Developing the data-based organisation: Exploring Data Work and Human-Data Interactions in a Cross-Organisational context

Cathrine Seidelin
IT University of Copenhagen
Rued Laangaards Vej 7, 2300 Copenhagen, Denmark
cfre@itu.dk

Organisations are looking for new ways to develop their services through innovative use of data. However, current tools and techniques for data-based service innovation often demand that people have highly-specialised skills, which allow them to design with data. This creates a barrier for many SMEs to engage with and explore the innovative potential of (Big) data. In this paper, I describe how my doctoral study aims to tackle this challenge. My work aims to create methods that can enable and support non-IT experts to explore and interact with data when developing new service offers. This paper reports on the industrial setting in which this action research project is carried out. It furthermore elaborates on the research objectives and the initial findings from my first case study. Finally, my contributions thus far and my research goals are discussed with consult sought on how to proceed with exploring and developing tools and techniques or data-based service design.

Big data. Service Innovation. Data as Design material. Human-Data Interaction.

1. INTRODUCTION

The research field targeting and exploring "Big Data" keeps evolving and, in part, mirrors the increasing societal focus on data's innovative potential and value (Manovich, 2011). However, current research within this field mainly focus on technical aspects of applying and making use of the vast amounts of data, e.g. storage, handling and analysis of heterogeneous and high-volume data (Hashem, 2015). Research that discusses Big Data from a socio-technical perspective, which considers the societal effects and ethical consequences of analysis and implementation of analytics and data-driven decision-making processes, is less prevalent Zwitter, 2014). One area that has yet to be further explored from a socio-technical perspective is how organisations can explore (Big) data in order for them to utilize their data resources, e.g. to improve existing services or discover new potential business opportunities and service offerings.

Boyd and Crawford argue it is important to consider how the currently available data analytics tools take part in creating our societies as we make use of them and their incorporated data streams (Boyd, 2012). Currently, tools and techniques for handling and analysing heterogeneous data sources are restricted to those with a computational background or computational skills. It requires highly specialised skills to make sense of and innovate with vast amounts of data from heterogeneous sources. This results in that most people are excluded from designing and developing new solutions and services that incorporate Big Data (Manovich, 2011, Uddannelser, 2018).

My doctoral study targets this research gap and aims to develop methods that allow non-IT experts to explore and engage with (big) data during the process of designing (new) services. In this context, I understand and use the term 'service' to describe social, material and technical configurations that together provide value for users and stakeholders. One of the main objectives of this research is to consider how the analysis and application of data analytics can become more accessible in an organisational context in order to include more people in the process of using and designing with (big) data.

The remainder of this paper is structured as follows: First, I describe the primary and secondary research settings. Secondly, I present my main research question and sub-questions. Then, I elaborate on my research approach and methodology, which is followed by initial results from my first case study. Finally, future research is suggested and I propose two areas for which I seek advice for my research.

2. PRIMARY RESEARCH SETTING

This PhD research is founded and will evolve within an organisational setting. Thus, it is relevant to elaborate on IU and the context in which the organisation exists and operates. As mentioned, IU is an education secretariat for AVT and VET in the industrial sector in Denmark. The organisation navigates within a strong network of the leading organisations in the Danish Labour market (Uddannelser, 2018).

By operating within this network, IU also functions as a knowledge broker between various stakeholders such as unions, manufacturers, VET students, VAT participants, and industrial organisations.

IU is a service organisation, which takes on two general roles: The organisation functions both as a service developer that create e.g. new educational programs, and as a service provider in that the organisation also maintain e.g. existing educational programs and many other administrative tasks (Seidelin, 2017). Due to the advances in technology, the context in which services take place and are experienced have become increasingly complex (Ostrom, 2015).

