

PUBLIC AND PRIVATE STANCES IN OPEN DATA SERVICE DESIGN

AMALIA DE GÖTZEN
AALBORG UNIVERSITY COPENHAGEN
DEPARTMENT OF ARCHITECTURE, DESIGN
AND MEDIA TECHNOLOGY
AGO@CREATE.AAU.DK

ABSTRACT

This paper will highlight some of the problems that participatory design and the design of services are facing while dealing with Open Data. How to build trust among stakeholders? How to preserve the individual privacies? To support the need of design care with respect to the above mentioned issues, two case studies will be briefly presented, showing the added complexity layer that attempted solutions involve. The first case study is the design of a Time banking service, while the second one is about involving the commuters in the maintainance of S-Train Station. Both case studies have been developed in Copenhagen by students from the Service Systems Design Master at Aalborg University Copenhagen.

INTRODUCTION: PRIVACY ISSUES IN THE INFORMATION SOCIETY

One of the main issues in designing participatory services which provide added benefits to users and society at large is the one concerning the thin boundary that divides open data from breaches of personal privacy.

Open Data is a perspective that can provides the most brilliant and visionary opportunities concerning governance and democracy (Buckingham Shum 2012) in our ever expanding “liquid” society (Bauman 2000).

But the cases in which “opening data” clashes with our personal lives and our right to privacy (intended in this paper as individual autonomy over disclosure of personal information) are hardly rare. And for good reason: the acquired “freedom” that our societies exhibit cannot but stress the fact that our individual freedom is of some use only when it is placed in relation with the society that surrounds us, lest it become an existence “solitary, miserable and full of needs” (Maritain 1947).

In order to fully appreciate our modern societies, we need to relinquish a bit our grip over our data as our contribution to society. This goes against our rights to privacy, but most of us are so used to donate bits and pieces of our privacy here and there that we hardly notice the issue any longer.

In some cases this “spontaneous donation” of bits and pieces of our privacy is part of an implicit agreement between us and the service provider. When travelling, for example, security checks systematically infringe individual privacies: they are considered necessary to reach the goal of providing security, but we hardly know what use is being done of the private information we are leaving behind us.

In other cases they are part of our social attitude to share experience and information about our daily life. This happens everyday in our online social networking and it is being used to create alternative forms of organization, where the official organization or information channels are not meeting our needs in a satisfactory way. Social networks can become therefore the backbone for new protest movements or simply a way to control existing private and public services.

The new opportunities offered by open data do not come without concrete risks for our privacy, freedom and democracy.

Participatory service design must use extreme care to this delicate aspect in providing services that allow each individual to be in some form of conscious control of its own private life. Since the implementation of such services often involve many different disciplines and specialties, this care must inform all of them -and this is often harder than it seems at first sight.

SOME DILEMMAS OF OPEN DATA

Legitimate interested stakeholders (such as governments, banks, etc.) are not the only source of trouble when analyzing Open Data flows. There are of course many other much less legitimate parties that may use our private streams of data for malicious purposes.

The current use of vast amount of data available on the internet to accomplish terrorist or other criminal acts is well documented, and it goes much beyond what is usually termed as “cybercrime”: terrorists communicate and program coups on social networks, burglars use public webcams to organize their thefts, while frauds that do not use some form of information technology have practically disappeared. While helping well behaved citizens in their daily lives, Open Data happens to be very useful to these criminal activities too, and there is little that can be done about it without jeopardizing the usefulness of the whole system.

However, this is not the only danger which is deeply rooted in Open Data. Much softer (but all the same annoying) activities can take place using the massive quantities of data publicly available. Stakeholders of all kinds, ranging from pharmaceutical companies to banks, from food producers to travel agencies, can track our habits and general profiles to generate massive amounts of promotional material which may clutter some of our sources of information (such as e-mail or other forms of messaging) to the point of turning them into useless sources of unwanted noise.

And even when the sources are coming from individuals with the best possible intentions they may still be plagued by unwanted errors and noise. Notable cases are the rapidly raising movements of “citizen science” (Chamberlain 2013 and “crowd sourcing” (Papangelis 2014). While these usually provide high quality data and information, it is difficult to be sure that noise will not crop up out of them.

Since precisely noise (as opposed to information) is **the** source of pollution in the information society we live in, substituting carbon oxide or other poisonous substances which constituted the pollution of the industrial society (Castells 2011), it is conceivable to imagine that this

kind of “spam” is a sort of poison which severely boycotts our relation to our surroundings.

