauthier¹, Carlos Roberto De Rolt² and Gisiela Hasse Klein²

¹*Departamento de Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina - UFSC
Campus Universitário Reitor João David Ferreira Lima – Florianópolis – Santa Catarina, Brasil*

²*Centro de Ciências da Administração e Socioeconômicas, Universidade do Estado de Santa Catarina - UDESC
Campus Av. Madre Benvenuta, 2037 - Florianópolis – Santa Catarina, Brasil*

ABSTRACT

Nowadays modern enterprises have to face the problem of handing and sharing huge amount of data, expressed in different incompatible formats and also scattered in distributed sources, aiming at improving the quality and effectiveness of their processes. These requirements are becoming more and more relevant to meet the new challenges posed by massive spread of Big Data, mobile technologies and Internet of Things (IoT). Under these conditions, as the information exchanges are complex, it is necessary to adopt a new model of collaboration between the various involved sources, regardless of the platform on which they run. Among the methodologies designed to solve these incompatibilities, semantic interoperability represents a consolidated conceptual paradigm for joining heterogeneous distributed data sources into a synergistic view. The advantage of Ontology-based Big Data Integration based on semantic technologies is enabling enterprises to integrate and manage data with a high degree of automation and make decisions during the integration process. This paper investigates these issues related to the semantic web as a key component for data integration over different data models because the characteristics of semantic technologies in creating and representing complex information and relationships between concepts can help to interpret information by identifying the corresponding context. The challenges of Big Data management continue growing currently and require innumerable efforts to solve the interoperability across a diversity of data types, platforms and patterns of systems as well as the exponential demand of society for new applications, quality of services and velocity of information.

KEYWORDS

Big Data, Semantic Web, Ontologies, Data Integration, NoSQL, OBDI

1. INTRODUCTION

Nowadays modern enterprises have to face the problem of handing and sharing huge amount of data, expressed in different incompatible formats and also scattered in distributed sources, aiming at improving the quality and effectiveness of their processes. Under these conditions, as the information exchanges are complex, it is necessary to adopt a new model of collaboration between the various involved sources, regardless of the platform on which they run. Among the methodologies designed to solve these incompatibilities, semantic interoperability represents a consolidated conceptual paradigm for joining heterogeneous distributed data sources into a synergistic view that facilitates the exchange of information and the use of the information that has been exchanged.

This is a Big Data era, Gartner (2014) identified three dimensions of big data, on which business and information technology leaders have to focus: information volume, variety and velocity. These requirements are becoming more and more relevant to meet the new challenges posed by massive spread of Big Data, mobile technologies and Internet of Things (IoT) also as part of enterprises. In this context the development of robust interoperable tools is also required to combine, integrate and process large amounts of live streaming data generated from highly distributed devices.

More recently, NoSQL systems are viable solutions for applications that require scalable data repositories, which can easily scale out over multiple servers and support flexible data model and storage schema. In general, NoSQL systems are recommended for newly developed applications, but not for migrations of existing applications which are written on top of traditional relational database systems. The

current NoSQL trend is motivated by applications stemming mostly of the Web 2.0 domain. Some of these applications have storage requirements which exceed capacities and possibilities of relational databases.

Since the relational data model is considered as not suitable for certain new applications, the integration of different data models offered by NoSQL systems and operating in an environment with relational legacy systems remains a challenge.

Ontology-based Big Data Integration based on semantic technologies enable enterprises to integrate and manage data with a high degree of automation. Also, less manual mistakes can occur because this approach minimizes the manual interference in the integration process. The advantage of Ontology-based Big Data Integration is that the user/enterprise can make decisions during the integration process. The ontology will support the user by limiting the possibilities the user has to choose from mistakes. Hence, there will also be a decrease in the total number of manual decisions the user will have to make during the integration process. This will be enabled by a pre-selection automatism based on ontology alignment, and will reduce time and cost for integration process significantly (Eine et al. 2017).

This paper investigates these issues related to the semantic web as a key component for data integration over different data models because the characteristics of semantic technologies in creating and representing complex information and relationships between concepts can help to interpret information by identifying the corresponding context. Semantic technologies can make it easier to understand the meaning and purpose of data (e.g., symbols, words etc.) and complex concepts, as well as share knowledge for humans and machines.

2. SEMANTIC WEB DATA INTEGRATION CHALLENGES

One of the biggest current research challenges in the human-computer interaction and information retrieval is to provide users with intuitive access to the growing amount of data present in different database management systems (DBMS). These data express real world facts, attributes and interrelationships. The origin of data is their history which covers source collection, processing technologies, executors, etc. Semantic technologies in data integration solutions allow representing relationships of data definition area, relating data using datasets and identifying relationships for new associations. The integration of data is one of the most complicated tasks which need to be addressed by IT researchers. Despite the critical importance, the current approaches to semantic interoperability of heterogeneous databases have not been sufficiently effective (Uzdanaviciute and Butleris 2011).

Data integration is concerned with unifying data that share some common semantics but originate from unrelated sources. Necessarily, when we work on data integration, we must take into account a more important and complex concept called “heterogeneity”. Heterogeneity might be classified into four categories: (1) structural heterogeneity, involving different data models; (2) syntactical heterogeneity, involving different languages and data representations; (3) systemic heterogeneity, involving hardware and operating systems; and (4) semantics heterogeneity, involving different concepts and their interpretations (Buccella et al. 2005).

Ontology based semantic integration systems for Relational Database Management Systems (RDBMS) already exist that satisfies the above requirements. Many commercial systems are operational based on the above technique, tools and solutions for the above approach are very mature (Lake and Crowther 2013). The same system architecture used for RDBMS integration system can be used for NoSQL data stores like for example HBase, MongoDB, Cassandra, etc, i.e., in different data models with different implementations (e.g., Uzdanaviciute, and Butleris 2011; Huang et al. 2015).

Some surveys on ontology-based systems for data integration can be found in literature. For example, in Wache et al. (2001) authors provide a survey especially focusing on the ontology use, mappings and tools, but without providing a comparison of other elements, such as query resolution or architectural issues. Another different survey can be found in Corcho and Gomez-Perez (2000), but only languages to represent ontologies are compared in this case. Another important work Buccella et al. (2005) shows the state-of-the-art in ontology-based systems for data integration by describing seven systems – SIMS, OBSERVER, DOME, KRAFT, Carnot, InfoSleuth and COIN.

3. ONTOLOGY-BASED DATA INTEGRATION (OBDI)

Ontology-based data integration (OBDI) is an effective method to cope with the heterogeneous data. This solution is based on the idea of decoupling information semantics from the data sources. Moreover, ontologies support dynamic domains better. For this reason, it is necessary to analyze data source elements: data, schema, schema elements and content, values, entities and attributes, query result classes. It is known that ontology-based search system gives the user more meaningful query results than the normal search system, which queries data with syntactic parameters (Uzdanaviciute and Butleris 2011).

NoSQL databases are suitable for network traffic loads experienced by today's websites. The important attributes that make them a favorable choice are inherent distributed nature, horizontal scalability and flexibility in data models. Hence, more data is now getting handled through NoSQL databases and sensible information extraction from these data stores is becoming a requirement. Information extraction may require sourcing data from multiple data sources, establishing relationship among them and querying across these data sources together (Kiran and Vijayakumar 2014). "Not only SQL" (NoSQL) platforms are no longer questioned today. Driven by major Web companies, they have been developed to meet requirements of novel applications, hardly available in relational databases (RDB), such as a flexible schema, high throughput, high availability and horizontal elasticity. They are emerging as an efficient alternative to relational database management systems (RDBMS) at the core of many applications in the context of big data, but also they are increasingly used as a generic-purpose database in many domains (e.g., social networks, biotechnology, smart cities, etc) (Michel et al. 2016). Furthermore, there are still many challenges for NoSQL platforms, particularly concerning data query mechanisms, data searching, security, and data migration among relational database systems and NoSQL systems (Kurpanik and Pańkowska 2015).

Some research papers today argue that the integration of these features needs to consider the semantics of the elements of the application domain. This could be a major breakthrough for both NoSQL stores and the Semantic community since RDBMS is not really reactive in integrating semantics (Curé et al. 2012). Thus, this overwhelming success makes NoSQL databases a natural candidate for RDF-based data integration systems, and potential significant contributors to feed the Web of Data (Michel et al. 2016).

The emergent context of semantic web integration approach (e.g., OBDA, RBDA) indicates the integration of various semi-structured data, emergence of very large and semi-structured knowledge bases and emergence of databases with different NoSQL data models.

Non-relational database management systems are classified according to the data model defining the logical organization of data. Columnar database, document database, key-value pair database and graph database models constitute the basic models group. For example, Graph databases store data entities and connections between them as nodes and edges. Document databases store text, media, and JSON (Java Script Object Notation) or XML data. Key-value pair databases store data as simple key-value pairs. The keys are unique and do not have any foreign keys or constraints. Column-oriented databases have a huge number of columns for each tuple. Each column also has a column key. The new models group covers multi-model, grid model, multidimensional model, multi-value model, network model and others, which cannot be explicitly classified (Kirti and Maan 2015).

Some concepts are related with the OBDI such as Ontology-Based Data Model (OBDM) and especially Ontology-Based Data Access (OBDA) that are introduced ahead.

3.1 Ontology-Based Data Model (OBDM)

The term "ontology" refers to a machine-readable representation of knowledge, particularly for automated inference. Ontology is a data model which consists of these parts: classes, properties and relationships between them. The power of ontologies lies in the ability to represent relationships between the classes. The main benefit of using the ontology-based model is its runtime interpretation. One of the major advantages of the ontology model is an assumption of open-world. The reason for the popularity – clearly interpreted dissemination of knowledge between people and applications. Moreover, ontology supports the integration task as it describes the exact content and semantics of these data sources more explicitly.

The integrated data takes into consideration the four following aspects designed in the composition model: a data model, architecture, the content and the representation. The data model of a particular data source definition. The model must allow extension with new data, retrieval of important, highest-quality, semantically meaningful information, and there use of data (Uzdanaviciute and Butleris 2011).

The architecture which is the core of integration and has to perform the high level autonomy to data sources. The architecture must provide semantic interoperability for the systems. It must allow the system architect to manage the development of data collections when data sources have different formats (text files, XML schemas, relational models) and to re-aggregate the application (Uzdanaviciute and Butleris 2011).

A neutral representation abstracts specific syntax; therefore, all the structured and semi-structured data sources first need to be expressed in a neutral format. A set of content data elements must be able to receive high-quality, semantically meaningful information. The content is heavily affected by the semantic conflict types to be resolved. The content, i.e., the meaning of the information that is interchanged, must be understood. Data and relations have to be visualized and represented in the best, most appropriate way. It follows that each representation must bind a single expression to a single meaning using the Resource Description Framework (RDF) language (Uzdanaviciute and Butleris 2011).

3.2 Ontology-Based Data Access (OBDA)

Ontology-Based Data Access (OBDA) may be a good fit in this direction since it aims to represent the concepts and properties of a domain with a formalized ontology. OBDA provides a semantic conceptual schema over a repository of data and, due to its logical formalism, it is likely to support formal analysis, optimization and reasoning (Curé et al. 2012).

OBDA is a new data management paradigm that exploits the semantic knowledge represented in ontologies when querying data stored in a database (Curé et al. 2012). OBDA is a well-established paradigm for querying incomplete data sources while taking into account knowledge provided by domain ontology (Poggi et al., 2008). Today, the main applications of OBDA can be found in data integration as well as in querying the Semantic Web (Mugnier et al. 2016).

The interest of OBDA is to allow the users to ask queries on high-level ontology vocabularies and to delegate to algorithms (1) the reformulation of these high-level queries into a set of low-level databases queries, (2) the efficient computation of their answers by native data management systems in which data is stored and indexed, and (3) the combination of these answers in order to obtain the final answers to the users' query (Mugnier et al. 2016). The advantage of OBDA is that, since the query reformulation step is independent of the data, ontology-mediated query answering has the same data complexity as the query engines equipping the underlying native data storage systems, and can benefit from the many low-level optimizations making them efficient and scalable (Mugnier et al. 2016).

OBDA has been studied so far for relational structures and deployed on top of relational databases. The database queries used in OBDA are (unions of) conjunctive relational queries, while the ontologies are specified in either a description logic (e.g., the lightweight DL-Lite (Calvanese et al. 2007), or more generally, a suitable fragment of first-order logic (e.g., Gottlob and Pieris 2012) and existential rules (Baget et al. 2011). Within this framework, decidability and complexity results have been obtained for ontology-mediated query answering, and many algorithms have been designed and implemented (Mugnier et al. 2016).

This paradigm can be used to query data sources that are non-relational and is still an open question in the research efforts. One possibility to deal with non-relational data sources is to define mappings for translating them into relational structures, and then use the classic OBDA framework as it is. However, this approach would induce a significant performance degrade as it would add a step for converting the data using the mappings and, most importantly, it would make impossible to take advantage of the low level query optimizations provided by native systems. This can be particularly acute for NoSQL systems, like key-value, that have been specifically designed to scale when dealing with very large collections of data (Mugnier et al. 2016). In this way, the solution to the problem of diverse NoSQL systems integration are possible using ontologies and the Ontology-Based Data Access (OBDA) approach developing simple and effective query methods.

In the literature, some research works are implementing OBDA to integrate NoSQL databases. Botoeva et al. (2016b) proposed the use of NoSQL in OBDA applications. Their work is an extension of Ontop (Bagosi et al. (2014), which is an OBDA system for relational databases. The new proposed architecture was tested using a document-oriented MongoDB database.

In related work, Botoeva et al. (2016a) presented a formal evaluation of a subset of data access queries available on MongoDB. This evaluation has shown that it is very hard to build a completely generic framework that is able to query any kind of NoSQL DBMS. Unlike relational databases that use a common query language (SQL), NoSQL DBMS share few query patterns, so every NoSQL DBMS needs a specific query translator (Araujo et al. 2017).

In another interesting work, Michel et al. (2016) have built a SPARQL to MongoDB query mapping tool in order to turn a legacy database into a publicly available data source. The database should be exposed as a virtual RDF database: stored documents would be published as RDF triples (Araujo et al. 2017).

In this way, the Onto-mongo project (Araujo et al. 2017) is a working prototype that applies the query translation method, generating a set of RDF triples as a result that is then used to populate an OWL ontology.

The Onto-mongo project has the following main components: an OWL Ontology, a NoSQL database, an Access Interface, an Ontology to Database collections mapping, a SPARQL to NoSQL query translator, and a RDF export (Araujo et al. 2017).

