DOCTORAL THESIS IN
INFORMATION AND COMMUNICATION TECHNOLOGY
STOCKHOLM, SWEDEN 2015

Health Data

Representation and (In)visibility

PEDRO SANCHES

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Health Data: Representation and (In)visibility

Pedro Sanches

(*sanches@kth.se*)

TRITA-ICT/ECS AVH 14:17
ISSN 1653-6363
ISRN KTH/ICT/ECS/AVH-14/17-SE
ISBN: 978-91-7595-403-5

SICS Dissertation Series 72
ISSN 1101-1335

Akademisk avhandling som med tillstånd av Kungl. Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen i informations- och kommunikationsteknik den 29:e januari 2015 klockan 14.00 i Sal B, Electrum, Kistagången 16, Kista.

Abstract

Health data requires context to be understood. I show how, by examining two areas: self-surveillance, with a focus on representation of bodily data, and mass-surveillance, with a focus on representing populations. I critically explore how Information and Communication Technology (ICT) can be made to represent individuals and populations, and identify implications of such representations. My contributions are: (i) the design of a self-tracking stress management system, (ii) the design of a mass-surveillance system based on mobile phone data, (iii) an empirical study exploring how users of a fitness tracker make sense of their generated data, (iv) an analysis of the discourse of designers of a syndrome surveillance system, (v) a critical analysis of the design process of a mass-surveillance system, and (vi) an analysis of the historicity of the concepts and decisions taken during the design of a stress management system. I show that producing health data, and subsequently the technological characteristics of algorithms that produce them depend on factors present in the ICT design process. These factors determine how data is made to represent individuals and populations in ways that may selectively make invisible parts of the population, determinants of health, or individual conception of self and wellbeing. In addition, I show that the work of producing data does not stop with the work of the engineers who produce ICT-based systems: maintenance is constantly required.

Sammanfattning

För att förstå hälsodata krävs sammanhang. Jag visar hur detta kan erhållas, genom två fallstudier: en om självövervakning, med fokus på representation av kroppsdata, samt en om massövervakning, med fokus på representation av populationer. Jag granskar kritiskt hur informationsteknologi (IT) kan fås att representera såväl individer som populationer och vilka följder det får. Mina bidrag är: (i) utformningen av ett självövervakningssystem för stresshantering, (ii) utformningen av ett massövervakningssystem baserat på data från mobiltelefonanvändning, (iii) en empirisk studie av hur användare av en hälsosensor begriper det data som sensorn genererar, (iv) en diskursiv analys av hur syndromövervakningssystem utformas, (v) en kritisk analys av processer kring att utforma ett massövervakningssystem, samt (vi) en analys av den historiska korrektheten i begrepp och beslutsfattande i samband med utformningen av ett stresshanteringssystem. Jag visar att produktion av hälsodata, liksom tekniska beskrivningar av de algoritmer som används i den processen, beror av faktorer som hänger samman med IT-utformningsprocessen. Dessa faktorer avgör sedan hur data kan fås att representera individer och populationer på sätt som kan rendera delar av en population, hälsodeterminanter, eller individens självuppfattning och förståelse av välmående osynliga. Jag visar också att arbetet med att producera data inte är avslutat i och med det ingenjörsarbete som krävs för att IT-systemen ska byggas: konstant underhåll krävs också.

Acknowledgments

This thesis is the product of a great journey. Many have contributed to this work, directly and indirectly, and I am eternally thankful.

I would like to thank my supervisors. Dr. Markus Bylund for introducing me to the mad world of research, and for helping me take the first steps, Dr. Nina Wormbs for her keen eye, precise criticism, and mind-blowing perspective shifts, Dr. Louise Barkhuus for all the encouragement and help, and Prof. Magnus Boman for all the lunches, espressos, stimulating conversations, and unrelenting trust and motivation. All of you provided me with a broad view of what research is, and stimulated my intellectual curiosity in all kinds of exciting directions. What a team.

I am grateful to Prof. Harald Kjellin, Prof. Christian Schulte, Prof. Gerald Q. Maguire Jr., and Viktoria Lindström for reading early drafts of the thesis and providing constructive, decisive comments.

I would also like to thank my co-authors and those who contributed to the work included here. First, a very special thanks to Baki Cakici, whom I have learned immensely from. Our conversations, readings, and exchanges of ideas are the basis for this thesis. Also, you taught me a very valuable lesson: always keep up the momentum. I would also like to thank Kristina Höök, because without her I would never have thought about doing research in the first place. I would also like to thank the Affective Health team, especially Elsa Vaara, Ana Ståhl, Claus Weymann, Nathalie Peira, and Marie Sjölinder. And I would like to thank the Consider8 team: Olof Görnerup, Richard Carlsson, and Simon Moritz. In particular, I want to thank Eric Oluf-Svee for always being an excellent office roommate, and more importantly a good friend. To the other co-authors, I am truly thankful for all the work you have put into it, all the suggestions, and all the improvements. I am grateful to the people who participated in studies, and read early drafts of each paper. I also thank Thomas Sjöland and Sandra Gustavsson for important assistance in my SCS department life at the School of ICT, KTH.

