A Framework for Classification of Social Machines

Authors
Web and Internet Science Group
University of Southampton
Southampton, UK
{a,b,c,d,wh,nrs}@ecs.soton.ac.uk

ABSTRACT

This paper provides a framework for...

1. INTRODUCTION (0.5 PAGE)

What are social machines and why are they important

An emerging, interdisciplinary field of research and development with a growing number of social machine instances and studies about them

We propose a classification framework for social machines. Targeted audiences

* research community: shared terminology, set of key features and their possible instantiations based on the analysis of a relevant set of 50 (?) social machines (both popular systems as well as long tail systems); facilitates the study of design and evolution patterns; supports research on how communities emerge and develop and organisational studies (mechanism design, incentives); different research approaches can be better compared and aligned.

* developer community: provides a theoretical grounding for the realisation of new systems, informs system design and operation (what features should be implemented, what features are more important than others); enables predictions on user behaviour and community development.

The framework is based on interviews with XXX researchers covering XXX systems and was preliminary evaluated with respect to ...

2. METHODOLOGY (1 PAGE)

In order to identify and classify social machines, we first embarked on a principled knowledge elicitation process [11], to determine the *elements* (i.e., the social machines) and the *constructs* (i.e., the classifying factors) that we can use to classify social machines. We used the *repertory grid* elicitation technique [5] in order to derive an initial set of elements and constructs. We asked ten Computer Science researchers familiar with social machines to create their own repertory grids, and populate the elements from their own knowledge, and to create the constructs using the standard repertory grid elicitation technique. This exercise led to 10 grids, the union of which comprised a total of 56 unique elements (social machines) and 117 different constructs.

To make the set of social machines easier to refer to and more approachable, we performed an initial grid analysis of the machines, clustering them into related user-centered themes, of: Social Networks, (Micro)Blogging, News Aggregators, Image Boards, Crowd Science, Answer Gardens, Community Watch, Health and Wellbeing Support, Action and Investigation, Opinion Sharing, Video Sharing, Photo Sharing, Code Sharing, Art Sharing, Crowdsourcing Platforms, Mash-up Systems, Crowdsourcing Toolkits and Platforms.

While determining the intersection of elements was straightforward, the consolidation of the constructs required a more thorough process. To consolidate the constructs we embarked on a manual process of grouping the constructs into rough clusters, based around the areas they cover. For example, we used construct clusters including: Users, Motivation, Popularity, Technology, and Purpose. We then examined each construct to determine which were equivalent, and whether we could re-word or subsume existing constructs to cover the same aspects as a construct, with the aim to remove any redundancy, to cut the list of constructs down to a manageable number, but to ensure that all of the constructs that were elicited were somehow represented in the final set. This process involved four of the authors discussing the choices, and finally agreeing on a consolidated set of 31 constructs.

Even with this reduced set of constructs, if a single person were to classify all 56 elements, this would still require 1736 individual decisions, and hence is unfeasible to expect full coverage of the elements using all of these constructs, despite how representative they are.

[something here about doing 2 grids each and verifying the results?]

To extract greater usefulness, and the ability to use our constructs to classify social machines we re-evaluated the constructs and the semantics that they represent into richer constructs that are not constrained by the 1–5 rating of repertory grid constructs.

3. CLASSIFICATION FRAMEWORK (2 PAGES)

To create a classification framework, it is necessary to first determine the aspects of classification. We will first discuss the relationships between social machines, enabling technologies and platforms; we move on to discuss the multifaceted nature of motivation within social machines; next we

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Social Machines Workshop 2013, Paris, France
Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...$15.00.

evaluate the artefacts of social machines: contributions, actions, activities and purpose; finally we discuss the metrics of participation, and the difference between designed and emergent properties of social machines.