4. OTHER SEMANTIC WEB APPROACHES FOR BIG DATA INTEGRATION

We can find in the literature other approaches proposing NoSQL data integration without OBDM and OBDA schemas but developing specific methods to address database integration using Semantic Web patterns such as RDF data model and SPARQL adapting NoSQL database for semantic data integration. Some examples are: i) the Transformed Table. To improve the scalability and the performance of triple query, which is called SPARQL query processing, MapReduce programming model and NoSQL database system such as HBase are well-known solutions for large scale data processing. However, in general case, the subject of a triple is regarded as RowKey in the table. In some queries, finding matched triple patterns is a time-consuming job. Therefore, was designed in another table with different storage schema called Transformed Table (Huang et al. 2015) to reduce the time cost for read operation. The experimental results show that using Transformed Table can improve the triple query performance significantly; ii) Mapping-based SPARQL Access to a MongoDB. In recent years, MongoDB has become the leader in the NoSQL market. Some methods have been proposed to translate MongoDB documents into RDF (Michel et al. 2015), or to use MongoDB as an RDF triple store. The xR2RML mapping language is an extension of the R2RML and RML addressing the mapping of a large and extensible scope of non-relational databases to RDF. Some works have been proposed to use MongoDB as an RDF triple store, and in this context they designed a method to translate SPARQL queries into MongoDB queries (Tomaszuk et al. 2010).

5. CONCLUSION

We consider that these main issues related with Big Data and heterogeneous NoSQL databases integration can be deal in the context of Ontology-Based Data Integration (OBDI) efforts and the Ontology-Based Data Access (OBDA) techniques. OBDA is a new data management paradigm that exploits the semantic knowledge represented in ontologies when querying data stored in a database. Furthermore, another techniques using RDF data model and SPARQL are possible to use directly enabling a transformation of the original NoSQL schemas in an integrate data model over platforms. The challenges of Big Data management continue growing currently and require innumerable efforts to solve the interoperability across a diversity of data types, platforms and patterns of systems as well as the exponential demand of society for new applications, quality of services and velocity of information.

ACKNOWLEDGEMENT

This research was supported by the program CAPES 3º Cronograma – Chamadas de Projetos nº 09/2014, processo nº 8745-13-7 for Carlos Roberto De Rolt and nº 88887.122829/2016-00 for Danilo Hasse.

REFERENCES

- Araujo, T.H.D., et al, 2017. OntoMongo - Ontology-Based Data Access for NoSQL. *Proceedings of the IX Seminar on Ontology Research in Brazil and I Doctoral and Masters Consortium on Ontologies*. Brasília, Brazil.
- Baget, J.F. et al, 2011. On Rules with Existential Variables: Walking the Decidability Line. In *Journal Artificial Intelligence*. Vol.175, No.10, pp.1620–1654.
- Bagosi et al., 2014. The Ontop Framework for Ontology Based Data Access. *Proceedings 8th Chinese Semantic Web and Web Science Conference*, pp.67–77.
- Botoeva et al., 2016a. *A Formal Presentation of MongoDB (Extended Version)*. CoRR Technical Report abs/1603.09291, arXiv.org e-Print archive.
- Botoeva et al., 2016b. OBDA Beyond Relational DBs: A Study for MongoDB. *Proceedings of the 29th Int. Workshop on Description Logics*, (DL 2016), 1.
- Buccella, A. et al, 2005. Ontology-Based Data Integration Methods: A Framework for Comparison. In *Revista Colombiana de Computación* Vol.6, No.1.
- Calvanese, D., et al, 2007. Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family. In *Journal Autom. Reasoning*. Vol.39, No.3, pp.385–429.
- Corcho, O. and Gomez-Perez, A., 2000. Evaluating knowledge representation and reasoning capabilities of ontology specification languages. *Proceedings of the ECAI 2000 Workshop on Applications of Ontologies and Problem-Solving Methods*. Berlin.
- Curé, O. et al, 2012. On The Potential Integration of an Ontology-Based Data Access Approach in NoSQL Stores. *Proceedings Third International Conference on Emerging Intelligent Data and Web Technologies*. Bucharest, Romania, pp.166-173.
- Eine, B., et al, 2017. Ontology-Based Big Data Management. *Systems In International Open Access Journal on Systems Engineering and Systems Management*, Vol.5, No. 3, 45.
- Gartner, 2014. IT Glossary Big Data [Online]. Available: <http://www.gartner.com/it-glossary/big-data/>
- Gottlob, G. and Pieris, A., 2012. Towards more Expressive Ontology Languages: The Query Answering Problem. In *Journal Artificial Intelligence*. Vol.193, pp.87–128.
- Huang, S.H., et al, 2015. Efficient and Scalable SPARQL Query Processing with Transformed Table. *Proceedings IEEE Wireless Communications and Networking Conference (WCNC) - Workshop - Next Generation WiFi Techonology*. New Orleans, USA, pp.103-106.
- Kiran, V.K. and Vijayakumar, R., 2014. Ontology Based Data Integration of NoSQL Datastores. *Proceedings 9th International Conference on Industrial and Information Systems (ICIIS)*. Gwalior, India, pp.1-6.
- Kirti and Maan P., 2015, Database for Unstructured, Semistructured data-NoSQL. In *International Journal of Advanced Research in Computer Engineering & Technology*. Vol. 4, Issue 2, pp.466-469.
- Kurpanik, J. and Pańkowska, M., 2015. *NoSQL Problem Literature Review*. Uniwersytetu Ekonomicznego w Katowicach, ISSN 2083-8611 Nr. 234.
- Lake P. and Crowther P., 2013. *Concise Guide to Databases, A Practical Introduction*. Springer-Verlag. London.
- Michel, F. et al, 2015. Translation of Relational and Non-Relational Databases into RDF with xR2RML. *Proceedings WebIST2015 Conf.*, Lisbon, Portugal, pp.443–454.
- Michel F., et al, 2016. *Mapping-based SPARQL Access to a MongoDB Database*. Research Report CNRS.
- Mugnier, M.L. et al, 2016. Ontology-Mediated Queries for NOSQL Databases. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*. Phoenix, USA, pp.1051-1057.
- Poggi, A. et al, 2008. Linking Data to Ontologies. In *Journal Data Semantics*. Vol.10 pp.133–173.
- Tomaszuk, D., et al, 2010. Document-oriented triplestore based on RDF/JSON, In *Log. Philos. Comput. Sci.*, University of Białystok, pp. 125–140.
- Uzdanaviciute, V. and Butleris, R., 2011. Ontology-based Foundations for Data Integration. *Proceedings BUSTECH 2011: The First International Conference on Business Intelligence and Technology*. Rome, Italy, pp.34-39.
- Wache, H., et al, 2001. Ontology-based Integration of Information - A Survey of Existing Approaches. *Proceedings of IJCAI-01 Workshop: Ontologies and Information Sharing*. Seattle, WA, pp.108-117.

INVESTIGATION OF eHEALTH LITERACY AMONG PATIENTS WITH DIABETES

Sophie Huey-Ming Guo

Mackay Medical College

Mackay Medical College, Department of Nursing, New Taipei City, Taiwan

ABSTRACT

People with diabetes are challenging for caring this lifelong disease since minimize possibilities of complications determined by core capabilities of self-care. In the digital age, eHealth technology and related services offer a promising vision for assisting people with chronic diseases. The ability to use eHealth technologies and health applications, known as 'eHealth literacy', is crucial to facilitate eHealth adoption. This paper describes the investigation of eHealth literacy, behaviors of using health websites and health applications among patients with diabetes.

KEYWORDS

eHealth, mHealth, Diabetes, eHealth literacy

1. INTRODUCTION

People with diabetes have lifelong learning needs to cope with this disease in order to minimize possibilities of complications. Differences in self-care behaviors may be a key contributor to the disparity in outcomes. The disease's course is more likely toward deterioration with people who have less capability of processing health information. Since technologies have been indispensable part of our life, an attempt of using electronic resources to improve health behavior therefore enormous increase. Successful and extensive self-care behaviors often come well-learned or supported by several ways such as eHealth service and patient education programs. The electronic health tools include personal computers, tablets, web-based applications, and mobile applications. In sum, health services and information delivered or enhanced through the Internet, mobile and related technologies, have been defined as eHealth (Eysenbach 2001) and mHealth (Adibi 2015). The eHealth and mHealth tools now are more fully embedded in hospitals' workflows but less on consumer health, especially for people who need to cope their chronic diseases (Charlene, Graff et al. 2014, Capozza, Woolsey et al. 2015).

Different kinks of tools offer vary resources and functionalities, enabling operations on a board range of mobile eHealth tasks. Thus, different types of challenges arise depending on the tool and users' skills. One example, searching information about treatment options for diabetes condition requires identifying appropriate and reliable sources and assessing quality of information (Kaufman and Woodley 2011) (Chan & Kaufman, 2011). Another quantitative study reported that older adults experienced 'frustrating, overwhelming, and challenging' and but also fascinating and fun at the initially learning stage of a mobile health application (Grindrod, Li, & Gates, 2014).

Health literacy is the degree to which individuals have the ability to obtain, process, and understand health information in order to make appropriate health decisions (Neter and Brainin 2012). The concept of eHealth literacy focuses 'online' health information, consisting of eight elements: traditional literacy, health literacy, information literacy, scientific literacy, media literacy, computer literacy, mobile literacy, and Internet literacy (Norman and Skinner 2006). To facilitate their adoption, the ability to use eHealth technologies and applications, known as 'eHealth literacy', is required to study for facilitate the adoption of eHealth services to diabetes care (Jahns 2014, Ahmed 2015, Chung and Nahm 2015, Kim and Xie 2017).

2. METHODS

A self-administered questionnaire was completed by 254 patients with diabetes during a 12-month period from Jan 16, 2017, to Dec 15, 2017, in three outpatient department in Taiwan. Individuals were provided a NT100 gift card for participation. This study was approved by the institutional review board of target hospitals. Patients diagnosed with Diabetes were recruited in this study. The inclusion criterion was ages 20 - 75 years, no vision defects, and willingness to participate this study.

Demographics such as age, gender, education, health status, and duration of diabetes were collected. Knowledge and skills about Mobile and Internet were adapted from Xie (2011) was modified in the study. The eHealth Literacy questionnaire was revised by our team members according our research purposes; the original version developed by Skinner (2006) measures perceived skills and comfort with using the Internet for health information and decision making. The researchers added mobile-related items about perceived skills concerning health mobile apps for self-management. There are a total 20 items in which 8 for eHealth and 12 for mHealth, scoring on 1- to 5-point Likert scale with the following anchors: (1) strongly disagree; (2)disagree; (3)undecided; (4)agree; (5)strongly agree. Higher score indicates higher mobile eHealth literacy efficacy.

The study was approved by the Institutional Review Board of the targeted hospitals. The participants referred by metabolic physicians or certificated diabetes educators first signed the consent form and then were administered the questionnaire. Data Analysis used the software IBM SPSS 21.0. The main statistical methods were descriptive analysis, one-way ANOVA etc.

3. RESULTS

Out of 256 participants, 165(64.5%) were male and mean age was 46.05 years ($SD = 11.4$, $Min = 18$, $Max = 63$). The mean duration of diabetes was 6.55 ($SD = 6.17$, $Min = 0$, $Max = 34$) (Table 1). An average of four hours daily use smartphone ($SD = 3.798$) was found from participates response (Table 2). Sixty-four percent (164/256) had been searched online diabetes information. Only 4% (4/252) had been used health applications. One-way ANOVA found education has significant effect on different eHealth score, but not for duration of diabetes (Table 3).

Regarding evaluation of eHealth literacy, 58.0 to 79.7 agreed they has confidence to use eHealth (Table 4). Of these respondents, 72.2% agreed (10% disagreed) that they know what health resources are available on the Internet, 75.3% agreed (7.4 % disagreed) that they know where to find helpful health resources on the Internet, 79.7% agreed (3.9 disagreed) that they know how to find helpful health resources on the Internet, and 72.6 % agreed (6.7 disagreed) that they know how to use the health information I find on the Internet to help me.

Table 1. Use of the Internet to Seek Online Health Information (N = 256)

Characteristics	Frequency	Percentage (%)
Search for diabetes information		
Have	164	64.3
Not have	92	35.7
Search for health information*		
Have	23	17.6
Not have	108	82.4

* indicate missing value

Table 2. Use of Smart Devices and Health Applications (N = 256)

Characteristics	Frequency	Percentage (%)
Use of app		
Ever used	42	16.6
Use	26	10.3
No	185	73.1
Use of Health app		
Use	4	1.6
No	252	73.1
Use of smart devices (yr), M(SD)		
Smartphone (n = 239)	6.39(3.38)	0~20 (range)
Tablet (n = 93)	4.679(3.28)	0~15(range)
Daily use (hr), M(SD)		
smart-phone (n = 239)	4.465(3.798)	0~20(range)
Tablet (n = 78)	2.272(2.65)	0~12(range)

Table 3. eHealth literacy between different sociodemographic characteristics (N = 256)

Characteristics	Frequency	Percentage (%)	p
Education level			.032
①Elementary school	34	13.3	
②Senior high school	73	28.6	
③College or university	110	44.0	
④Master or PhD	36	14.1	
Duration of Diabetes (mean, SD)	Mean = 6.55 SD = 6.17	Range 0~34	> 0.05
①<1 year	50	27.0	
②1~4.9years	62	33.5	
③5~9.9 years	54	29.2	
④>10 years	19	10.3	

Table 4. Evaluation of eHealth literacy (N = 256)

Items of eHealth literacy	Respondents No. (%)		
	(1) disagree	(2) undecided	(3) agree
1.I know what health resources are available on the Internet	23 (10)	42 (16.4)	176 (72.7)
2.I know where to find helpful health resources on the Internet	19 (7.4)	39 (15.2)	183 (75.3)
3.I know how to find helpful health resources on the Internet	10 (3.9)	36(14.1)	204 (79.7)
4.I know how to use the Internet to answer my questions about health	21 (8.6)	51 (19.9)	177 (69.2)
5.I know how to use the health information I find on the Internet to help me	17 (6.7)	47 (18.4)	186 (72.6)
6.I have the skills I need to evaluate the health resources I find on the Internet	28 (11.4)	67(26.3)	153 (59.7)
7.I can tell high quality health resources from low quality health resources on the Internet	38 (14.8)	62 (24.2)	150 (58.6)
8.I feel confident in using information from the Internet to make health decisions	22 (8.6)	67 (26.2)	160 (62.5)

Note: The eHealth literacy questionnaire was originally in 5-point Likert scale in which were converted into 3-point scale
- Disagree, Undecided, Agree.