In fact, the problem of “noise” in our society is in itself so multi-faceted and complex that it would require a separate treatment.

Furthermore, it is only related to the issues of participatory design using Open Data simply because it undermines the basic structure of our societies. Nevertheless, it is useful in our discussion to introduce the notion of **trust**.

Noise and trust are intimately related by logic. Since noise can be qualified as an “unwanted distortion that degrades the quality of signals and data” it fits well enough the notion that “reputation noise can be considered an unwanted distortion that degrades the quality of information about others” (Falcone 2008).

We inherently cannot trust noisy information, while its sole abundance does not grant any form of quality. And, as in the case of thermodynamics, since noise is comparable to heat exchange in raising the general entropy of information (Shannon 1949), and reducing this entropy requires energy (e.g. refrigeration), reducing noise entropy in information (i.e. building trust) requires energy in terms of (often purely human and manual) labor.

This energy expense may be particularly evident in the activities of (human) moderation and review of specific sets of data, but it is in fact always present when we actually take a decision of some sort based on the data we have available at a given point in time.

THE QUEST FOR SOLUTIONS

One of the most important problems concerning Open Data is whether anyone can reasonably trust any form of data or not. As a matter of fact, the ability of trusting (or mis-trusting, if needed) any information is a vital necessity in our society as much as the capability of creating eco-friendly sources of energy or to purify water was essential for survival in the industrial society.

Trust is fundamental to reduce information--age pollution and in general to augment the signal-to-noise ratio (i.e. the quality) of its main asset -- precisely data.

And a sound theory of trust provides at least one of the key solutions to the problems we have outlined in the previous section.

In the industrial society, trust was based on reputation and accountability (Swift, 2001). Companies of all sorts have spent unlimited budgets and energies in advertising with the specific purpose of building confidence and trust in consumers. This trust was thus based on surface, exterior knowledge: it was based on brand recognition rather than real experience, and vague ideas of quality were mistified for actual knowledge.

In the wake of the late industrial society, (Klein 2009) was just one of the report texts, if the most successful, to show how futile and fragile had such trust become.

In the information society, the old trust model is rapidly fading away. In assessing huge sources of data (such as the ones available through Open Data projects), any human being -- whether it be a common citizen or a specialist, a child or an adult or an elder -- needs to be able to quickly build a sense of trust (or lack thereof) out of actual - and not mistified - data.

Furthermore, reputation and accountability, which were an essential asset in the industrial society, cannot be used any longer because of privacy issues. Therefore, in our society trust is based on **correlation of multi-agent data** (Castelfranchi Falcone 2010).

We build our trust over the information provided by Open Data (and data in general) by actively seeking the congruence of data coming from a multitude of sources, be them cognitive, cultural, institutional, technical, or normative. This kind of trust integrates very well with technical solutions and in fact it can be (and actually is) used in autonomous computing systems as well - thus providing humans with the necessary help in

building it rapidly enough for any practical purpose.

Open Data becomes therefore the problem and, at the same time, the solution. The availability of data from multiple sources is essential to be able to build **our** sense of trust: tapping single sources of information is no longer enough for us. In order to be able to trust data, we need to be able to see it under different (possibly diverging) angles, as it were, evaluating the existing correlations between them.. The success of product evaluation websites or of citizen speciality blogs are a good example of this: rather than simply trusting products based on promotional information, we search for multiple (and often contrasting) sources of information to build trust.

Open Data, along with crowd sourcing, citizen participation, and helped out by smart technical solutions can provide the key to augment the quality of information on which our societies rely upon.

Of course, there is no silver bullet when it comes to decide how many sources are required in order to build up sufficient trust, nor when a malicious attack on data sources is prepared well enough to provide wrong (but plausible) data correlation.

However, these may be considered extreme cases which should not jeopardize the validity of the quasi--totality of the rest.

To summarize: a wealth of data is trustable (and therefore qualitatively relevant) **only** when it is provided by a multitude of (possibly diverging and/or contrasting) sources. Large-scale correlations may be used to reduce informational noise (simply discarding correlation outliers) and to increase data accountability.

No single source of data will be able to grant trusting quality per se any longer.

CASE STUDIES

To illustrate these issues we discuss some case studies from the Service Design master at AAU CPH. The paper will focus on projects that emphasized the opportunities for citizens to create parallel flows of information that support the transport system in and outside Copenhagen or to share resources (time, skills, knowledge) to improve our daily life while strengthening our communities. Both case studies present issues that are discussed from the designer and from the various stakeholders perspectives (commissioner, provider, final user, etc.).