3.1 The Hierarchical/Polyarchical relationship of social machines

While determining a set of social machines, one clear distinction was between social machine frameworks, such as MediaWiki and Ushahidi, which enable social machines to be created, and instantiations of those frameworks into social machines that have been created for a specific purpose, such as Wikipedia and Ushahidi Haiti. However, through further investigation, we discovered that the distinction between framework and instantiations was not the only differentiator. Specifically, that some social machines could also have sub-machines or communities within them, that are working to solve different problems, within the context of the larger machine. For example, Wikipedia is an instantiation of the MediaWiki framework, but there are a large number of communities within Wikipedia that work on domain-specific areas, such as those contributing in areas of science and technology[1]. However, we must also note that wikipedia is primarily split into communities of language, with the English, German and French wikipedias having the largest number of articles. It is therefore possible to analyse wikipedia into a hierarchy, from MediaWiki, to the languages of Wikipedia, down to the domain-specific communities. We explored this idea further, to include other social machines, and links where social machines utilise other social machines in order to function. For example, the social machine used by the Obama '12 US Presidential campaign [10], which relied on social machines such as Twitter and Facebook.

Looking at the set of social machines in a polyarchy leads to a broad/specific relationship emerging that lets us talk about behaviours at various levels of granularity. We propose looking at nested machines with "The Web" as one of several potential roots, with the next level down consisting of sub-platforms (such as Facebook, Twitter, MediaWiki, Ushahidi platforms) that spawn more specific social machines. A resulting polyarchy is shown in Figure 1, which classifies levels as Infrastructure, Frameworks, Services and Causes/Groups.

This approach enables us to start with a more detailed analysis of certain levels over others; and seeing what similarities flow up and down the polyarchy. For example, what do specific instances of Ushahidi/Zooniverse/MediaWiki have in common with other instances, and how to they differ dynamically?

3.2 Multi-faceted motivation, and who-gets-what benefit

We identified a hierarchy of constructs pertaining to the motivation people have for using social machines with top level constructs of:

1. Hedonistic
2. Financial
3. To be informed
4. To help others

The sub constructs are illustrated in Figure 2, and we relate them to Maslow's hierarchy of needs [7], (e.g., self-actualisation, security, respect of others), in various ways.

For each of these motivations, different roles that participate may have different *core motivations*, such as:

1. Benefit to contributor
2. Benefit to moderator
3. Benefit to system operator/host (e.g, Google, Facebook, Amazon)
4. To affiliates of the host (e.g., Amazon affiliates, YouTube partners)
5. To society as a whole
6. To a contributor's social network

In addition to relationships to Maslow's hierarchy of needs, which drive the motivations, we suggest that there are also resulting actions from each motivation. Thus, it should be possible to follow basic human needs to motivations in social machines, through to the actual actions (and therefore the necessary software feature support) that are made by participants in social machines. Our theory is that this path may enable designers of social machines to satisfy human needs by implemented specific features into their social machines, or to identify potential gaps in their serving of needs and motivations due to lack of implemented features.

3.3 Artefacts of Social Machines

Our next observation was that there are a number of shared artefact types that can be classified within the context of social machines. Specifically that there are "intentions" that motivate the use of social machines, "actions" that are directly performed by participants of social machines, and "contributions" that result from the use of social machines. We have taken these abstract concepts, and mapped them to a number of well-known social machines in order to demonstrate that these concepts can be used to compare and contrast different machines.

[more here]

3.4 Metrics of participants and usage

Finally, we identified several constructs that are relevant for all social machines. First, there are construct that relate to "analytics" that can be measured at a point in time, and may change over time. We have documented these constructs in Table 1. In the left column we list construct regarding the design affordances of the social machine. In the middle column we have a mixture of the result of how the design affordances actually were realised by the social computation, such as how users co-opted / incorporated / interpreted the constraints. Finally, in the right column we list analytical measures of social machines which are subject to change over time.

The result is what we call 'participation' constructs. These constructs arise from use of a social machine, and are not necessarily predictable; they include:

[1]Visualisation of Wikipedia Science Communities: `http://www.olihb.com/WikiCommunities/`