4. DISCUSSIONS AND CONCLUSION

4.1 Discussions

This paper describes the investigation of mobile eHealth literacy among patient with diabetes. Analytic results revealed that education was the important factor of eHealth literacy. Sixty percent of subjects have been used the Internet to seek diabetes information. Further evaluation of eHealth literacy, more than half of the subjects said they has confidence to use eHealth.

In contrast to eHealth survey, there are large portion of subjects have behavior of using mobile app in life but only was small portion on use of mHealth app. Benefits of using mHealth app were well established in huge number of studies. It is possible that people with diabetes have limited ability of using mHealth app.

Although eHealth and mHealth interventions offer a promising vision for assisting people with chronic diseases.

There is still a gap between mHealth app concept, delivery, and translation into health behavior change (Zhao, Freeman et al. 2016). Many factors can restrain individual's effective use of mHealth and eHealth tools, including environmental barriers, physical access barriers, and individual-level barriers (Holtz and Lauckner 2012). Some barriers to use new health technology will impede its benefits to spread widely. People with diabetes are often middle-aged or older with limited capability of computer and Internet operation, or less education. Moreover, fragmented or errant health information found online will direct people make 'unwise' health decisions. This obstacle makes it challenging for them to work well in eHealth and mHealth era while technology becoming used fast and wide in healthcare service. The numbers of potential users are continually increasing; however, they will not benefit from well-established health technologies due to lack of mobile eHealth literacy.

4.2 Implications for Future Research and Practice

Deploy a simple questionnaire mobile eHealth literacy of patients and caregivers in clinical practice may help healthcare providers understand patients' ability before starting disease education. Facilitate the adoption of mHealth app still need more efforts not only improve user interface but also user ability. Especially, patients with chronic conditions have been slow to embrace mobile health technology did which requires additional support from health professionals.

Healthcare sectors are the major stakeholder to educate and engage patients in self-management, but they more emphasize on disease care rather than education concerning health technology. Educational strategy is essential to solve the obstacle of limited mobile eHealth literacy, not only enhancing health literacy but also incorporating the ability of using technology.

4.3 Conclusion

These results highlight the need for better adoption of mHealth app in the diabetes population. Collaborations between researchers, healthcare providers, app developers, and policy makers could enhance the process of delivering and conducting evidence-based apps to improve health outcomes.

ACKNOWLEDGEMENT

The project was supported by grants from partly supported by the Ministry of Science and Technology (MOST 105-2511-S-715-001).

REFERENCES

- Adibi, S. (2015). Mobile health: a technology road map, Springer.
- Ahmed, S. (2015). mHealth Literacy: Characterizing People's Ability to Use Smartphone-based Health-related Applications, School of Information, University of Michigan.
- Capozza, K., S. Woolsey, M. Georgsson, J. Black, N. Bello, C. Lence, S. Oostema and C. North (2015). "Going Mobile With Diabetes Support: A Randomized Study of a Text Message-Based Personalized Behavioral Intervention for Type 2 Diabetes Self-Care." *Diabetes Spectrum* 28(2): 83-91.
- Charlene, C., K. Graff, J. Harris, A. McQueen, M. Smith, M. Fairchild and M. Kreuter (2014). "Evaluating diabetes mobile applications for health literate designs and functionality." *Prev Chronic Dis.* http://www.cdc.gov/pcd/2015/14_0433.
- Chung, S.-Y. and E.-S. Nahm (2015). "Testing Reliability and Validity of the eHealth Literacy Scale (eHEALS) for Older Adults Recruited Online." *Computers Informatics Nursing* 33(4): 150-156.
- Eysenbach, G. (2001). "What is e-health?" *Journal of medical Internet research* 3(2).
- Holtz, B. and C. Lauckner (2012). "Diabetes management via mobile phones: a systematic review." *Telemedicine journal and e-health: the official journal of the American Telemedicine Association* 18(3): 175.
- Jahns, R. G. (2014). "Today's diabetes apps are far away from meeting the seven best practice standards." Retrieved 12.12, 2015, from <http://research2guidance.com/todays-diabetes-apps-are-far-away-from-meeting-the-seven-best-practice-standards/>.
- Kaufman, N. D. and P. D. P. Woodley (2011). "Self-management support interventions that are clinically linked and technology enabled: can they successfully prevent and treat diabetes?" *Journal of Diabetes Science and Technology* 5(3): 798.
- Kim, H. and B. Xie (2017). "Health literacy in the eHealth era: A systematic review of the literature." *Patient Education and Counseling* 100(6): 1073-1082.
- Neter, E. and E. Brainin (2012). "eHealth Literacy: Extending the Digital Divide to the Realm of Health Information." *J Med Internet Res* 14(1): e19.
- Norman, C. D. and H. A. Skinner (2006). "eHealth Literacy: Essential Skills for Consumer Health in a Networked World." *J Med Internet Res* 8(2): e9.
- Zhao, J., B. Freeman and M. Li (2016). "Can mobile phone apps influence people's health behavior change? An evidence review." *Journal of medical Internet research* 18(11).

Reflection Papers

REFLECTIONS ON DIGITAL IDENTITY IN THE ONTOLOGICAL CYBERSPACE

Ziyed Guelmami¹ and François Nicolle²

¹*LaRA*, ²*LIRSA, France*

ABSTRACT

This essay introduces an ontological discussion of cyberspace. Challenging the objective, subjective and relational perspectives of spaces and place, the Heideggerian precognitive approach contends that “lived spaces” are constituted by an individual’s purpose, objectives and aspirations. An existential analysis is carried out to show how a different kind of spatialization is achieved by human beings in cyberspace (“cyber-spatiality”) and how it affects the way they disclose themselves online. Heidegger’s ground-breaking perspective proves itself useful to provide a deeper understanding of the way consumers experience online space.

KEYWORDS

Cyberspace, Self-Disclosure, Digital Identity

1. INTRODUCTION

Today, people dwell regularly in virtual worlds for many purposes : shopping, entertainment, social bonding, information searching or dating for instance. The nature of the space where our conscience is thrown in does not seem relevant to the way we experience reality. Navigation on the “cyberspace” has become commonplace in Western societies and changed the way we interact with the world and the way we construct our identity. The traditional perspectives of space, namely the objective, subjective and relational approaches of space, seem inadequate to explore the peculiar spatiality characterizing cyberspace, that is, how people phenomenologically relate to their environment through technology and virtual interfaces. Philosophical debates on the nature of space have been durably affected by the Cartesian perspective opposing the subject (conceptualized as a self-contained inner entity) to the object (the world “out there”). Consequently, space is often understood as an objective and immutable “container” or as a subjective representation produced by the human mind. The lack of conceptual tools to understand the elusive nature of cyberspace is especially problematic in a time of frenetic technological development, where new spaces of human interaction flourish online. Thus, an important issue emerges concerning the spatial nature of these online spaces and their role in transforming identities and communities.

2. DISCUSSING THE NEBULOUS NATURE OF CYBERSPACE

Although digital consumption has been the focus of many marketing papers, very few addressed the problem of the nature of cyberspace. There is no consensual definition of this concept neither in consumer research nor in computer sciences despite the wide use of the word. The few existing definitions are often descriptive and narrative and consequently lack scientific precision (Cicognani, 1998). Therefore, the word is becoming more and more meaningless (Strate, 1999). Cyberspace is in fact often seen as a network of interconnected electronic devices (e.g. Hemetsberger, 2005). The emphasis is sometimes put on the informational dimension of cyberspace, as in Cicognani’s definition, cyberspace as an “electronic fluxus of information” (1998, p.19). The social dimension of cyberspace has also been early identified by researchers and became an important point of focus for many researches (e.g. Venkatesh, Meamber and Firat, 1998 ; Kalay and Marx, 2001 ; Zhang and Jacob, 2012). One important particularity of cyberspace is that it allows a subject-decentered, fragmented and non-geographically bounded interaction (Venkatesh, Meamber and Firat, 1998), making it the most fertile ground for building and sustaining a postmodern *communitas*. The current development of

social media suggests that social bonding is becoming the most crucial feature of cyberspace, rather than information exchange. These researches seem to not fully appreciate the postmodern nature of cyberspace, which has penetrated every space of human existence, becoming “realer than real” (Venkatesh, Meamber and Firat, 1998). Some notable works discussed the possibility for cyberspace to being compared to “real space”. But conclusions somehow differ. For example, Zhang and Jacob (2012) argue that even if the comparison between cyberspace and space/place is legitimate, « cyberspace is not a ‘space’ but a ‘place’. The ‘space’ in cyberspace is no more than a metaphor for a ‘place’ constituted of ‘place-like’ units » (p. 99). Like some others researchers, they conclude that cyberspace must be understood as a spatial metaphor and is not space *per se*. On the contrary, Coyne (1995) precisely describes how cyberspace may be understood as a world, a space or places. Moreover, he contends that cyberspace shares many common features with physical space in that it is a container, it is measurable space and enabling and constraining human interaction (1995, p.155). This point of view really questions the intuitive distinction that many researchers make between space and cyberspace. This type of space, which is the product of past events, covers every object in relation with human existence and defines possibilities of action as much as interdictions. According to him, these levels are deeply interconnected, relying on one another. Just like “real space”, cyberspace can be described according to the interconnection of these dimensions. Cyberspace is largely depending on its physical constituents : an expansion of the number of connected devices, visual interfaces and servers automatically means an expansion of cyberspace. Like urban space, cyberspace is built according to logical considerations on human needs and thus depends on mental space. Finally, as a formidable medium for social interaction, cyberspace can be understood in terms of social space. Cyberspace is the sum of numerous online places defined by their social utility and by their own implicit and explicit social rules : social media for self expression and communication, online multiplayer games for entertainment, community forums for socialization, educating oneself, etc.

From different theoretical perspectives, cyberspace is equivalent to “real”, offline space. Nevertheless, as Venkatesh, Meamber and Firat's work suggests, the latter allows biological life and the former does not. The latter is directly and naturally accessible to our senses and the former is technology-mediated. As essential as these specificities may be, we should not put aside the fact that cyberspace allows and significantly expands human existence. This idea is closely related to Don Ihde's postphenomenological perspective of technology (1993). According to Ihde, technology allows human beings to access to qualities and dimensions of reality that would be otherwise out of reach. Ihde asserts that human beings develop relationships to reality in a precognitive fashion, before the theoretical constitution of the subject and of the object. Therefore, human beings and their world are always “mutually constitutive in their fundamental interrelation” (1993, p.3). Technology must not be regarded as a mere mediator, between the subject and the world, but rather as a constitutive element in the subject's relationship to reality. According to postphenomenology, virtual worlds must be apprehended as epistemological and ontological instruments (Gualeni, 2014). Following this approach, in the following, cyberspace will be considered from an ontological standpoint and defined as an Internet-mediated space of human existence.

Heidegger's theory of space is useful to apprehend cyberspace as “lived space” (Schatzki, 2007). According to Heidegger, the main constitutive feature of human existence is Being-in-the-world, that is to say, *Dasein*'s working out its self-understanding throughout existence. Therefore, there are many ways of Being-in-the-world. Heidegger also claimed that the primary way *Dasein* is Being-in is practical rather than cognitive. That means that people are primarily “busy” using “equipment” (*Zeug*, for instance a computer, files, softwares, etc.) rather than acquiring knowledge and theorizing about things. Indeed, Heidegger explains that the process of knowing the world is based on a practical way of Being-in that consists in using things for defined purposes. This approach tackles the problem of space in a very novel way : space is constituted by things that allow people to fulfill their goals. From this perspective, cyberspace appears as a legitimate and “real” space. In the next section, we call “cyber-spatiality”, the peculiar way for humans to apprehend cyberspace as a space of existence.

3. IMPLICATIONS OF BEING IN THE CYBERSPACE ON THE CONSUMERS' DIGITAL IDENTITY CONSTRUCTION

Cyberspace springs from technology that is created by humans for humans. Consequently, the technology and interface that support cyberspace are designed in an intelligible way for us : once the computer is turned on, the interface is supposed to make sense for the user. Cyberspace is underlain by an incalculable number

of softwares that mimic features of human existence (Thrift, 2002), however its fundamentally Internet-mediated nature implies that *Dasein* does not relate to cyber-objects and cyber-sociality in the same way than in the material world. According to Heidegger, *Dasein* is characterized by “de-severance” (*Ent-fernung*), which means that *Dasein* has a natural tendency to bring things close, to make them familiar and ready to use (no matter how far these things are situated). Consequently, de-severance tends to create familiarity with distant things such as virtual objects (forums, broadcasted events, online games, etc.) and people (online acquaintances, geographically distant friends). Moreover, as Heidegger suggests, media have the power to expand one's world-around by providing an increasing number of familiar and ready to use objects. Human consciousness and knowledge of the world are especially enhanced by cyberspace as it allows people to grasp circumspectively and bring closer an infinity of scattered objects related to their concernful dealings. Cyberspace is structured to provide this readiness and actionability of knowledge in the more efficient way possible. That is why cyberspace can be characterized as a “teleotopia”. Teleotopic spaces are phenomenologically constructed and structured according to specific projects and purposes.

A large literature on online behavior shows that people tend to conduct themselves somehow differently in cyberspace than in “real” world (Christopherson, 2007). This difference may be explained by the peculiar spatialization people experience in cyber social spaces. Obviously, the non-physical presence and the relative anonymity characterizing online browsing — and therefore, cyber-spatiality — play an important part in the specificity of cyber-communication. These two factors influence the degree to which one frees himself from social norms and roles (Postmes and Spears, 2002). Cyber-spatiality is ubiquity : one can be present in different cyberplaces at the same time, but necessarily in an ephemeral and ghost-like way. Freedom and ubiquity in cyberspace allow people to easily express a multiplicity of identities online (Iteanu, 2008). People undertake an identity play in cyber social places implying the superposition or the interchangeability of these multiple identities (Pierre, 2011). Digital identities are constructed as representations as much for other people than for oneself (Castells, 2009 ; Pierre, 2011). The construction of digital identities involves an introspection, a *cogito*, that questions one's true identity. For instance, the creation of a personal profile on a dating app leads to thinking about oneself and selecting scattered elements among many facets of one's identity. The *de-severance* characterizing *Dasein* allows a person to bind to different online communities in order to fit in. This phenomenon is quite interesting when compared to that of online activism for differentiating oneself (Donath, 2007). One's world-around is constantly shifting particularly because of the endlessness of information streams and the instability of online social relations. As much as in the offline world, people need to carry out different identity performances according to the various contexts they have to deal with. The freedom and ubiquity characterizing cyber-spatiality allow them to evolve easily and simultaneously in otherwise inaccessible social spaces. Consequently, people construct their identities online in a very peculiar way by binding with others from afar, by fitting in social groups they are not familiar with and by performing their identities as they see fit.