TIME BANKING

The time bank project was developed by four students of the Service Design Master in Copenhagen with the aim of creating a time bank service that could support and enhance an already existing community, the Copenhagen food co-op (KBHFF). An important structural detail of the organization is that it is entirely driven by volunteers.

KBHFF is a member-based and member-driven non-profit food co-operative in Copenhagen, Denmark. It is an alternative to the ordinary profit-driven supermarket chains. Any profit made, go towards lowering food prices, developing the organization or supporting social initiatives and projects in Copenhagen.

The organization focuses on offering organic and biodynamic seasonal products with good taste and quality at affordable prices. Within KBHFF, the customers are members, owners and co-workers.

The idea for a time bank in KBHFF started as a way to increase the impact of KBHFF in local Copenhagen communities as well as a way to increase the cooperation between members. One of KBHFF's major goals is to create grounds for a more sustainable economy through locally based, organic food production.

With its conscious focus on locally produced food, KBHFF seeks to build an alternative political economy of provision that supports local food production and the resilience of communities through self sufficiency. The actors map in Fig. 1 represents all the actors that are involved in the Time banking service (called HOUR COMMUNITY) and that are ensuring a parallel flow of information, building in this way the trust that is needed to share personal Open Data.

ACTORS' MAP

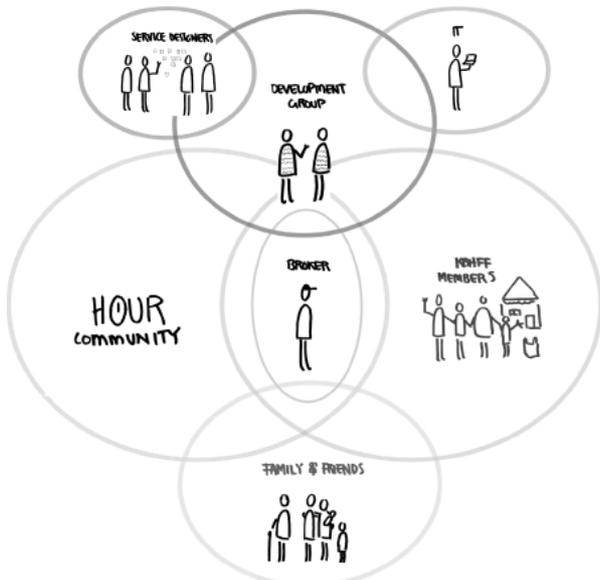


Figure : Actors Map for the Time Bank project

The schema shows links between each actor in the community internal actors (development group, Hour community, Broker, KBHFF members and their family and friends) including external actors (IT team and Service designers). Having family and friends as actors is a relevant element here because of the shared vegetables and the membership. This membership is shared because a member has to work three hours a month for the co-op, and these hours can be performed by their spouse if need be.

The internal service providers may access the private profile of members for particular uses but they also have a separate profile to carry out tasks specific to their role as service provider.

A member of the local KBHFF has an email address and must be able to access and use the Hour Community website.

The broker is the closest contact a member can have to the Hour Community, so she/he must be able to communicate clearly both verbally and in writing and becomes the diplomatic spokesman in the case of service breach. She/he must be able to explain about what Hour Community is, should have “normal” IT capabilities and does not need to understand every single technical detail of the system, but she/he should be able to use the Hour Community website and interface with the database if needed.

The event manager is able to plan and carry out events for members. He can define and divide the work needed to be done at events and use the Hour Account to

request help. He should communicate with members to have suggestions for events and he should also share infos with the KBHFF event group.

The development group should translate member feedback into both technical and non-technical solutions. She/he should set goals for further development of Hour Community and handle the main communication with KBHFF and other external actors.

Based on how active the community is, there might be a need for more internal service provider roles so the development group should be able to decide and assign new roles.

The last actor, the external IT maintainer, should be able to understand and manage all IT systems that have been developed to carry out small corrections and handle possible database maintenance and breakdowns.

Fig. 2 shows the global blueprint of the Service. Narrowing down to the main steps, we get a simple view of how a whole general service functions with the backend office. The actors who are active in this service process are also shown in the blueprint: the broker, the one who signs a member up and supports her/him with help and guidelines, while the event manager manages the event and its requirements.