Designed Affordances
Domain specificity [1 - Specific / 5 - General] vs Generality of Purpose
Visibility and importance of a participants 'reputation' [1 - Not important / 5- Important] Explicit representation of participant reputation [1 - No Reputations / 5 - Reputations]
Participants can define new types of contributions [1 - Not possible / 5 - Complete freedom]
Degree of open source software used/created [1 - Proprietary / 5 - Entirely open source]
Degree of openness and availability of data created in the system (APIs/Dumps) [1 - Closed/unavailable data / 5 - Completely open data]
Variety of actions participants can perform [1 - Single Action / 5 - Large variety of actions]
Participatory Affordances
(Level of social features) - Supports social interaction [1 - No social features / 5 - Large variety of social features]
How often a participant participates (times a day, etc) [1 - Many times a day / 5 Seldom]
Importance of Timely Participation [1 - Unimportant / 5 - Timely]
Generality of Audience [1 - Niche / 5 - General]
Participant anonymity [1 - No Anonymity / 5 - Complete Anonymity]
Participant autonomy (1 - No autonomy / 5 - Complete autonomy)
Quality of participant contributions [1 - Low quality / 5 - High quality]
(Different roles clearly defined) - Clear separation of roles/responsibilities among participants - [1 - Single role - Everyone is the same) / 5 - People assume different roles/responsibilities]
Extent of hierarchical organisation of roles [1 - No Hierarchy / 5 - Deep/Well-defined Hierarchy]
Variety of types of contributions [1 - Single Type / 5 - Large variety of types]
Participants are motivated by extrinsic award [1 - Disagree / 5 - Agree]
Participants are intrinsically motivated [1 - Disagree / 5 - Agree]
Service/Platform derives benefit from participant participation [1 - Disagree / 5- Agree]
Participants directly benefit from participating [1 - Disagree / 5 - Agree]
Participants' friends/social network benefit from participation [1 - Disagree / 5 - Agree]
Participation is for the benefit of a specific person/group other than the participant or platform owner [1-Disagree / 5-Agree]
Participation is for the benefit of society/the world at large [1 - Disagree / 5 - Agree]
Participation is done to 'get something done' [1-Disagree/5-Agree]
Participation is done 'for fun' [1-Disagree/5-Agree]
Participation is done to exchange knowledge [1-Disagree/5-Agree]
Participation is done to be social [1-Disagree/5-Agree]
Participation is done via mobile devices [1-Never / 5 - Often]
Participants' (geographic) location is used by the service [1-No / 5 - Very]
Analytics
Popularity [1 - Unpopular / 5 - Very Popular]
Maturity [1 - New / 5 - Mature]
Ratio of passive to active participants [1 - Passive / 5 - Active]
Number of competitors [1 - Not Many / 5 - Many]
Extent of global geographic coverage [1 - Highly localised / 5 - Global]

Table 1: Constructs of social machines, listing affordances that are designed into machines, affordances that have resulted from use, and analytical constructs that can be measured and change over time.

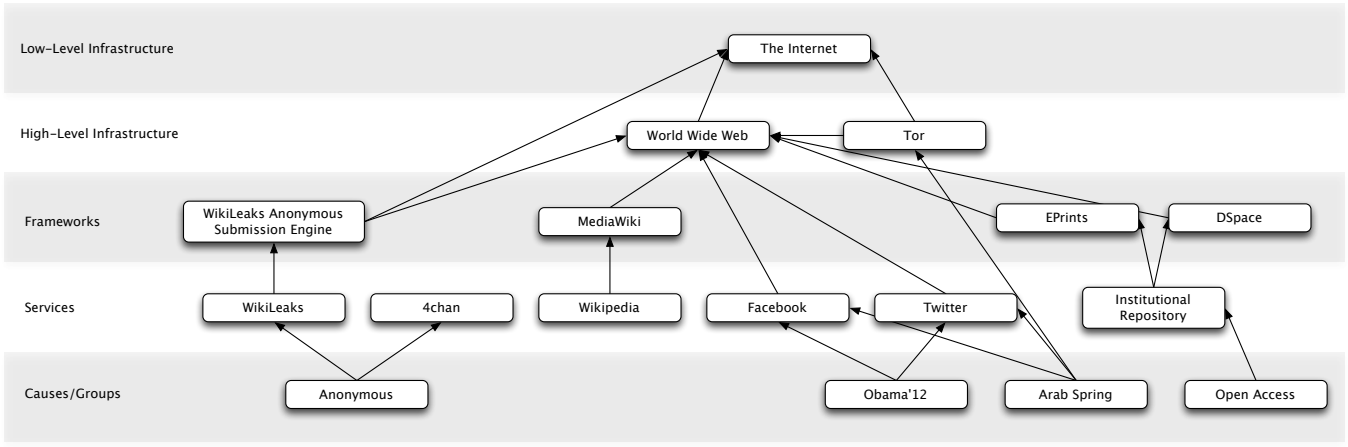

Figure 1: A polyarchy of Social Machines, illustrating the infrastructure and frameworks used by social machines, and machine-machine usage.