The evolving information technologies creates new opportunities for ways in which service providers, such as IU, can deliver and develop services. IU wants to address these challenges and increasing complexity by exploring new ways of making use of data, in order to be able to continuously develop state-of-the-art educational schemes. Furthermore, IU wants to develop new competences for how the organisation prepares and provides its services. In other words, IU wants to engage with Big Data and its promising potential, while also inducing change that enable the organisation to develop and innovate services in new ways. On this basis, IU is a very relevant organisation for this project.

2.1 Secondary Research Setting

In addition to the study's focus on IU and its related network of external organisations and stakeholders, I am currently involved in the "Data Science for Local Government" at the Internet Institute at Oxford University.

This project investigates how governmental agencies and organisations initiate, work with and implement Data Science (that being for example new ways of using data to automate service processes). This thus inform my doctoral research in terms of data practices in small and medium-sized organisations (see research objectives below).

3. RESEARCH QUESTIONS

The main research question, I plan to answer is:

How can small and medium-sized organisations, such as IU, explore the innovative potential of (Big) Data and design data-driven services?

Sub questions:

- What are the current data practices in small and medium-sized organisations? What are the problems?
- How can we explore representations and design elements to make complex data sources comprehensible?
- How to organise a data-centred Service Design process?
- Can we establish such processes in the organisation to support continuous innovation of data-centred services?

Each of these research questions are empirical questions which will be founded in and investigated with the case organisation.

4. RESEARCH APPROACH AND METHODOLOGY

This PhD research is based on a participatory action research approach. I am employed by IU as an Industrial PhD student, and I therefore work as an embedded researcher. Action Research is a methodology which implies that the research aims to induce change and improve certain aspects within the target research domain (Robson, 2002, McNiff, 1996, Stringer, 1999).

This research draws on the understanding that Action Research on a general level consists of a three-stage cyclical process: 1) Planning – 2) Implementing – 3) Evaluating. Within this framework, Action Research is a recursive iterative process that often does not proceed in a linear manner (Stringer, 1999). Moreover, Action Research aims to create, e.g. organisational change through a "bottom-up" approach, where people within the organisation is involved and take part in inducing the change (McNiff, 1996). The participatory "twist" in this project's Action Research approach lies in the planned, continuous and increasing involvement of the people at IU. Practitioners at IU as well as relevant stakeholders are invited to participate in the methodological design of the research. The inclusion of the organisation and its employees can help counteract the treat of a single perspective also known as researcher bias (Robson, 2002). An inclusive methodological design process might also promote decisions that induce organisational change. In the following paragraphs, I elaborate on the methodological framework and what is meant by 'increasing involvement' of the people at the case organisation.
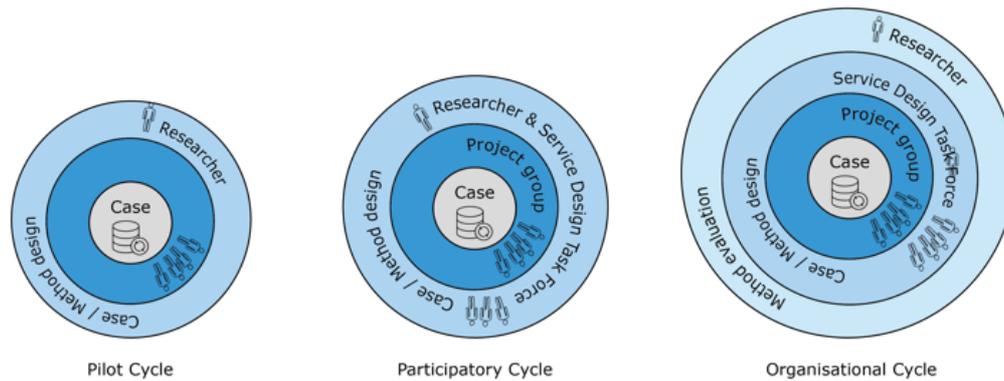
Figure 1 Action Research Cycles and Method development

4.1 Pilot Cycle

The initial Action Research cycle focuses on exploring and developing specific design representations for Big Data in Service Design. Together with the management at IU, I selected a presumably simple case for the first Service Design process. The case revolves around the organisational processes that are carried out in order to maintain and update data in one, single database, which primarily stores information on members of Local Education Committees. The process and insights of the Pilot cycle is further described below in 5. Initial Results.