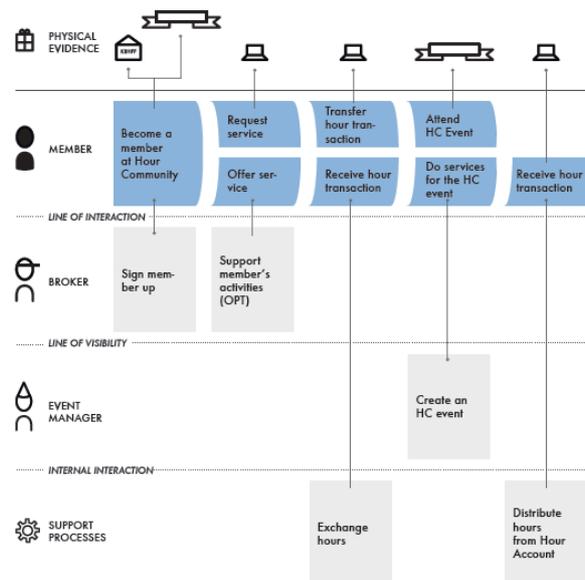


Figure : The global blueprint of the Hour Community service

This example shows a real-life deployment of the issues and dilemmas outlined in this paper. In our society, time is perhaps our most precious asset. Time crosses over all our agendas, public and private alike. However, a time bank is a service to which anyone is most likely lending time from its own private agenda, since the public one (work, family, etc.) is probably entrenched in a strongly fixed system which is generally hard to subvert. This private agenda may be divided in two parts: on the one hand, all those micro-moments of

idleness which accompany our daily lives (waiting for the bus, queuing at the supermarket, waiting for a document to print, a computer to finish, the network to respond, etc.); on the other hand, substantial moments which are generally devoted to relax and recover from the fatiguing routine (evenings, vacation periods, sickness periods, etc.).

The first ones are hardly usable with current technology, while the second ones are really too close for comfort with privacy – after all anything that is not a “public” moment is obviously by definition “private”.

This case study tried hard to handle this delicate issue with the utmost care and attention, but it was hard not to notice and analyze the language barrier issue taking place between service designers and computer engineers.

Trust and accountability issues were partially helped out by the strong reputation of seriousness and integrity that the KBHFF community has gained in the past years. However, providing a new service like the one described here (Time bank) has implied building a new and different kind of reputation. This latter kind started building when a critical mass of followers (both providers and users) was reached, because the amount of data correlation available to the final user were large enough to allow trust building.

PROBLEM PRIORITIZING SERVICE

This case study explores maintenance in S--train stations, and investigates whether commuters could be dynamically involved in this.

This service system has been designed in response to the need of having better maintenance of S-train stations in the Copenhagen area.

The main objective of this project is to obtain a better alignment of stakeholders priorities without violating the privacy of commuters but also of the company that is in charge of the S-train station maintenance.

This service creates information on commuter perception and prioritizes it. It is an independent service to which large organizations in need of such information can subscribe. In the case of the S-train network in Copenhagen, the service could be subscribed to by DSB, BD, and the individual municipalities.

A panel of commuters will be recruited and they will be helped with basic instruction and some motivation - e.g. subsidised travel.

The commuter panel would be given access to a back-end application. It would then incorporate the service into its individual commutes. It would be able to receive information from this service about suggested issues to photograph.

The service would base these recommendations on requests from participating organizations. The commuter panel would take pictures of maintenance issues using the back--end app, tagging them with essential information (location and time within the station and nature of the problem), and uploading them to the collected image pool. The images would be geotagged, linking them to the station they belong to.

Members of the commuter panel could contact the service through the back--end app. When the images are uploaded to the image pool, they would become accessible to the general public via a front-end app.

This app would be available for free on mobile platforms, so any commuter, citizen, or work organization could gain access to it. Users would interact with photographs nearest to their geographical location.

A promotional campaign within the S--train stations and vehicles would alert commuters of the presence of this service.

Fig. 3 shows the motivation matrix with all the stakeholders involved (where LENK is the name of the service). It explains quite well expectations and roles.