In the last section, the implications of these theoretical considerations for consumer research are discussed.

4. CONCLUSION: IMPLICATIONS FOR CONSUMER RESEARCH

First, this research provides epistemological contributions. By shedding light on the ontological dimension of cyberspace, we argued that considerations on the “real” or “virtual” nature of cyberspace are not relevant anymore. For instance, browsing should not be regarded as an activity (what most researches implicitly postulate) but as a technologically mediated way of Being-in. Consequently, connecting to Facebook before even getting out of bed (as related by Llamas and Belk, 2013) is not more surprising than gently telling your wife or husband “good morning” as you wake up. Being-in-cyberspace is not an activity, that is why cyberspace imbues every single part of our existence. It is just a way to *be*. This cyber-ontology also has theoretical implications about the identity of the digital consumer. The analysis of cyberspace according to Heidegger's philosophy pretty much agrees with Kozinets' (2013) judicious point : the digital consumer that we have been looking for with much determination is not some kind of unfathomable entity, it is just *us*. Nevertheless, this point seems to underestimate the ontological value and influence of cyberspace. This paper argues that, in cyberspace, *Dasein* is characterized by cyber-spatiality. This means that the way we relate to our environment — the world-around — is necessarily altered by cyberspace. The cyber-spatialized consumer is an ontologically augmented person (Gualeni, 2014) because he has an almost unlimited and

immediate access to “equipment”. But as Kozinets (2013) suggests it, we must not engage ourselves in some kind of post-humanist poetics and praising blindly the advent of a technologically transcended “cyborg”. As we argued earlier, cyberspace is a plan of human existence, and therefore a place of mistrust as much as a place for sharing. Many people see their lives enhanced by cyberspace but many get also scammed and cyberbullied in cyberspace (Moore et al. 2012). Cyber-spatiality makes us much more aware of the world, but it can be for the worst.

In conclusion, this essay aimed at showing that cyberspace can and should be regarded as a real space, a lived space where identity projects are carried out in a very peculiar way. We should embrace cyberspace as the important part of human existence that it is and leave behind the outdated question of its reality. By focusing on its growth as a legitimate plan of human existence and its very real influence on people's identities, we can hope for tremendous developments in our understanding of this fascinating phenomenon.

REFERENCES

- Castells, M. and Llodrà, B., 2009. *Who are you on the internet ? Digital identity : an exhibition concerning technosocial uses of the net*. Proceedings of the IADIS international conference e-society 2009, Barcelona, Spain, pp. 253-257.
- Christopherson, K. M., 2007. The positive and negative implications of anonymity in Internet social interactions: "On the Internet, Nobody Knows You're a Dog". *Computers in Human Behavior*, 23(6), 3038-3056.
- Cicognani, A., 1998. On the Linguistic Nature of Cyberspace and Virtual Communities. *Virtual Reality*, 3 : 16-24.
- Coyne, R. (1995). *Designing information technology in the postmodern age*.
- Descartes, R., 1983. *Principles of Philosophy*. R. P. Miller and V. R. Miller (trans.). Dordrecht : D. Reidel.
- Donath, J., 2007. Signals in social supernets. *Journal of Computer-Mediated Communication*, 13(1), 231-251.
- Franck, D., 1986. *Heidegger et le problème de l'espace*. Paris : Editions de minuit.
- Gualeni, S., 2014. Augmented Ontologies or How to Philosophize with a Digital Hammer. *Philosophy and Technology*, 27(2) : 177-99.
- Heidegger, M., [1954] 1977. *The Questioning Concerning Technology and Other Essays*. William Lovitt (trans.). New York: Garland Publishing.
- Heidegger, M., [1927] 1962. *Being and Time*. J. Macquarrie and E. Robinson (trans.). NY, New York: Harper & Row Publishers.
- Heidegger, M., [1959] 1969. *Discourse on Thinking*. M. Anderson and H. Freund (trans.). NY, New York: Harper & Row Publisher.
- Hemetsberger, A., 2005. Creative Cyborgs: How Consumers Use the Internet For Self-Realization. In G. Menon and A.R. Rao (eds), *NA - Advances in Consumer Research*, Volume 32, pp. 653-60. Duluth, MN : Association for Consumer Research.
- Ihde, D., 1993. *Postphenomenology: essays in the postmodern context*. Evanston: Northwestern University Press.
- Iteanu, O. and Salvatori, O., 2008. *L'identité numérique en question*. Eyrrolles.
- Kalay, Y. E., & Marx, J., 2001. The role of place in cyberspace. In *Virtual Systems and Multimedia*, 2001. Proceedings. Seventh International Conference on (pp. 770-779). IEEE.
- Kozinets, R.V., 1999. E-tribalized Marketing ? The Strategic Implications of Virtual Communities of Consumption. *European Management Journal*, 17(3) : 252-64.
- Llamas R. and Belk R. W., 2013. Living in a Digital World. In R. Llamas and R.W. Belk, *The Routledge Companion to Digital Consumption*, pp. 3-12. NY, New York : Routledge.
- Moore, M. J., Nakano, T., Enomoto, A., and Suda, T., 2012. Anonymity and roles associated with aggressive posts in an online forum. *Computers in Human Behavior*, 28 : 861-67.
- Pierre, J., 2011. Génétique de l'identité numérique. *Les Cahiers du numérique*, 7(1), 15-29.
- Postmes, T. and Spears, R., 2002. Behavior online: does anonymous computer communication reduce gender inequality. *Personality and Social Psychology Bulletin*, 28(8) : 1073-83.
- Schatzki, T. R., 2007. *Martin Heidegger: theorist of space*. Stuttgart: Steiner.
- Strate, L., 1999. The varieties of cyberspace: Problems in definition and delimitation. *Western Journal of Communication*, 63(3) : 382-412.
- Thrift, N. and French, S., 2002. The Automatic Production of Space. *Transactions of the Institute of British Geographers*, 27(4) : 309-335.
- Venkatesh, A., Meamber, L. and Firat, F., 1998. Cyberspace: The Next Marketing Frontier (?) Questions and Issues. In S. Brown and D. Turley (eds.), *Consumer Research: Postcards from the Edge*, pp.301-21. London: Routledge.
- Zhang, G., & Jacob, E. K., 2012. Reconceptualizing Cyberspace:" Real" Places in Digital Space. *International Journal of Science in Society*, 3(2).

PRIVACY IN LOCATION-BASED APPLICATIONS FOR AIRPORT USERS

Nayra Rodríguez-Pérez and Pino Caballero-Gil
Department of Computer Engineering and Systems, Spain

ABSTRACT

Nowadays, privacy is considered a necessity in the world, as a consequence of the fact that users are increasingly sharing their personal information and location data at all times. The case study of this paper is based on the lack of privacy regarding the identity of users in various types of location-based services. These services are usually deployed in mobile applications related to the management of airport users and their travel plans. A study of many existing mobile applications and different research orientations is here conducted to address the privacy issue in location-based systems. The huge increase in apps with these services makes identity privacy and data protection a big challenge because in many of those apps, users and their personal information are monitored at all times.

KEYWORDS

Airports, Location-Based Services, Applications, Privacy

1. INTRODUCTION

Mobile devices, and in particular smartphones, offer a wealth of personalized services that enrich the user experience. This is mainly a consequence of the possibility of access to personal information stored in the devices through different applications. Thanks to that, those applications can be customized taking into account data collected through different systems such as locating or providing information from other applications. Mobile devices enable snapshot distribution of information and may broadcast location information. There are multiple types of applications that integrate location as a service to provide the user about information available in their environment through providing personalized information for each user. In this sense, some of them are especially remarkable because they make use of these data to offer resources and functionalities based on this information (social networks, recommendations, climate conditions, health, travel, games, e-commerce, emergencies, etc.). Tourism applications are an example of using location-based services to help users in their travel planning (Yu, Chang, 2009).

Accessing to location data of users involves a privacy issue because this information may reveal sensitive personal data. In addition to provide information about the user's location, other types of information can also be deduced, such as movements, place of residence, ideologies, actions, etc. Current systems must prioritize the need to protect data to avoid the unauthorized access to their information and identity and ensure that applications can provide the same functionality without compromising users' identity or linking all this related information. Location-based applications provide a service to search for environmental resources through location information. The protection of user information may be compromised by an unreliable mobile application that makes use of this data. Location on mobile devices can be obtained through the GPS system or wireless networks. Despite that is private information, it might be used, infringing user privacy (Junglas, Johnson & Spitzmüller, 2008).

Regarding airports, there are systems that monitor the position of users inside and outside the terminals and add this information to their travel plans. This fact supposes a privacy problem for user identification and location at all times. This paper includes a brief state of the art of the study of users privacy in different airport applications, together with a survey of different methods that are currently being proposed as a possible solution to protect the identity of users when using location-based applications.

2. ACCESS TO LOCATION DATA

The main case study in this work is the access to location data of users in location-based airport applications. In particular, it is especially interesting the exchange of information and data related to flights or travel plans between different applications and services. Today, the large majority of these mobile applications obtain their location from user devices in real time, both inside and outside the airport. A user who regularly uses a mobile application to monitor his or her travel plans, typically wants real-time information about flight status, so it is important to find a solution that combines both goals: to provide useful information to the user, and to integrate a high security layer into the location-based functionality to protect location information from possible illegitimate uses. On the one hand, location-based functionalities allow enriching the lives of people thanks to a better knowledge of their preferences and desires (Cormode & Srivastava, 2009). On the other hand, when that knowledge concerns very private and sensitive aspects of people lives, location-based services may suppose a privacy problem that needs to be seriously considered and analyzed.

In general, regarding the perspective of the large number of users who use location-based applications, there is no excessive concern about their privacy considering their location data. Users use services without worrying about how those service manage the protection of their data. Different studies include analyses of this issue. On the one hand, statistics obtained from a survey done to airport users (Hansen, Alapetite, Andersen, Malmborg & Thommesen, 2009) show that users may want different levels of protection. On the other hand, in addition to the study of possible privacy threats, different ways to get the location have also been studied in (Krumm, 2009) to obtain the most accurate location values.

Applications with access to email have also access to the personal and identification data of users regarding trips, such as name, surname, identity number, reservation number, travel itinerary, etc. In addition, they do not only have access to the data of the user, but it may also get information via email regarding other passengers sharing the journey. This is a problem because of email associates directly with the identity of a particular user. In this case, one of the most important privacy problems appear, corresponding to the case when different applications share data and private information related to journeys and routes of users.

In addition to this, it is essential to highlight the introduction of the new data protection law as of this year. In this law, users are given greater power over their personal data, which will allow them to have greater control over them. Service providers must, therefore, improve and update their security practices, eliminating vulnerabilities and protecting their user personal data. It must be ensured that personal data are processed in such a way as to guarantee their security, that there are no losses or unauthorised and illegal procedures. Among the most important measures to implement the new law are pseudonymization and encryption of personal data. Finally, the confidentiality and integrity of personal information must be guaranteed.

3. STUDY OF ANONYMIZATION SOLUTIONS

In order to provide information about privacy in airport mobile applications or information that may be related to travel planning, a study has been carried out on the access requirements of these applications when they are installed on the mobile devices of users. As explained in (Ghinita, 2013), despite all the benefits of location-based applications, the disclosure of location in different services raises serious privacy concerns since they may disclose confidential information about users.

There are many research papers highlighting the need to apply privacy and confidentiality techniques to private user location data, as well as a diverse set of solutions to protect user identity. Anonymization of users in location-based applications and services is proposed as one of the possible methodologies to solve this problem (Figure 1. Anonymization). The anonymization process reduces the risks of identifying the data while maintaining the veracity of the results in their treatment, so any operation can be carried out without any distortion with the real data (AGPD, 2017). To this respect, data and information relevant to different systems (e.g. passenger flow) can continue to be available without associating such data with a particular user identity. A method proposes the removal or encryption of personally identifiable information as a technique to ensure the protection of the private data of users. It is important to establish an algorithm to meet the user anonymity restrictions used by location-based services in a given area.

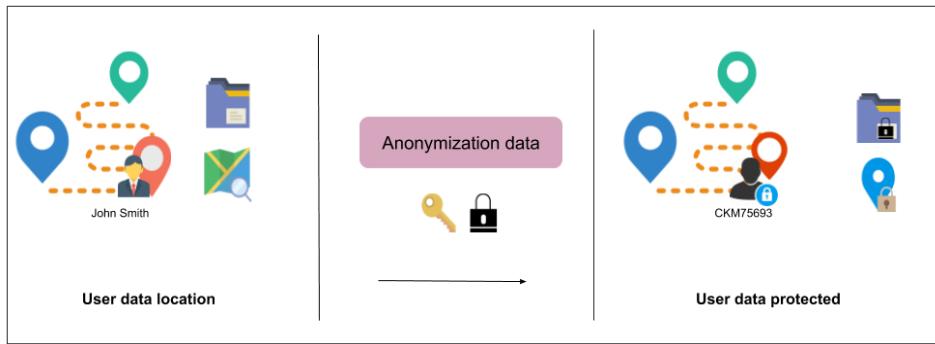


Figure 1. Anonymization

Ensuring the data anonymous usage in location-based services requires that the location information transmitted by a user cannot be to re-identify the user. Anonymization is usually used through a series of different techniques that alter real user data to hide sensitive information about each person. These anonymization techniques can be used from different approaches. In general terms, the following forms of anonymization protection stand out:

- Data encryption, in order to provide confidentiality to the identity of users.
- Grouping information into sets.
- Dissociation or exchange. The user information is isolated from all of his or her personal data.
- Deleting information from a data set.

An opponent can model the confidential information obtained through the user location data, in addition to obtain private information about their residence or movements. Many aspects participate in ensuring a user can obtain a complete location-based service, so the private user information may be potentially known by many different services or content providers. As consequence, increasing the amount of personal information that is shared among different elements or services represents a difficult problem to be solved. This issue requires a control mechanism for information accessing and protecting, in addition to an appropriate authentication system (Jagwani & Kaushik, 2017).

Information obfuscation, mainly of the user identity and location information, is a way to deal with the problem. By using anonymization techniques, data is obscured to make it difficult to relate identity with personal information. The application of anonymization is based on using all the information available on mobile devices and their location-based applications while hiding all private information to the attacker who tries to link it to the data on the device. In this way, it is possible for applications to provide location-based functionality without affecting the privacy of user identity.

Another alternative of anonymization solution is based on sending multiple location data to the servers instead of uniquely sending the actual location values of the user. This method allows the user to send a set of location data instead of a single location value with the actual data. The work (Sun, Chen, Hu, Quian & Hassan, 2017) makes an extensive and detailed study of this technique with the aim of protecting the real location of users by generating a number of fictitious positions around their location. It would also be possible to provide users the possibility of selecting their level of privacy regarding the applied anonymization degree. On each mobile device, users can customize privacy requirements while respecting a minimum level of protection. It is necessary to identify and model accessible and public information so that the privacy techniques used may preserve data usefulness. This aspect will allow service-based applications to remain functional without affecting the privacy of users.