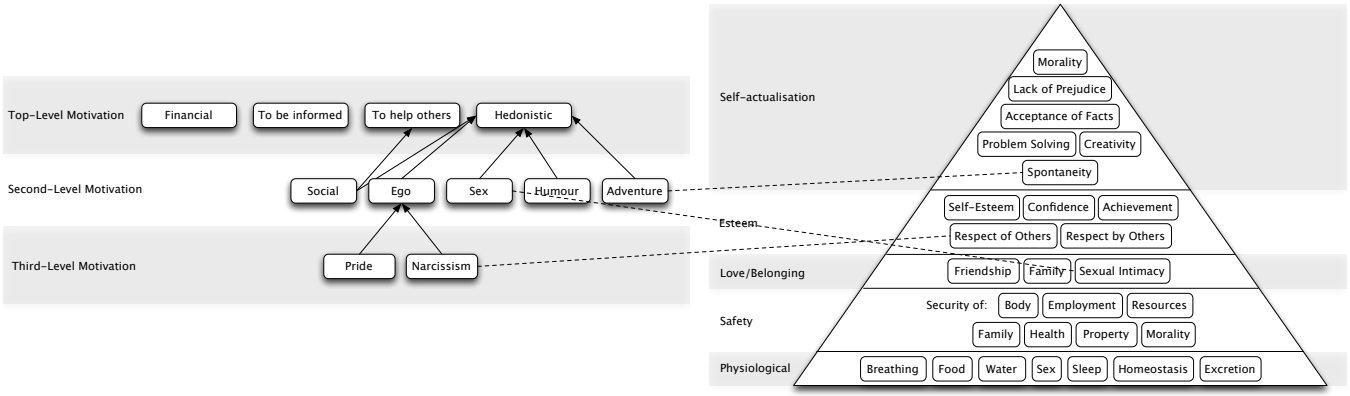

Figure 2: Motivation hierarchy of participants using Social Machines, related to Maslow's hierarchy of needs.

1. **Roles emerging in the social machine**
   In addition to designed and implemented roles such as *moderator* and *administrator*, emerging roles include motivations of the participant, such as *trolls*, *spammers*, *domain experts*, *social nexus*, and *re-blogger*.

2. **Quality of contributions**
   For any social machine, the quality of contributions will vary. While some issues with quality can be predicted and protected against (e.g., through multi-user result agreement on Mechanical Turk [4]), some particular issues with quality cannot always be predicted ahead of time.

3. **Content of contributions**
   While a system operator may design their social machine to generate specific types of content, participants may subvert the system in order to generate other types, or to concentrate on areas that were unforeseen at the time of launch, but are now seen as useful to the operators. It is also possible that multiple social machines can be combined/linked by users to further expand the types of contributions they produce.

4. **Global / cultural reach**
   A social machine may naturally focus on a narrow geographic area, for example an Ushahidi-based social machine for local elections [8] or natural disasters [9]. Such machines are unlikely to be re-purposed by participants outside of that area. However, some machines have been created for a single cultural/global area, but have potential global relevance that was unforeseen early on. An example is the fall-off of English-language users of the social network "Orkut," which gained significant usage elsewhere, particularly in Brazil.

5. **Task reach — what do people use it for?**
   In addition to breaking cultural barriers, user have also shown that they can use social machines for tasks that were not set up by the system operators. For example the identification of "Green Pea Galaxies" using the GalaxyZoo social machine for astronomy [1]. Users were not originally asked to identify these types, but a number of users identified them, and the software was modified to include them, resulting in their discovery as a new type of galaxy.

6. **Active/Passive participation roles**
When it comes to participation in social machines, there are levels of participation that can be classified as being "active" and being "passive." These roles may not be particularly designed into systems, but may be exhibited over time. Particularly noteworthy examples can be seen in news and link aggregation social machines such as Reddit and Digg, where users can both submit links and vote on links. The frequency of each of these activities can vary wildly, with a small minority of users controlling the links that rise to the top [6].