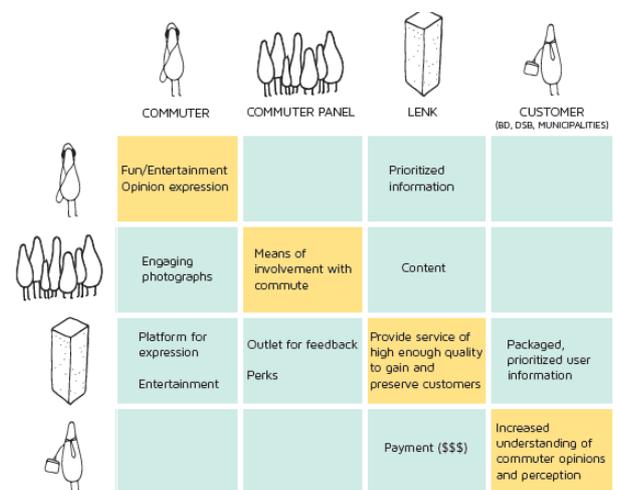


Figure : In order to consider the motivations of the different stakeholders, and to iterate what they expect of the service, the motivational matrix method was employed (Manzini 2004).

This service would gradually gather a large pool of “clicks” as users interact with the pictures. These would then be sorted resulting in useful elements of information that could be passed back to the participating organizations.

This case study outlines the actual complexity that crops up while attempting to build useful participatory services while preserving the rights to privacy of each

stakeholder (in this case the commuters, the main service provider, the provider of accessory station maintenance services, etc.). The open participatory design ensures that trust can be built over the service itself through the abundance of sources and open data, while the anonymized involvement of all actors contributes to the needed privacy stances.

CONCLUSIONS

This paper outlines the issues that are inherent to participatory design using Open Data. These issues mainly involve privacy and trusting.

Design (particularly participatory design) must be aware of these issues and must conceive services which are able to build trust among stakeholders while preserving individual privacies. The case studies show the added complexity layer that attempted solutions to these issues involve, outlining however the idea that the use of Open Data in participatory design clearly offers proper solutions to most of them.

ACKNOWLEDGMENTS

The author would like to thank and mention the students that have been working to the development of the two case studies taken into consideration in this paper:

Fanny Giordano, Helen Phuong Nguyen, Veronika Ondrikova, Jonas Rauff Mortensen, Lara Claire Casciola, Esben Grondal, Nea Alina Annika Kosonen, Katharina Maria Stark.

REFERENCES

Bauman, Z. 2000. *Liquid Modernity*. Wiley.

Buckingham Shum, S., Aberer, K., Schmidt, A., Bishop, S., Lukowicz, P., Anderson, S., Dunwell, I., Edmonds, B., Grey, F., Haklay, M., Jelasity, M., Karpitenko, A., Kohlhammer, J., Lewis, J., Pitt, J., Summer, R., and Helbing, D., 2012. Towards a global participatory platform. *The European Physical Journal Special Topics*, 214(1):109-152.

Castelfranchi, C. and Falcone, R. 2010. *Trust Theory: A Socio-Cognitive and Computational Model*. Wiley Publishing, 1st edition.

Castells, M. 2011. *The Rise of the Network Society: The Information Age: Economy, Society, and Culture, volume 1 of Information Age Series*. Wiley.

Chamberlain, A. and Griffiths, C. 2013. Moths at midnight: Design implications for supporting ecology-focused citizen science. In *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia 2013*, Luleå, Sweden.

Falcone, R., Barber, K., Sabater-Mir, J., and Singh, M. 2008. *Trust in Agent Societies: 11th International Workshop, TRUST 2008*, Estoril, Portugal. Revised Selected and Invited Papers. *Lecture Notes in Computer Science / Lecture Notes in Artificial Intelligence*. Springer.

Hemment, D., Smith, K., Tait, J., and Dornan, A. 2013. Datagm: curating an environment for change. In *Proceedings of the Participatory Innovation Conference 2013*, Lahti, Finland.

Klein, N. 2009. *No Logo*. Picador.

Manzini, E., Evans, S., and Collina, L. 2004. *Solution oriented partnership: how to design industrialised sustainable solutions*. Cranfield University.

Maritain, J. 1947. *The person and the common good*. Charles Scribner's sons, New York.

Papangelis, K., Chamberlain, A., Velaga, N., Corsar, D., Sripada, S., Nelson, J., and Beecroft, M. (2014). People, plans and place: Understanding and supporting responses to rural public transport disruption. In *Proceedings of the 11th International Conference on the Design of Cooperative Systems*, Nice, France.

Shannon, C. and Weaver, W. 1949. *The Mathematical Theory of Communication*. University of Illinois Press.

Swift, T. (2001). Trust, reputation and corporate accountability to stakeholders. *Business Ethics: A European Review*, 10(1):16–26.