4. CONCLUSION

Service-based localization is present in a multitude of mobile applications. It is important to provide this functionality in multiple applications that make life easier for people, but it is also imperative to guarantee the privacy of information and user identity. Currently, the geolocation functionality is fundamental in many applications that are widely used. Especially in airport environments, travellers use more and more

applications that facilitate their travel plans through location-based services. This means that these users are located at all times, both inside and outside the airport during their travels.

The protection and privacy of user data and their location it is really important. For this reason, a study has been carried out that proposes the anonymization of data as a possible solution to protect the personal information of individuals and their movements. In this way, the information generated by the applications is decoupled with the identity of the users. For the use of anonymization in different application providers it is necessary to apply different techniques that unlink this information. This line of research is a reflection work in progress with the aim of applying anonymization in the field of information protection regarding location-based services.

ACKNOWLEDGEMENT

Research supported by the CajaCanarias Foundation, the Centre for the Development of Industrial Technology, the Ministry of Economy Industry and Competitiveness, and the European Regional Development Fund, under Projects DIG02-INSITU, IDI-20160465, TEC2014-54110-R, RTC-2014-1648-8 and MTM2015-69138-REDT.

REFERENCES

- AGPD. (2017). Orientaciones y garantías de anonimización. 20/02/2018, from Spanish Data Protection Agency Web Site:https://www.agpd.es/portalwebAGPD/canaldocumentacion/publicaciones/common/Guias/2016/Orientaciones_y_garantias_Anonimizacion.pdf
- Cormode, G., & Srivastava, D. (2009, June). Anonymized data: generation, models, usage. In Proceedings of the 2009 ACM SIGMOD International Conference on Management of data (pp. 1015-1018). ACM.
- Ghinita, G. (2013). Privacy for location-based services. *Synthesis Lectures on Information Security, Privacy, & Trust*, 4(1), 1-85.
- Hansen, J. P., Alapetite, A., Andersen, H. B., Malmborg, L., & Thommesen, J. (2009, August). Location-based services and privacy in airports. In IFIP Conference on Human-Computer Interaction (pp. 168-181). Springer, Berlin, Heidelberg.
- Jagwani, P., & Kaushik, S. (2017, March). Privacy in Location Based Services: Protection Strategies, Attack Models and Open Challenges. In International Conference on Information Science and Applications (pp. 12-21). Springer, Singapore.
- Junglas, I. A., Johnson, N. A., & Spitzmüller, C. (2008). Personality traits and concern for privacy: an empirical study in the context of location-based services. *European Journal of Information Systems*, 17(4), 387-402.
- Krumm, J. (2009). A survey of computational location privacy. *Personal and Ubiquitous Computing*, 13(6), 391-399.
- Sun, Y., Chen, M., Hu, L., Qian, Y., & Hassan, M. M. (2017). ASA: Against statistical attacks for privacy-aware users in Location Based Service. *Future Generation Computer Systems*, 70, 48-58.
- Yu, C. C., & Chang, H. P. (2009, September). Personalized location-based recommendation services for tour planning in mobile tourism applications. In International Conference on Electronic Commerce and Web Technologies (pp. 38-49). Springer, Berlin, Heidelberg.

MEDICAL E-LEARNING PLATFORM "IMAGE OF THE WEEK": ADVANCES AND CHALLENGES

Luísa Bernardino Valério, Giovanna Vieira Moreira, Laio Bastos de Paiva Raspante,
Lucas Augusto Carvalho Raso, Bruno Campos Santos and Viviane Santuari Parisotto Marino

Universidade Federal de Minas Gerais

Avenida Professor Alfredo Balena, 190 – Santa Efigênia, Belo Horizonte, MG – 30130-100, Brazil

ABSTRACT

Image of the Week (IW) Project is an open access e-learning platform established in 2010 at the Faculty of Medicine of the Federal University of Minas Gerais (FM-UFMG), Brazil. A clinical case elaborated from daily medical experiences, with diagnostic images associated to a multiple-choice question is presented weekly. By marking the correct answer, one has access to the theoretical discussion of the case, complementary propaedeutics, treatment and relevant differential diagnoses. In the last seven years, the project has expanded and reached many accesses on social networks, with regular publications of articles in scientific journals and publication of two textbooks. A partnership with a supplementary healthcare provider cooperative was also established, leading to a dissemination of the project content for the continuing education of its members. This work analyzes, in a systematized way, the evolution and results achieved by this project, reflecting its potentialities and perspectives.

KEYWORDS

Image of the Week, e-Learning, Medicine, Continuing Education

1. INTRODUCTION

Image of the Week (IW) Project is an open access e-learning platform established in 2010 at the Faculty of Medicine of the Federal University of Minas Gerais (FM-UFMG), Brazil, accessible at www.medicina.ufmg.br/imagemdasemana, available in Portuguese and English. A weekly clinical case drawn from daily medical experiences of undergraduates and physicians, with diagnostic images associated to a multiple-choice question is presented. The current members of the project annually select a group of medical students to join in and be responsible to elaborate cases for the website based on medical care they helped provide. Teachers or other physicians orient the content of the texts and correct it making sure it is according to the literature. By marking the correct question answer, one has access thus to a theoretical discussion of the case, the workup and therapeutics instituted as well as of its differential diagnoses.

IW purpose is to share and democratize the knowledge generated in the university with the medical community in general. Epidemiological aspects and current trends of study are considered through problem-based learning, seeking to develop clinical reasoning, optimize learning time and stimulate self-learning and critical thinking.

Visitors have access to all published cases in the previous weeks, being able to use the page as a studying tool, a revision material and as to incorporation of concepts. It has recently been added a "Test Question" section, with questions from real medical examinations to further contribute to the establishment of the learned knowledge, demonstrating the relevance of that content in quizzes of different contests, encouraging the complete reading of the presented content and the interest until the last years of the medical course.

As social networks, Facebook (@imagemdasemanaufmg) and Instagram (@is_oficial) are used to alert the publication of clinical cases and dissemination of the project. There's also an interactive app for Android mobile devices (http://bit.ly/app_IS) and an e-book (http://bit.ly/ebook_IS). Another form of access is through the website of UNIMED Medical Healthcare Cooperative of Belo Horizonte (UNIMED-BH), which reproduces the content of IW's website as an instrument of continuing medical education for its cooperating physicians (http://bit.ly/UNIMED_IS). The Revista Médica de Minas Gerais (RMMG), an indexed Brazilian medical journal, is another partner of the project that publishes, every three months, a compact version of the best case of the quarter (last edition: http://bit.ly/RMMG_IS).

2. RESULTS

2.1 Statistical Accesses

IW platform presented a significant increase in accesses from 2011 to 2016, going from 139,768 to 301,095 accesses / year on its official website (Figure 1), accounting for the total of 1,369,406 accesses through all 239 published cases in the last seven years.

In spite of the absolute quantitative increase in the number of visits to the website, there was a reduction of about 77% in the number of accesses to the discussion of the case when comparing the years of 2011 to 2016. This reduction was more pronounced after 2014, which determined the adoption of measures such as strict compliance with the number of words/lines, in order not to make the text too long, which would require a lot of time on the website.

In 2012, the content began to be available also in English, which practically tripled the number of international accesses (Figure 2). Today, IW is accessed by visitors from all continents, which 2% of all our visitors are from Portugal, 0.8% visitors from the United States and 0.61% visitors from Bolivia. Together, international accesses represents 7% of hits to the website.

2.2 Actions and Partnerships

IW has been awarded five prizes so far. The first was a national award for best work presented at the congress of the Brazilian Association of Medical Education (ABEM) and the others prizes were in events at our institution, including honorable mention and academic relevance. Concurrently, the project organized interactive symposiums for physicians and academics annually. In addition, the project produced an e-book available at: http://bit.ly/ebook_IS (Figure 3) containing the compilation of the best cases among the first hundred published.

The participation of the IW on social networks began in November 2015 with the aim to approach the target audience and increase dissemination of the project through Facebook. A progressive increase in the number of followers was observed, reaching a total of 9,548 supporters upon the completion of this work (Figure 3) and national projection as shown in table 1.

Instagram, on the other hand, was created in April 2016, showing little impact when compared to Facebook, accounting for only 798 followers and an average of 20 Likes/publication. We believe that the profile of this network and the time spent on it justifies the smaller scope of the project when compared to other means.

2.3 Future Challenges

The evolution of e-learning technologies and modalities has also revolutionized medical education. Currently, students study connected to the Internet, getting a lot of information in a short time, which can compromise the retention of its content. IS visitors spend, on average, less than three minutes on the website, preferring succinct cases and therefore concise discussion. Thus, more complex and broader-based cases, which require greater skill in their scientific dissertation for publication, are being underused by the visitor. Thus, the proper use of this content is demanding the elaboration of new strategies, whose first initiative was to insert a “Test Question” at the end of the complete text.

The main challenge of this project is to incorporate the widely accepted technological advances. In other words, the aim is to use the same language of the young people so that the learning is attractive. Improve the quality of the presented images and their interactivity is also a priority to the project, since one of its objectives is to discuss the incorporation of new technologies in health assistance.

2.4 Figures and Tables

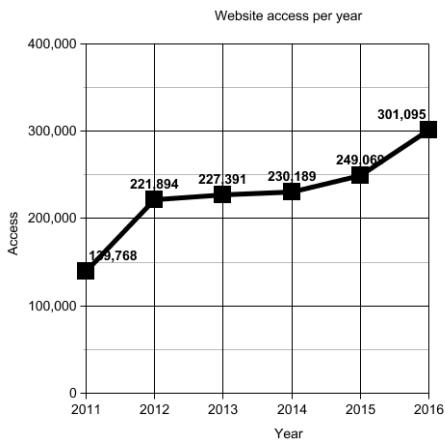


Figure 1. Graphic: Number of accesses to the Image of the Week platform, from 2011 to 2016

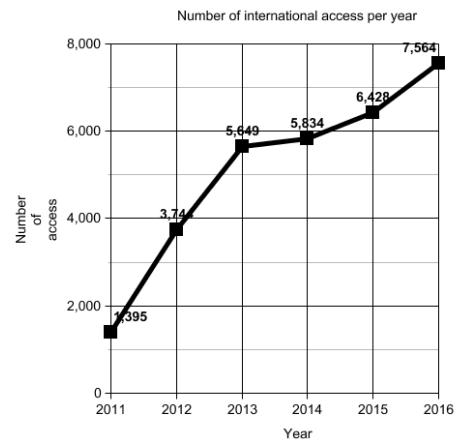


Figure 2. Graphic: Number of international accesses to the Image of the Week platform, from 2011 to 2016

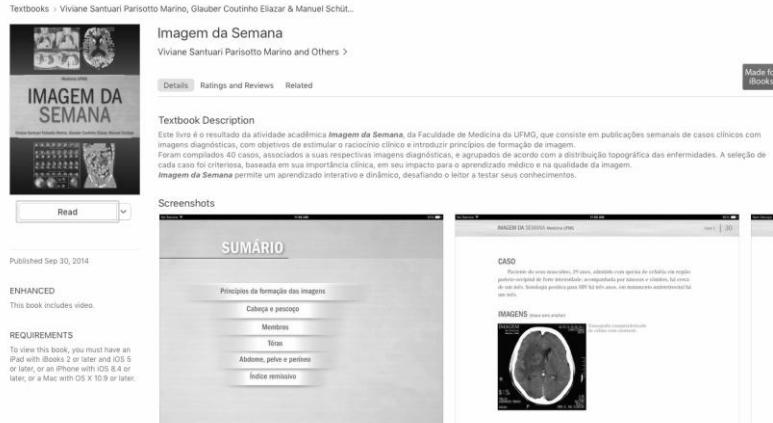


Figure 3. Reproduction from the App Store page to download the e-book "Image of the Week", published in 2014

Table 1. Followers of the Image of the Week Facebook page, by Country

Followers on Facebook Per Country	
Brazil	9.202
Bolivia	114
Paraguay	37
Argentina	31
United States of America	16
Angola	15
Mexico	14
Peru	12
Portugal	11
Germany	7

3. CONCLUSION

Discussion of clinical cases, a result of the academic experience at the Faculty of Medicine-UFGM, associated with diagnostic images has been well evaluated by IW users as a tool for continuing education for students and medical assistants. Such an acceptance is reflected by the progressive increase in the number of visitors from different states in Brazil and abroad, certainly boosted by the dissemination of the project in social networks. The reflection on the performance of IW over its seven years of activity evidenced the need to: broaden its dissemination; increase its interactivity and connectivity; and incorporate technological advances without loss of quality on its content in order to reach its main objective, which is its recognition as a tool of continuing education in medicine.

ACKNOWLEDGEMENT

We thank supervising teachers, academic community of FM-UFGM, FM-UFGM Board and UNIMED-BH. We also acknowledge the importance of former members of the project, who provided insight and expertise that greatly assisted our work.

REFERENCES

- Farooq, A. and White, J. (2014). #Nomoretextbooks? The impact of rapid communications technologies on medical education. *Canadian Journal of Surgery*, 57(4), pp.E119-E120.
- Frehywot S. et al, 2013. E-learning in medical education in resource constrained low- and middle-income countries. In: *Human Resources Health*. Vol. 11, No 4.
- Masic, I. (2008). E-Learning as New Method of Medical Education. *Acta Informatica Medica*, 16(2), p.102.
- Schütze, M. et al, 2012. Projeto Imagem da Semana (Image of the Week Project). In: *Revista Brasileira de Educação Médica*. Vol. 36, No 3, pp 423-430.
- Whitcomb ME, 2004.. More on competency-based education. In: *Academic Medicine*. Vol.79, No 3. pp 493–494

TECHNOLOGICAL EXURBIA: LIMINALITY AND KNOWLEDGE CITIES' CONCEPTUAL BOUNDARIES

Lucia Elizondo

Tecnológico de Monterrey

Av. Eugenio Garza Sada 2501 sur, Col. Tecnológico, Monterrey, NL 66247 Mexico

ABSTRACT

Knowledge based urban development (KBUD), an offspring of Knowledge Societies, considers the transformation of knowledge resources into local development. Spatial configuration is not considered within the four development domains that are the key pillars of KBUD: economy, society, environment and management. This reflection paper has the aim to analyze what are the repercussions of not considering spatiality within its domains. By focusing on knowledge clusters, KBUD contributes to the marginalization of areas in Knowledge cities. Exurbia is one such area, which if considered physically and metaphorically, is in a liminal state, a zone of transition, in which knowledge plays no role. The paper concludes that KBUD should acknowledge areas as these, which in turn could leverage the homogenizing gaze it imposes and contribute to knowledge construction and exchange.