4. USAGE EXAMPLES (0.5 PAGE)

In order to demonstrate usage of our framework's constructs, we ranked our social machines by their Alexa ranking, and performed a full repertory grid elicitation exercise. The results are illustrated in a grid and accompanying dendrograms for element and construct similarity in Figure 3.

4.1 Causes as Social Machines

One originally unforeseen classification is that of the causes that utilise Social Machines. Through a particularly popular cause, large numbers of participants can be mobilised, for either a short period of time, or for long-term engagement, using multiple social machines in order to further their cause. Examples of causes can be seen in Figure 1, at the lowest level of the polyarchy. One such example is the cause of "Open Access" to academic publications [2]. This cause mobilises academic authors to self-archive their own papers online so they are available for free, in addition to hosted versions behind publishers pay-walls. In order to achieve this aim, the social machine of "open access" typically uses the social machines of institutional repositories provided by the authors' associated institutions. As with other social machines, these social machines typically use off-the-shelf implementations (in this case the most popular are EPrints [3] and DSpace [12]), which run as web sites on the world wide web. This cause is mature, growing year-on-year and has global coverage.

[more about this?]

5. RELATED WORK (0.5 PAGES)

existing classification systems in CSCW, collective intelligence, social computing systems

6. CONCLUSIONS AND FUTURE WORK (0.5 PAGES)

Detail evaluation plans

Feedback from different disciplines, and from the developers community

7. ACKNOWLEDGMENTS

This work is supported under SOCIAM: The Theory and Practice of Social Machines. The SOCIAM Project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1 and comprises the Universities of Southampton, Oxford and Edinburgh.

8. REFERENCES

[1] C. N. Cardamone, K. Schawinski, M. Sarzi, S. P. Bamford, N. Bennert, C. M. Urry, C. Lintott, W. C. Keel, J. Parejko, R. C. Nichol, D. Thomas, D. Andreescu, P. Murray, M. J. Raddick, A. Slosar, A. Szalay, and J. VandenBerg. Galaxy Zoo Green Peas: Discovery of A Class of Compact Extremely Star-Forming Galaxies. *Monthly Notices of the Royal Astronomical Society*, 399(3):1191–1205, 2009.

[2] S. Harnad. The Self-Archiving Initiative. *Nature*, 410:1024–1025, 2001.

[3] S. Hitchcock, T. Brody, C. Gutteridge, L. Carr, W. Hall, S. Harnad, D. Bergmark, and C. Lagoze. Open citation linking: The way forward. *D-Lib Magazine*, 8(10), 2002.

[4] P. G. Ipeirotis, F. Provost, and J. Wang. Quality management on amazon mechanical turk. In *Proceedings of the ACM SIGKDD workshop on human computation*, pages 64–67. ACM, 2010.

[5] G. A. Kelly. The Psychology Of Personal Constructs. 1955.

[6] K. Lerman. Social information processing in news aggregation. *Internet Computing, IEEE*, 11(6):16 –28, nov.-dec. 2007.

[7] A. H. Maslow. A theory of human motivation. *Psychological review*, 50(4):370, 1943.

[8] P. Meier and K. Brodock. Crisis Mapping KenyaâĂŹs Election Violence: Comparing Mainstream News, Citizen Journalism and Ushahidi. *iRevolution Blog, October*, 23, 2008.

[9] N. Morrow, N. Mock, A. Papendieck, and N. Kocmich. Independent evaluation of the Ushahidi Haiti project. *Development Information Systems International*, 8, 2011.

[10] K. O'Hara. Social Machine Politics Are Here to Stay. *IEEE Internet Computing*, 17(2), March/April 2013.

[11] N. R. Shadbolt and A. M. Burton. Knowledge Elicitation. *Evaluation of Human Work: Practical Ergonomics Methodology*, 1990.

[12] M. Smith, M. Barton, M. Bass, M. Branschofsky, G. McClellan, D. Stuve, R. Tansley, and J. H. Walker. DSpace: An open source dynamic digital repository. *D-Lib Magazine*, 9(1), 2003.

Figure 3: Dendrogram from a repertory grid exercise of the top 20 (ranked by Alexa) social machines from our element set, against our consolidated constructs.