4.2. Participatory Cycle

The second cycle is planned to build on the learnings and insights from the Pilot cycle. This cycle aims to develop descriptions of more general principles for methods for Service Design for Big data and to develop the data representations from the first cycle. The Participatory cycle was initiated in the beginning of 2018, where management from IU and key stakeholders were invited to workshops, which aimed to identify the case for the second cycle in a participatory manner. As shown in figure 1, this cycle aims to involve and engage relevant employees in the case organisation to a greater extend. During the second cycle I will not act as the main driver of the design process. Rather, the field- and design work will be conducted in a cooperation with people at IU.

4.3 Organisational Cycle

The third and final Action Research cycle in this project will explore hoe people (who are not educated in Service Design or applied mathematics) work and manage the co-developed Service Design for Big Data methods and techniques. Moreover, this cycle will explore whether (and if so, how) this new approach can be anchored within the case organisation. During this cycle the established "Service Design Task Force" will drive the design process, while I will take a more observant, evaluating role (cf. Figure 1).

5. INITIAL RESULTS

As described above, my doctoral study consists of three overall action research cycles, which each focus on a data-based (either existing or new) service. The first cycle investigated the current data practices related to service provisioning of Local Education Committee (LEC) members, and the process of redesigning the current services related to the LEC members. The data about individual LEC members are used across organisational boundaries, by various stakeholders with different knowledge interests and "data priorities". However, the data is currently stored in a single database, and IU is the knowledge-broker of this data (Seidelin, 2018a).

The initial data collection consisted of approximately 20 hours of field observations and 15 semi-structured interviews with administrative "data workers", who represented 12 different organisations. During the data collection, I focused on how data was produced, collected, analysed and applied in order to maintain and update the LEC database. What became visible from the analysis of the fieldwork material was that data quality and data transparency is vital in a complex collaborative network, where actors partly have different interests. These particular dimensions become highly important to consider because data and data management needs to be accountable for all actors that take part in the collaborative processes of maintaining and updating the data (for a more detailed analysis and discussion, see (Seidelin, 2018a).

I applied the notion of "Human-data Interaction" (HDI) as an analytical lens to analyse this case, (Mortier, 2014). The notion of HDI is a perspective that aims to emphasise the increasing ethical and practical challenges that exists when collecting, analysing and trading data (Crabtree, 2015). So far, the field of HDI studies has mainly focused on

personal health related data, and thus the interaction between an individual and his or her personal data (Mortier, 2014). The HDI perspective arguably takes a first step in order to recognise data as an entity in its own right - an entity that needs to be considered and represented in design. However, I argue when taking a cross-organisational perspective the level of complexity increases: It is no longer 'just' the interaction between a person and his or her personal data that is at play. In a cross-organisational setting it is necessary to consider and make visible that multiple stakeholders interact with the 'same' data across organisational boundaries (Seidelin, 2018b).

The empirical work and theoretical insights informed a subsequent string of three participatory design workshops. During the first workshop, I proposed a notation for representing the data entities, which exists in the LEC database. Together with the participants, I explored how non-IT experts where able to engage with the data representation (Seidelin, 2018b). The second and third workshops focused on developing new service concepts for to improve the current data practices and inform the subsequent system design.

6. PROPOSED CONTRIBUTIONS OF CURRENT WORK

This doctoral study is expected to contribute both to the scientific community and to industrial practice. As mentioned in the introduction, there is very little research on how to design and innovate services and related data analytics to make use of existing and new data and realise its potential value. This research will contribute with methods and techniques that enable development of new services based on heterogeneous data coming from various sources. More specifically, the results of this projects (and future projects) will contribute to the HCI community, by considering how to render and represent data in design.