KEYWORDS

Knowledge Society, Knowledge-Based Urban Development, Knowledge Exchange, Exurbia, Liminal

1. INTRODUCTION

Cities magnify humanity's strengths (Glaeser, 2011). Yet 21st century cities are facing important challenges. As Yigitcanlar (2011) points out, cities are required to be entrepreneurial, eco-friendly, more inclusive and transparent in their decision-making processes while supporting social equity. And as he further states: "All these new requirements are leading to the development of a new planning and development approach or paradigm for cities, one that is applicable for different geographic and political contexts" (2011, p. 5). Yet, it is precisely the spatial configuration by which cities emerge that limits the homogenizing gaze of a paradigm through which cities can be understood. Characterizing a city today is by no means an easy endeavor. It requires not only comprehending its locality but also sizing it up in order to manage it. However, as Thrift and Amin state: "the city's boundaries have become far too permeable and stretched, both geographically and socially, for it to be theorized as a whole" (2002, p. 8). This leads us to ask then, how are we to tackle and approach the challenges that cities meet if it can't be theorized as a whole? Knowledge Societies, and one of its offspring, Knowledge Based Urban Development (KBUD), has attempted to undertake these considerations with success, except that in their approach, spatial configuration of cities and its repercussion on knowledge exchange has been avoided. So the aim of this critical reflection is to further the understanding in spaces of engagement for KBUD considering the spatiality of cities and its contributions to human development.

2. INCREASING KNOWLEDGE BASED URBAN DEVELOPMENT SPACES OF ENGAGEMENT WITHIN KNOWLDEGE CITIES

Knowledge societies are about "capabilities to identify, produce, process and transform, disseminate and use information to build and apply knowledge for human development" (UNESCO World Report, 2005, p. 27) while promoting integration of all its members, for knowledge is a public good available to each and every individual (2005, p. 18). However, "control of knowledge can go hand in hand with serious inequality,

exclusion and social conflict" (2005, p.17). Moreover, guaranteeing the access to such information, and closing the digital divide does not translate to knowledge gain or mastery, for it is more than a matter of infrastructure 2005, p. 22). Yet not closing the digital divide at all would be exponentially detrimental for some, as they point to the fact that knowledge participates in a virtuoso circle where knowledge and its consequent innovation produces more knowledge in the long term (2005, p. 28). Marginalized territories on knowledge societies face serious consequences then, if not taken into account.

Several years later, on a revision to the UNESCO's original aims they introduced the concept of inclusion to their revised definition that states as follows: "By inclusive knowledge societies, UNESCO means societies in which people have the capabilities not just to acquire information but also to transform it into knowledge and understanding, which empowers them to enhance their livelihoods and contribute to the social and economic development of communities" (Building Inclusive Knowledge Societies, 2014, p. 18). As can be seen, this later revision acknowledges three things: first that knowledge is constructed, not handed down, second, that knowledge empowers and third, that it emphasizes the responsibility of individuals to contribute to their communities. What is not taken into account is spatial inclusion, which is paramount to knowledge exchange and community building. The rise of the concept of network societies, advanced by Castells (1996) contributed to the understanding of human relations as de-territorialized, yet many cities from the developing world are still in its industrial robe, while at the same time, embarking in this knowledge exchange and dependent on locality, or spatial configuration, for establishing human relations. As Carillo (2014) points out: "socially generated learning and knowledge based development are determined by people's degree and capabilities for relationships." Many areas of developing cities are still the physical locus for knowledge transference, construction and exchange, and as May and Perry explain cities are critical for knowledge based growth in the 21st century (quoted in Yigitcanlar, 2011). This all begs the question then, as to where knowledge and cities meet.

The crossing point of both has been the field of study of Knowledge Based Urban Development (KBUD). Originally introduced as Knowledge Based Development by Richard Knight, it considered the transformation of knowledge resources into local development, while providing a basis for sustainable development (Yigitcanlar, 2011). Later on, taken by the OECD, this paradigm was seen as beneficial to foster more sustainable objectives and improve the quality of life and, on 2004, was defined as a medium for the creation of Knowledge Cities. By the mid 2000's the phenomenon gained substantial approval "for enhancing the competitiveness of cities within the context of expanding knowledge-based economy and society" (Yigitcanlar, 2011) and began to be regarded not only as a strategy but as a process as well. In order to achieve its goals, as Yigitcanlar (2011) states, it has four development domains that are the key pillars of KBUD: economy, society, environment and management, which in turn focus on economic prosperity, a just socio-spatial order, environment, sustainability and good governance respectively. So then, how are these development domains translated into cities' functioning? From these four pillars I would like to dwell on the second and the third, on the socio-cultural development domain of society and environ-urban development, which seek a just socio-spatial order, because it is precisely the purpose of this reflection to establish what this approach is lacking. Both, institutional and economic development do not take into account the spatiality of the built environment altogether.

The goals of socio-cultural development, within a KBUD perspective, aim to "progress towards establishing a society – a knowledge society – in which the generation, distribution, diffusion, use, integration and manipulation of knowledge and information is a significant economic, political and cultural activity" (Yigitcanlar, 2011, p. 6). Increasing the skills and knowledge base of residents is prioritized for, social and human capital, as Frane (2005) states, and are seen as highly interrelated with achievement within this domain (quoted in Yigitcanlar, 2011). As can be seen, the goals of this domain, although previously stated as a just socio-spatial order have nothing to do with spatial configuration of cities as such. Spatiality and order are considered here as abstract and metaphorical, not geographical and material. The city is seen as a network of many actors of which spatiality and materiality are not included, and I ask myself how is justice to be dealt with in spatial terms if space is ignored? Even on 2011, when Yigitcanlar based on further research findings redefined KBUD¹, the domain of the relationship between the configuration of space and

¹ "A new development paradigm of the knowledge era that aims to bring economic prosperity, environmental sustainability, a just socio-spatial order and good governance to cities, and produces a city purposefully designed to encourage the production and circulation of knowledge in an environmentally conserved, economically secure, socially just and well governed human setting" (Yigitcanlar, 2011, p. 11)

the city was left out. Many theorists believe the city is made of flesh, not of concrete (Glaeser, 2011), and it may well be the case, but a urban paradigm that disregards the effects of spatial configuration and materiality on its “flesh” relations is myopic.

Furthermore, the enviro-urban development, “aims to meet human needs while preserving the environment so that these needs can be met not only in the present, but also for generations to come. Enviro-urban development ties together concerns for the natural systems with the social challenges facing humanity and builds a strong spatial network relationship between urban development clusters while driving an urban development that is ecologically friendly” (Yigitcanlar, 2011, p. 7). Although this developmental approach includes space, at least as network, it poses a serious dichotomy: that which in the city’s spatial configuration could be labeled as a urban development cluster and that which is not. I don’t see this being a ‘socio-spatial just order’ in any way. Just as I specified earlier, Amin and Thrift stated that cities could not be theorized as a whole because of its expansion, then we could also state that we could not theorize cities as homogeneous wholes because of the many faces and places our cities have. So, urban clusters of any kind, are a city’s natural consequence and are not defective on their own, but defective by what they exclude essentially: every other city space outside their margins. While the city is said to borderless, urban clusters are clearly defined. If knowledge based urban development is to focus on ‘clusters’ it follows maybe, as Knowledge societies have pursued, that cluster developments will pull the rest of the urban areas into their domain, yet this has not been the case. It would be ignorant not to acknowledge that a leveled city development would be difficult, but in order for cities to thrive and become resilient maybe this could be another goal of KBUD. By precisely focusing on clusters, the socio-spatial justness is challenged.

Furthermore, in the practice of knowledge exchange tensions are produced. These problems derive from Knowledge societies’ homogenizing gaze that imports success stories to a variety of contexts “in the face of which the distinctiveness of place evaporates” (May and Perry, 2010). According to Carillo et al. (1998), for a knowledge event to occur, three necessary conditions must be met: a knowledge object, that which is known, a knowledge subject, he or she who knows, and a knowledge context, axiological and semiotic references providing value and meaning and, as he states “critical at this stage is to emphasize the rarely considered knowledge context dimension for it is this dimension that provides of economic relevance and cultural significance” (2014, p. 5) that KBUD undertakes. Knowledge transference is of no use, “knowledge needs to be actively received, understood and interpreted” (Perry and May, 2010, p. 19), all of which is dependent on context, which is in turn dependent on spatial configuration.

The city, as an heterogeneous whole, needs to be locally understood, and in implementing KBUD “sensitivity to place-based dynamics is critical, as well as the need to actively intermediate between different parties to enable a productive relationship between knowledge and the space to be forged” (May and Perry, 2010, p. 20). May and Perry not only acknowledge the importance of locality, but also finally relate it to space, ‘the space to be forged’. Furthermore, they emphasize the importance of face-to-face interactions that are carried out in real physical places, not abstract knowledge networks. Finally, May and Perry stress the importance of “the development of more holistic, rather than piecemeal, frameworks for knowledge exchange at the urban level” (2010, p. 16). Therefore, a holistic framework would necessarily ground other intangible dimensions of KBUD in a tangible infrastructure dimension where space would have to be included.

Up until now, I have been stressing the importance of spatial configuration on knowledge exchange and its avoidance in KBUD spaces of engagement. Cities are non-homogenous incomplete wholes made up inevitably by clusters. Knowledge cities have been treating the city as if where similar to knowledge. Technology has permitted to have knowledge-on-demand, and KBUD principles expect the city to behave the same. Yet there is no such thing as a city-on-demand or architecture-on-demand. As long as physical transition spaces connect the spatial configuration of cities, there will be no such advantage as that which technology has supplied for knowledge access. Therefore, marginalized spaces on the spatial configuration of cities present a disadvantage that KBUD has not foreseen. One such marginalized space is exurbia.

The term exurbia was originally coined by Spectorsky (1955) as a space beyond the edge of cities. As a place “it has been studied under the banner of urban fringe studies, amenity migration studies, and political ecology” (Taylor, 2011, p. 324) and “there is more to characterizing it than identifying the zone it occupies” (2011, p. 326). In developed countries, many people are exurbanites by choice, ranging from a an anti-urban wish to live and work on the country or seeking environmental amenities by living in the country and commuting to the city. Yet, in many developing countries the growth of cities is determined by economic imperatives that push the cities’ boundaries out of its limits and become the only available housing option

space for many. Exurbia in the developing world then, differs from exurbia in the developed world in many ways: 1) inhabitants populate it by affordability not by choice; 2) it consists of alienated high density islands of low income housing instead of low density high income nature seekers; and 3) while exurbia is post-productivist in the developed world, exurbia in developing countries is productivist and characterized by a work migration population. Exurbia has become a necessity, and unfortunately as Glaeser (2011) states, there is too little government on the edge, which in turn means little space of opportunity for knowledge and the city to meet. Exurbia affords few chances of knowledge exchange, and it is not on the radar of either KBUD or local governments in most developing countries. The center-periphery model has dominated the evolution of most cities further increasing the unjust socio-spatial order while privileging islands of knowledge close to the center decreasing the opportunities for knowledge exchange in the periphery, in zones like exurbia.

The concept of exurbia though, offers more than an understanding of a spatial zone. So how then, would we characterize the relationship of knowledge to exurbia conceptually? Liminality is a concept borrowed from cultural studies that helps characterize and understand the essentials of a zone like this in conceptual terms. It can be understood as a “withdrawal from the ordinary forms of social interaction, ambivalence and being on a state of transition” (Turner 1977, p. 167). Exurbia as a liminal zone is a process: a state of transition between urbanity and marginalization, between knowledge and ignorance, leftover of the creation of a “smattering of ‘knowledge islands’ in a great sea of marginalized outsiders” (Esping-Anderson, quoted in Perry and May, 2011, p. 15). A technological exurbia as a concept then, could be described as the fringe marginalized space of knowledge societies.

3. CONCLUSION

A technological exurbia is an everyday constructed and negotiated space, a space in its own right, and as Kurki (2014) affirms, this border crossings outside of the city, also function as leverage and source of empowerment. Could KBUD acknowledge spatialities like these? Spatialities that represent a critical voice against the ‘homogenizing gaze’? Would not this be more enriching for knowledge construction and exchange? I believe so, and in conclusion, a focus on marginalized space, on technological-knowledge exurbias, can inform and increase KBUD spaces of engagement for further research and understanding. Cities aren’t wholes; they are not available con demand, so why should they be treated by KBUD as such?

REFERENCES

- Amin, A., & Thrift, N., 2002. *Cities: reimagining the urban*. Blackwell Publishing, Malden, MA.
- Bindé, J., 2005. *Towards knowledge societies: UNESCO world report*.
- Carrillo, F. J., Yigitcanlar, T., García, B., & Lönnqvist, A., 2014. *Knowledge and the city: concepts, applications and trends of knowledge-based urban development* (Vol. 52). Routledge, London.
- Castells, M., 2011. *The rise of the network society* (Vol. 12). John Wiley & Sons, New York.
- Glaeser, E. L., 2011. *Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier*. Penguin, New York.
- Kurki, T., 2014. Borders from the Cultural Point of View: An Introduction to Writing at Borders. *Culture Unbound: Journal of Current Cultural Research*, 6(6), 1055-1070.
- Perry, B., & May, T., 2010. Urban knowledge exchange: devilish dichotomies and active intermediation. *International Journal of Knowledge-Based Development*, 1(1-2), 6-24.
- Souter, D., 2010. *Towards inclusive knowledge societies: A review of UNESCO's action in implementing the WSIS outcomes* (p. 66). Paris: UNESCO.
- Spectorsky, A. C., 1955. *The Exurbanites*. Lippincott: Philadelphia.
- Taylor, L., 2011. No boundaries: Exurbia and the study of contemporary urban dispersion. *GeoJournal*, 76(4), 323-339.
- Yigitcanlar, T., 2011. Position paper: redefining knowledge based urban development. *International Journal of Knowledge-Based Development*, 2(4), pp. 340-356.

PROJECTION OF KNOWLEDGE SOCIETIES TOWARDS THE PAST

Arnulfo Felix Jr.

ITESM Monterrey , 9065 Sydney Ct. Unit 11406, San Diego, CA. 92122-1752, USA

ABSTRACT

This reflection paper examines the role of History in Knowledge Societies. It simultaneously follows two parallel premises. First, it discusses Knowledge Societies' reactions derived from the idea of a non-existent present as theorized by German writer and statesman Johann Wolfgang von Goethe. The fact that this unformed future, which cannot be reached until the future comes to fruition, has been overlooked in the development of Knowledge Society theories. It also looks at the dilemma that Knowledge Societies are not self-sufficient sciences, and thus they must be examined within the context of an already recognized science, such as History, to develop exact, stable and fully-explored procedures. My argument suggests that much of what we today call "Knowledge Societies" was not bred by a KS paradigm but rather by a projection towards the past that is understood by way of a historical approach.

KEYWORDS

History, Knowledge Societies, Future; Present, Self-sufficient, Sciences

1. INTRODUCTION

In his 1953 initiatory and renowned treaty, *The Historians Craft*, Marc Bloch¹ states that "whenever our exacting Western society, in the continuing crisis of growth, begins to doubt itself, it asks itself whether it has done well in trying to learn from the past and whether it has learned rightly." (p. 11) The Occident still finds it untenable to revive all the 'states of the soul' parted by vast distances in time; it is clear that rationality hasn't, altogether, pervaded the structured composition that includes both knowledge and memory.

Nonetheless, History is not merely the 'science' of the past. Bloch (1953) eloquently asked, "how can one make of phenomena, having no other common character than that of being accommodated in the past, the matter of rational knowledge?" (p. 22). History is, without a doubt, the will to know 'better'; restricting ourselves to the idea that sciences are the plain and concise description of how are they practiced now as well as how they were performed in the past, requires us to give in to a fair amount of self-deception. In doing so, we may consider it more reasonable to suggest that History would improve itself over time. However, we find ourselves captive in this perennial and unending spectacle that stretches across time. As past transforms into present, we seek not only to unite the two in some sensible, rational manner, but also strive to see a connection with, even a prediction of, what the future holds. Will there not be some recognizable progress draped across the tapestry of History, where a universal and repetitive macro-pattern is seen as progress?² "Could human beings" writes Lucien Febvre (1982), "ever be forced by perfectly established historical laws?" (p. 27)

¹ Marc Bloch was a French historian who cofounded, alongside Lucien Febvre, the highly influential Annales School of Historiography.

² Oswald Spengler in *The Decline of the West* of 1922, recalls these ideas by undertaking, for the first time, the attempt to predict history.

2. BODY OF PAPER

Out of pleasure, and indeed the necessity, History never ceases to tell its own story. “To have been” is a prerequisite for “being.” Reflecting on the role History plays often facilitates the comprehension, even the prediction, of the future. In turn, when applying such an analysis to the “Societies of Knowledge” one accomplishes no easy task, recognizing that the actual inquiry being posited is the very questioning of the value of History itself.

It is advisable, in principle, to outline the subject of this study. According to Stehr & Böhme (1986), “The historical emergence of ‘knowledge societies’ does not occur suddenly; it represents, not a revolutionary development, but rather a gradual process during which the defining characteristics of a society change and new traits emerge” (p. 7). Relatively modern societies from 150-500 years ago, were driven, primarily by a need to work the land, perform useful labor and operate unsophisticated machinery. However, recently, a new sector of our societies, a new sector of human exercise, specifically, science, has emerged as an alternative and complement to that previous focus on land, labor, and machinery. Science, for obvious reasons, implies knowledge, and knowledge for Stehr & Böhme (1986) has always had a leading role in social life. Firstly, “as an anthropological constant where human actions are knowledge-based” (p. 8), secondly, “as a carrier of power, which, has been frequently based on advantages in knowledge” (p. 8) and, thirdly, through the fact that “societal reproduction is not merely physical reproduction, but in the case of humans, always implies cultural reproduction of knowledge” (p. 8). Knowledge Societies are, most certainly, not scientific theories; neither are they descriptive theories. Instead, they form an explanatory framework that is fed back empirically from experimental models. In actuality, they are a series of actions, and not a science, because they propose ends and offer value judgments. They do not pretend to be, nor do they have the attributes to be explanatory theories. However, this fact does not exclude them from being susceptible to scientific criticism.

Knowledge societies are not self-sufficient sciences. It would be futile to expect that their observed variables could construct repetitive models, for if that were the case, the question of historical prediction would be instantaneously resolved. However, Knowledge Societies are better understood as societies of action, which restrict them to the task of constructing theoretical models. Those models do not contain variables that are self-sufficient or self-defined. This is in contrast to, as stated by English economist G. L. S. Shackle (1972), those sciences that “generate stable, exact and fully explored stereotypes” (p. 62) and are not founded on a strictly empirical basis.

In any case, the problem of the cohesion between History and Societies of Knowledge, given the evolutionary essence described by Stehr & Böhme, is revealed by observing the application of such historical tools to the present and pondering if those same tools could be extended into the past. Looking into the past requires an understanding of a historical context. As stated by Marc Bloch (1937) “if History has frequently been discredited as an instrument of knowledge, it is because, on many occasions, it has required something, which by definition, it could not give” (p. 6). It is true that History has its teachings. However, that’s not to say that determined, constitutive mechanisms of society, such as the utilization of labor and land, will bring about natural, predictive results equally in the present as they did in the past. Indeed, certain types of factors have led, on previous occasions, to specific results; and, if those factors were to be modified, a differing effect would follow. We know that the dominant elements of social life are continually evolving.

Perhaps to affirm that History highlights ‘change’ above all else, is equivalent to condemning History itself as a mere apathetic utility. Bloch (1937) wonders, “What is the practical interest of studying that which can only be seen once and will not be seen again?” (p. 12). Some theorists intentionally and fallaciously attempt to understand the experiences of past knowledge societies based solely on those theorists’ observations set in their own limited, present-day perspective. In reality, this “present” is a fragment of our contiguous past – *Nichts gegenwärtig, alles vorübergehend*³ -. These scholars commit a fatal error in trying to imply, without basis, a practical uselessness of the past when compared to the presumed primer relevance of the present. Such thinking invalidates any scrutiny of the past, for it demonstrates a deficient grasp of real History, and how can one determine the inefficiency of the past without comprehending it in the first place?

It is a worthwhile effort to behold the past and recognize that we now live beyond that transition from a concrete type of society based on labor, land, and machinery. Marc Bloch (1937) maintains that “the past offers us an indispensable field of work because it is the only place that allows us to study complete

³ Nothing present, all future. Johann Wolfgang von Goethe, *Schriften zur Kunst. Schriften zur Literatur. Maximen und Reflexionen*.

experiences and measure the effects to the last extreme" (p. 12). To understand the present, it is necessary, in principle, to turn one's head. However, as we do with the past, we tend to talk about the present without truly knowing it. We can only hope that a genuine and honest study of the past will allow us to, one day, come to comprehend that type of evolutionary thought through which we can identify those breaks of equilibrium, those successions of phases, and those revolutions or schisms. Nevertheless, it is impossible to assess the importance of events whose results still await us in the future. German Economist and contributor to the Austrian School of Economics, Ludwig Maurits Lachmann's 1959, *Metroeconómica*, appears as an epigraph on English Post-Keynesian Economist G. L. S. Shackle's 1972 book, *Epistemics & Economics*. Lachmann proclaimed that "the impossibility of prediction in economics follows from the fact that economic change is linked to change in knowledge, and future knowledge cannot be gained before its time." Which of our historical theories, without detracting from Spengler's effort, has seen the authentic predicament of man face-to-face, and the inexistence of complete knowledge?

There is a cardinal, i.e., a primordial difference between describing what is seen, through the facts, and, 'seeing' what should be described. Adherence to methodology-based historical criticism, such as the one found in Annales School, allows us to witness the diachronic transition, from tangible capitals such as land and labor to intangible capitals as seen in knowledge. From that point, it should be considered advisory to those with a keen eye, that preponderance be given to results over processes. The past cannot be simply discarded because the transitory images – the images of societies organized around labor, land, and machinery – are forgotten under the light of current Knowledge Societies. A historical critic should consider, before offering his verdict, the anachronism that distorts one's interpretation of the past by allowing the present to cloud a proper perspective. As indicated by Marc Bloch (1953), "history should be read, as Maitland said, "backwards." For the natural progression of all research is from the best (or least badly) understood to the most obscure" (p. 45). Let us not forget that the observations that pretend to rise to the 'eternal' - that eternal to which Febvre, Braudel, and Bloch propose - are taken from brief moments in time, which, at most, encompass no more than the lifetime of a single being.

3. CONCLUSION

It is only through observation of the past that the advantages of Knowledge Societies are revealed. The present, as already stated, does not exist and future knowledge cannot be reached before the future comes is realized. Societies of Knowledge surreptitiously, if you will, are indirectly the product of questioning the past. Its projection, without fictitious distinction of epochs, is imposed as a consequence of three main motives: 1) only a study of the past provides the necessary sense of change, 2) just the experience projected towards the past allows us to analyze such different cases in which the effect of the various societal factors is clearly manifested; and, finally, 3) human evolution is a continuous channel whose waves are susceptible to propagate from the past to the future. This enrichment of our observation will only be possible with one condition: it is necessary that historical studies have an ever-greater awareness of their task - and this depends on historians. It is essential that historians keep in touch with the past - the source of all life. The price for experiencing this life is our willingness to put into practice this admirable and almost prophetic description of French Historian Fustel de Coulanges where he declares that "History is not the accumulation of events of all kinds that have occurred in the past. History is the science of human societies."

REFERENCES

- Bloch, M, 1953. *The Historian's craft*. Knopf, New York.
- Bloch, M, 1937. *Que demander à l'histoire?* Paris. Bulletin du Centre polytechnicien d'études économiques.
- Böhme, G., & Stehr, N., 1986. *The Knowledge Society: The Growing Impact of Scientific Knowledge on Social Relations*. Springer Netherlands, Dordrecht.
- Febvre, L, 1982. *Combates por la historia*. Ariel, Barcelona.
- Shackle, G. L. S, 1972. *Epistemics and Economics: A Critique of Economic Doctrines*. Cambridge University Press, London.

Posters

AUTOMATIC REFLECTIVE WRITING ANALYSIS BASED ON SEMANTIC CONCEPTS

Huda Alrashidi and Mike Joy

University of Warwick, UK

ABSTRACT

This study proposes a semantic-based approach for reflective writing analysis to overcome the limitations of existing shallow processing approaches. The semantic analysis examines the meaning of linguistic input that depends mainly on transferring words, phrases and sentences into concepts in knowledge sources, such as WordNet-Affect and analysing the relationships of mapped concepts in the underlying knowledge sources. The proposed reflective writing analysis approach focuses on the efficiency of using affectional and emotional concepts identified in WordNet-Affect to classify text into reflective or non-reflective.

KEYWORDS

WordNet-Affect, Classification, Reflective, Semantic-Based

1. PROBLEM STATEMENT

Reflective writing (RW) involves insight into and mental consideration of learned topics, past experiences and actions. RW has several definitions in the literature – for example, it is defined as a form of conceptual processing with a purpose applied to unstructured ideas in a case with an unobvious solution (Moon, 1999). The considerable effort required to analyse reflective writing has stimulated the need for automatic analysis (Corich, 2011). The problems of existing automatic reflection analysis include the inability to deal with the depth of reflection activities in the text (Ullmann, 2011) because the existing approach depends on the strings alone and ignores the semantic features of the text.

2. SEMANTIC-BASED REFLECTIVE WRITING ANALYSIS

The proposed approach for RW analysis depends on classifying each sentence based on its feature vector constructed with reference to a group of affectional and emotional concepts identified in WordNet-Affect. WordNet-Affect is a lexical resource created in a semi-automatic way by augmenting WordNet with effective labels. Synsets, which a set of words can be regarded as synonyms, in WordNet have each been annotated with one semantic affective label. In particular, the effective concepts representing emotional state are individuated by synsets marked with the A-label “emotion”. There are also other a-labels for those concepts representing moods, situations eliciting emotions, or emotional responses. The feature vector contains a set of features that correspond to the identified concepts in WordNet-Affect. These features are emotion, mood, trait, cognitive and physical state, hedonic signal, emotion-eliciting situation, emotional response, behaviour, attitude and sensation. The proposed approach is illustrated in Figure 1.

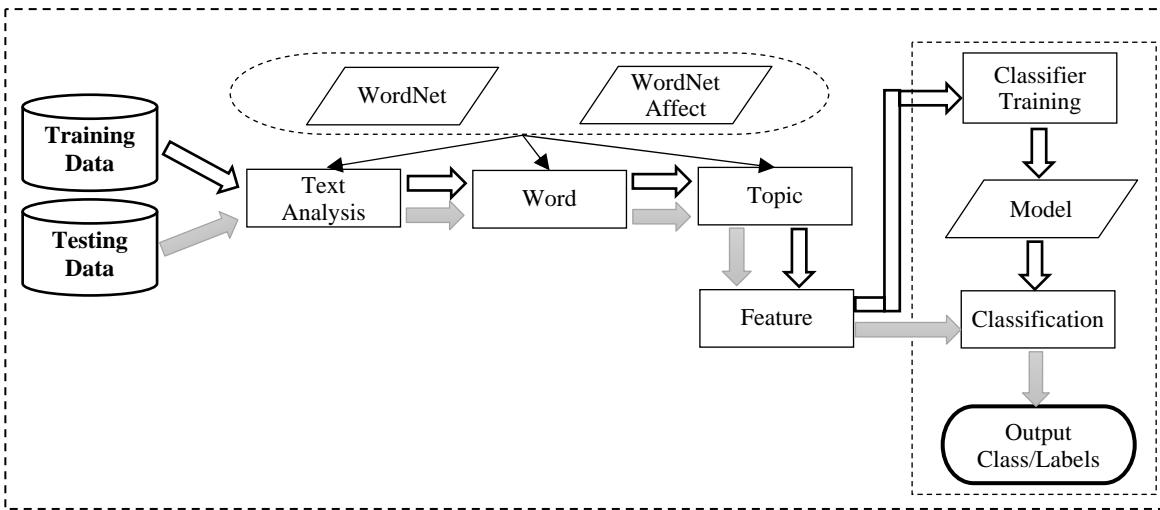


Figure. 1. Semantic-based reflective writing analysis approach

Firstly, the process is commenced by analysing the text and then mapping it to the corresponding concept in WordNet-Affect. Secondly, the word sense is disambiguated. The topic of the text is then modelled with reference to WordNet-Affect. A feature vector of WordNet-Affect concepts is constructed, and the semantic distance between the input sentence and each concept is calculated. Finally, with the use of SVM classification algorithms, the feature vector is used to classify the input text into reflective or non-reflective categories.