In terms of industrial practice, this research is expected to first of all radically change how IU – and potentially other organisations similar to IU (e.g. knowledge-broker and small and medium-sized service organisations) – prepares and provides services supporting development and use of educations. Secondly, the project is expected to improve IU's competences for better and smarter use of data to meet the needs of multiple stakeholders in a cross-organisational context. Finally, the long term societal and economic potential lies in the ability to continuously develop methods and techniques that develop small- and medium-sized services organisations' capabilities to work with heterogeneous data based on strategic considerations, and their ability to build competencies for sustained service innovation that responds to societal changes

7. ADVICE SOUGHT

For this doctoral consortium, I am seeking advice to one particular question: How can we develop a discussion that advance how we can talk and theorise about "interacting with data"? As mentioned above, the notion of HDI takes an important first step by emphasising that it is crucial to consider data in its own right in order to develop systems and data infrastructures that take privacy and ethical concerns into account (REF). During the initial stage of my doctoral study I have proposed that the HDI perspective has so far mainly been applied in cases concerning an individual user and his or her personal data. I therefore question whether this is applicable in a context where potentially multiple users interact with (not necessarily personal) data. Moreover, how – if relevant – do we consider the HDI perspective in regards to HCI? I would like to discuss and seek advice on how to address this issue.

8. REFERENCES

CRABTREE, A. M., R. 2015. Human data interaction: historical lessons from social studies and CSCW. *The 14th European Conference on Computer Supported Cooperative Work.* Oslo: Springer.

HASHEM, I. A., YAQOOB, I., ANUAR, N. B., MOKHTAR, S., GANI, A., & KHAN , S. U. 2015. The rise of "big data" on cloud computing: Review and open research issues. *Information Systems,* 47**,** 98-115.

MANOVICH, L. 2011. Trending: The Promises and the Challenges of Big Social Data. Available: http://manovich.net/content/04-projects/067-trending-the-promises-and-the-challenges-of-big-social-data/64-article-2011.pdf.

MCNIFF, J., LOMAX, P., & WHITEHEAD, J. 1996. *You and Your Action Research Project,* London, Routledge.

MORTIER, R., HADDADI, H., HENDERSON, T., MCAULEY, D. & CROWCROFT, J. . 2014. Human-Data Interaction: The Human Face of the Data-Driven Society. [Accessed May 2nd 2018].

OSTROM, A. L., PARASURAMAN, A., BOWEN, D. E., PATRÍCIO, L., & VOSS, C. A. 2015. Service Research Priorities in a Rapidly Chaning Context. *Journal of Service Research,* 18**,** 127-159.

ROBSON, C. 2002. *Real World Research - A resource for Social Scientists and Practitioner-Researchers,* Oxford, Blackwell Publishing.

SEIDELIN, C., DITTRICH, Y., GRÖNVALL, E. Identification of Data Representation Needs in Service Design. Selected Papers for the IRIS, 2017 Halden, Norway. AISeL.

SEIDELIN, C., DITTRICH, Y., GRÖNVALL, E. Data Work in a Knowledge-Broker Organisation: How Cross-Organisational Data Maintenance shapes Human-Data Interactions. The 32nd British Human-Computer Interaction Conference, 2018a Belfast, Northern Ireland. The ACM digital library and the British Computer Society.

SEIDELIN, C., DITTRICH, Y., GRÖNVALL, E. 2018b. Designing With Data: An Exploration of how Data can Become a First-class Citizen in Participatory Design. *Submitted to NordiCHI.* Oslo

STRINGER, E. T. 1999. *Action Research* Thousand Oaks, CA, USA, SAGE.

UDDANNELSER, I. 2018. *Industriens Uddannelser's website: Erhvervsuddannelser (Vocational Training)* [Online]. Available: https://iu.dk/uddannelser/erhvervsuddannelser/ [Accessed May 14th 2018].