Instead of using direct mapping into WordNet-Affect, these synsets are used to model the topic of the text. As reflective analysis depends on sentences, rather than whole text, short sentences classification cannot be implemented directly. Accordingly, each sentence is linked to all WordNet-Affects categories. As the synsets for the input text are extracted, the similarity of each synset with each category in the WordNet Affect is calculated using Lesk measure(Pedersen et al. 2004). Accordingly, each extracted synset in the input text will have a vector of values, whose length equal to the number of categories in WordNet. Given that there are multiple synsets for each category in WordNet-Affect, the similarity is calculated as the maximum similarity with any of these synsets, same in the previous stage.

3. EVALUATION

To evaluate the proposed approach for RW detection, a dataset has been constructed based on the British Academic Writing English Corpus (Heuboeck, 2010). The corpus is formed of a set of student writing assignments in various fields of study, including architecture, chemistry and computer science. Each assignment is graded with M (Merit) and D (Distinction). The corpus involves 13 different assignment formats, including case study, critique and literature survey. This dataset was not created for reflection studies and does not classify the type of text as either reflective or non-reflective.

A single file from each assignment format is selected from various fields of study. Only assignments with the distinction mark ‘D’ are considered in the experiments, as illustrated in Table 1. Each assignment is divided into separate sentences, which have been manually annotated by experts as either non-reflective or reflective. Examples are presented in Table 2. The total number of sentences used in the experiments is 979. The non-reflective and reflective categories have 529 and 450 sentences respectively.

The proposed RW detection approach is compared with the string-based approach, which is derived from the set of keywords listed and experimented by Ullmann (2015). The string-based approach depends on string indicators and uses matching processes to flag sections of text containing reflective material. The results of the proposed approach and the string-based approach are summarised in Table 3. The results of the proposed approach slightly outperform those of the string-based approach. Aside from possessing accuracy, the proposed approach can be extended to analyse the content because it is based on semantic concepts.

Table 1. Description of the constructed dataset

Format	Field	Language	Mark	# Sentences
Case Study	Engineering	English	D	42
Critique	Computer Science	English	D	105
Design Specification	Computer Science	English	D	115
Empathy Writing	Engineering	English	D	39
Essay	Economics	English	D	100
Exercise	Computer Science	English	D	64
Explanation	Engineering	English	D	49
Literature Survey	Philosophy	English	D	20
Methodology Recount	Engineering	English	D	31
Narrative Recount	Engineering	English	D	44
Problem Question	Engineering	English	D	114
Proposal	Engineering	English	D	140
Research Report	Economics	English	D	116

Table 2. Example sentences in the dataset

Sentence	Category
The operating system controls the allocation of memory to programs.	Non-Reflective
I believe that the Spiral model of the software development process would have been better suited to the type of Information System the UKPA was looking to develop.	Reflective

Table 3. Reflective text detection result comparison

	RW Approach	String-based
Accuracy	0.616	0.606
Precision	0.627	0.547
Recall	0.715	0.667
F-Measure	0.668	0.664

4. CONCLUSION

This paper presents an automatic RW detection approach based on semantic concepts. The proposed approach involves mapping words into concepts and then creating a feature vector for each tested document. The feature vector is then used as input to the classification algorithm, which labels the text as either reflective or non-reflective. The concepts in WordNet-Affect are evaluated and analysed to demonstrate their effects on classification and labelling tasks. The result shows that WordNet-Affect for RW detection outperforms the string-based approach reported in the literature. Future work will be devoted to evaluating other concepts related to WordNet-Affect for RW fine classification.

REFERENCES

- Corich, Stephen Paul., 2011. Automating the measurement of critical thinking in discussion forums, *PhD thesis, Massey University, Palmerston North, New Zealand.*
- Heuboeck, A.H., 2010. *The BAWE corpus manual for the project entitled, ‘An Investigation of Genres of Assessed Writing in British Higher Education’*, version 3. 2010.
- Moon, Jennifer A., 1999. *Learning journals: A handbook for academics, students and professional development* Routledge.
- Pedersen, T., Patwardhan, S. and Michelizzi, J., 2004, May. WordNet:: Similarity: measuring the relatedness of concepts. In *Demonstration papers at HLT-NAACL 2004* (pp. 38-41). Association for Computational Linguistics.
- Ullmann, Thomas Daniel., 2011. An architecture for the automated detection of textual indicators of reflection, *In 1st European Workshop on Awareness and Reflection in Learning Networks*, Palermo, Italy, pp. 138-151.
- Ullmann, Thomas Daniel. 2015. Keywords of written reflection-a comparison between reflective and descriptive datasets, *In CEUR Workshop Proceedings.*

LIFELONG LEARNING ON TELEMEDICINE AT MEDICAL INSTITUTE OF RUDN-UNIVERSITY

Elena Lukianova, Valery Stolyar, Tatiana Lyapunova,
Ekaterina Shimkevich and Maya Amcheslavskaya
RUDN-University, Moscow, Russia

ABSTRACT

In this article three educational programs on telemedicine are described. These programs allow obtain theoretical knowledge and practical skills of videoconferencing and distance education methods. The Department of Medical Informatics created the educational module "Telemedicine" for students of 4-5 courses. Two professional distant programs ("Telemedicine Technologies in Healthcare Practice" and "Telemedicine in the health care system") for physicians have been developed and are ready for inclusion in the list of educational programs posted on the Portal of Continuous Medical and Pharmaceutical Education of the Ministry of Health of Russia.

KEYWORDS

Telemedicine, Videoconferencing, Interactive Education, Distance Learning

1. INTRODUCTION

Renewed Federal Law (No. 242-FZ of July 29, 2017)¹ and normative acts of Russian Ministry of Health will create the necessary legal conditions for organizing and conducting consultations using telemedicine technologies.

Telemedicine consultations have a number of advantages, including the possibility of obtaining a "second opinion" in the best world clinics for clarifying the diagnosis and determining the optimal tactics for treating as well as the possibility of transferring a patient to the region during a teleconsultation with local specialists with the ability to monitor the course of rehabilitation

Telemedicine is actively being introduced into the doctors' practice, however, in order for it to become an effective tool in their hands, a basic knowledge of the possibilities and limitations of modern telemedicine technologies is needed, as well as practical skills in the preparation and conduct of video conferencing.

In Medical Institute of PFUR, the conditions necessary for the implementation of continuous medical education programs (CME) - a simulation training center where physicians can improve their practical skills and the Telemedicine Center for video conferencing equipped with all modern ITU standards - have been created.

2. EDUCATIONAL PROGRAMS ON TELEMEDICINE

The Department of Medical Informatics developed new curriculum module "Telemedicine" for students of 4-5 courses. It's designed for 36 academic hours (1 credit ECTS), 17 hours of which are for practice, and the rest is reserved for students' self-education and self-preparing. Also two distant professional educational programs for physicians have been developed and are ready for inclusion in the list of educational programs posted on the Portal of Continuous Medical and Pharmaceutical Education of the Ministry of Health of Russia.

¹ The Federal Law "On Amendments to Certain Legislative Acts of the Russian Federation on the Application of Information Technologies in the Sphere of Health Care" (No. 242-FZ of July 29, 2017) (in Russian)

2.1 Module "Telemedicine" for Students

This module was developed to teach students to use distance technologies in health care practice. Mastering of the module "Telemedicine" will allow to form the ability to analyze and be ready to act in non-standard situations and to be socially and ethically responsible for the decisions made. The main professional competences that are formed is the ability to use modern management methods to solve medical diagnostic problems.

Students will be able to competently formulate a qualified request for teleconsultation. During the teleconsultation they can to present the patient to the expert consultant and give exhaustive answers to his questions, to take part in conclusion's composing for a treatment plan, to organize the registration of the results of teleconsultation. They will be able to master the methods of preparing an electronic extract from medical-history and its transfer to a specialist-consultant and methods of organizing the interaction between the attending physician and the consultant in the teleconsultation process.

Training module "Telemedicine" includes the following topics: the fundamentals of telemedicine, and the world trends in its development; technological equipment of telemedicine events; hardware and software of telemedicine; economic and legal aspects of telemedicine; scenarios of telemedicine activities.

After theoretical lectures, students receive practical skills in the course of business games in preparation and conduct of video conferencing. During the classes we demonstrate to students the technologies of remote interactive learning, in particular television lectures and master classes from the leading clinics of Russia, countries of Europe, India, Brazil and Canada.

2.2 Distant Program "Telemedicine Technologies in Healthcare Practice"

Students will know basic concepts and goals of modern telemedicine, basics of the legislation concerning the use of telemedicine technologies, their current status in the world, the advantages and disadvantages of their use. They will know about the Internet portal as an environment for organizing telemedicine events and about the technical means of mobile telemedicine, will understand nuances of main forms of work in telemedicine: consultation, lecture, seminar, master class, scientific and practical conference, patronage, monitoring, supervising are considering. Since videoconferencing is the technological basis of telemedicine, the standards of encoding/decoding information, image and sound quality are considered in detail. Also, students will get knowledge about hardware and software of the telemedicine consulting center and the methodology for calculating the costs of its maintenance. Students receive basic knowledge of the legal and economic relations of subjects in telemedicine

2.3 Distant Program "Telemedicine in the Health Care System"

It is advanced program for those who already is familiar with telemedicine technologies. It includes special sections of telemedicine: tele-radiology and tele-cardiology, the protection of personal data during telemedicine activities, including remote access to the medical information system (MIS) for personal telemedicine. Also we pay attention for special standards of storage and transfer of graphic information about patients, the principles of Picture Archiving and Communication System (PACS), which is used for the transfer and archiving of medical images and documents in the Digital Imaging and Communications in Medicine (DICOM) standard. Since pathomorphological examination is carried out at a distance using a video monitor instead of light microscope image, attention to the technological equipment for telepathology is paid.

2.4 Technical Equipment for Educational Programs

At the Telemedicine Center of RUDN-University a class with videoconferencing, designed for 50 listeners is created. It is equipped with video conferencing complex. The camcorder provides Full HD video quality with a resolution of 1080p at 60 fps, a wide angle of 73 ° for capturing all participants in the audience, with 12x optical zoom and remote control of the camera's zoom, pan and tilt functions. The class has a professional document camera with built-in illumination plate for displaying x-rays with 240x magnification (16x optical and 15x digital), professional A3 flatbed scanner also allows to scan X-rays.

2.5 Methodology of Training

Videoconferencing tools help to solve the problem of access to the experience and knowledge of leading world specialists. Modern development of digital telecommunications gives a possibility to work in high-definition format and tele-training has moved to a new level and includes: cycles of remote telelections in frame of continuous medical education; interactive master classes with on-line operations and diagnostic procedures from the best Russian and foreign clinics; remote supervising of work of young doctors by experienced specialists ; remote interactive participation in international scientific and practical conferences on a wide range of medical problems.

3. CONCLUSION

Described teaching approach to study basics of telemedicine allows students to obtain theoretical knowledge and practical skills of video conferencing and distance education methods, to get familiarized with international experience and trends in the development of telemedicine technologies. Using videoconferencing for personification of interactive distance learning in CME, which will allow physicians to study at the workplace is perspective for distance learning development on the basis of modern telecommunication technologies

ACKNOWLEDGEMENT

The publication was prepared with the support of the “RUDN University Program 5-100”

REFERENCES

The Federal Law "On Amendments to Certain Legislative Acts of the Russian Federation on the Application of Information Technologies in the Sphere of Health Care" (No. 242-FZ of July 29, 2017) (in Russian)

AUTHOR INDEX

- Albán, M.....291
Almjally, N.319
Almutairi, F.11
Alrashidi, H.383
Amcheslavskaya, M.386
Angula, N.281
Annunzio, G.112
Araujo, F.185
Arbelaitz, O.83
Arndt, T.311
Arroyo-Blanco, E.265
Bardaki, C.55
Baroudi, S.47
Baumgartner, T.209
Belkhayat, N.330
Belkhayat, N.330
Besimi, A.273
Besnard, A.30
Bezerra, M.103
Bolts, G.296
Bontchev, B.193
Bottis, M.224, 341
Bouchagiar, G.224, 341
Bounabat, B.67
Boza-Monge, A.265
Brucker-Kley, E.112, 209
Buenaño-Fernández, D.301
Caballero-Gil, P.365
Chennouf, J.330
Cortés-Morales, R.265
Curvello, P.103
Darwish, E.247
das Chagas, M.315
De La Haye, F.30
De Rolt, C.129, 347
Degenhart, S.296
Delaere, S.92
Delgado-Cepeda, F.23, 40
Dlodlo, N.171, 281
Elizondo, L.373
Eppard, J.47
Farrugia, S.201
Felix Jr., A.377
Gauthier, F.129, 347
Guelmami, Z.361
Guo, S.353
Haj-Ali, A.3
Hardman, V.231
Hashi, A.47
Hasse Klein, G.129, 347
Hasse, D.129, 347
Helder, M.47
Hussein, B.3
Ivanov, D.193
Jordaan, D.306
Joy, M.319, 383
Karon, R.171
Keller, T.112, 209
Koiranen, I.120
Koivula, A.120
Kononova, O.147
Kotsopoulos, D.55
Kurtz, L.325
Lavandier, K.30
Le Bohec, O.30
Lemos, A.325
Lohmann, P.103
Lounis, S.55
Luján-Mora, S.301
Lukianova, E.386
Lyapunova, T.386
Marino, V.369
Mohr, S.257
Moreira, G.369
Moussalem, H.3
Muguerza, J.83
Mulder, B.163
Musmarra, P.178
Mzuku, K.239
Nazifi, S.273
Nicolle, F.361
Noël, Y.30
Nogues, J.30
Oriez, S.30
Özkoç, E.217
Pacheco, I.291

Papaioannou, T.	55
Patterson, D.	231
Pavlovskaya, M.	147
Pellanda, N.	315
Perona, I.	83
Pramatari, K.	55
Quaireau, C.	30
Raad, M.	3
Ranaivoson, H.	92
Raso, L.	369
Raspante, L.	369
Renzi, A.	185
Rhazale, A.	67
Roast, C.	231
Rodríguez-Pérez, N.	365
Ryoo, B.	296
Salazar, M.	291
Sampaio, A.	155
Santos, B.	369
Sarlak, M.	139
Seychell, D.	201
Shimkevich, E.	386
Souto Maior, A.	325
Stolyar, V.	386
Tavallaei, F.	139
Tomczyk, L.	335
Valério, L.	369
Van Belle, J.	239
Van Gorp, A.	163
Vassileva, D.	193
Veloz, M.	291
Veress, J.	75
Villegas-Ch, W.	301
Villessèche, J.	30
Virkki, K.	103
Vizcaíno, G.	291
Vuorinen, J.	120
White, S.	11
Yera, A.	83
Yescas-Martínez, F.	40
Zouain, D.	103
Zumstein, D.	257