This thesis has many homes. I would like to thank everyone at the Userware Lab (R.I.P.). What a truly wonderful group of people. I thank you for the reading sessions, for the dinners, and for the after works. The Software and Systems Engineering Lab picked me up when I was lost. It was really nice to meet you all, and your sympathy really kept me going when I was at the lowest motivation point. The Mobile Life has been the first and last home of this thesis. Thank you for putting up with my zombified state, with me taking up two desks with piles of books and printed paper everywhere, for the enforced fikas, and for all the good energy that you guys have. It's truly a magical and creative place, the best place one can be when writing a thesis. The lovely people of the History Division exposed me many times to different intellectual points of view, all while being nice and friendly. I would also like to thank the administration at SICS who has been a great support, ever since the very first days I came to Sweden.

I would specially like to thank those who lent me books to read, and those who have lent an ear to my ramblings and disconnected ideas, especially during the last weeks of writing. You have no idea how much you made a difference.

There are also the people who made this period extra special for me. Those are the climbers, those who know how to party, those who go to Lions, and those who prefer to go to fancier places, the Thailand 2014 crew, the musicians, the legendary Duvbo people, and the people whom I shared Lucidor with. And also the friends from Portugal, even those who like me have chosen to travel for a while, who make my visits feel like I am truly coming home. You guys are family.

I would like to thank Ulrika for doing this incredible journey together with me, through ups and downs, and for simply being an amazing person. I would also like to thank the Morris clan for welcoming me warmly and never failing with love and support.

I would finally like to thank my mum and dad for making me and forcing me to go to school. Well done! Vocês são espetaculares. Obrigado pelo vosso apoio. And I am grateful to my sister for, even far away, always cheering me up with nonsense photos. Um grande abraço para a minha avozinha que sempre me apoiou, e mesmo longe, senti-a tão perto! E para o avô, esta tese é para ele.

Papers included in this thesis

This thesis is primarily based on six papers:

I. Sanches, Pedro, Kristina Höök, Elsa Vaara, Claus Weymann, Markus Bylund, Pedro Ferreira, Nathalie Peira, and Marie Sjölinder. 2010. “Mind the Body!: Designing a Mobile Stress Management Application Encouraging Personal Reflection.” In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, 47–56.

II. Sanches, Pedro, Eric-Oluf Svee, Markus Bylund, Benjamin Hirsch, and Magnus Boman. 2013. “Knowing Your Population: Privacy-Sensitive Mining of Massive Data.” *Network & Communication Technologies* 2 (1): 34

III. Boman, Magnus, and Pedro Sanches. 2014. “Sensemaking in Intelligent Health Data Analytics”, *KI – Künstliche Intelligenz*, in press.

IV. Cakici, Baki, and Pedro Sanches. 2014. “Detecting the Visible: The Discursive Construction of Health Threats in a Syndromic Surveillance System Design.” *Societies* 4 (3). Multidisciplinary Digital Publishing Institute: 399–413.

V. Sanches, Pedro. 2014. “Positions in technology design: A case study of a stress management system”, *unpublished manuscript.*

VI. Sanches, Pedro. 2014. “Data-driven knowledge production in software engineering”, *submitted.*

The papers are referred to in the text by the above Roman numerals.

Papers not included in this thesis

This list includes papers that did not directly contribute to the results, but are nevertheless relevant to the work included in this thesis. This list does not include extended abstracts accepted at conferences.

- Ferreira, Pedro, Pedro Sanches, and Alexandra Weilenmann. 2013. “Awareness, Transience and Temporality: Design Opportunities from Rah Island.” In *Human-Computer Interaction--INTERACT 2013*, 696–713. Springer.

- Buscher, Monika, Markus Bylund, Pedro Sanches, Leonardo Ramirez, and Lisa Wood. 2013. “A New Manhattan Project? Interoperability and Ethics in Emergency Response Systems of Systems.” In *10th International Conference on Information Systems for Crisis Response and Management*. Baden, Germany.

- Sanches, Pedro, and Markus Bylund. 2011. “Experiencing Mobility Data.” In *1st International Workshop on Values in Design--Building Bridges between RE, HCI and Ethics*, 56.

- Silvăşan, Iuliana, Per Kreuger, Pedro Sanches, Elsa Vaara, and Marie Sjölinder. 2010. “Movement Identification in Affective Health--a Mobile Biofeedback Monitoring System.” In *18th Telecommunications Forum TELFOR*, 340–43. Belgrade, Serbia.

- Sanches, Pedro; Vaara, Elsa ,Sjölinder, Marie, Weymann, Claus and Höök, Kristina (2010). Affective Health – designing for empowerment rather than stress diagnosis. Proceedings of the workshop, *Know thyself: monitoring and reflecting on facets of one's life* at CHI 2010, Atlanta, GA, USA.

- Bylund, Markus, Pedro Sanches, and Eric Svee. 2010. “Collecting and Associating Data” European Union. EP2242292 A1

- Höök, Kristina, Petra Sundström, Jakob Tholander, Pedro Pereira Alves Ferreira, Anna Ståhl, Jarmo Laaksolahti, Elsa Kosmack Vaara, et al. 2010. “Design Processes for Bodily Interaction.” In *At the Workshop Artifacts in Design: Representation, Ideation, and Process, CHI, Atlanda, USA, April 2010.*

- Svee, Eric-Oluf, Pedro Sanches, and Markus Bylund. 2009. “Time Geography Rediscovered: A Common Language for Location-Oriented Services.” In *Proceedings of the 2nd International Workshop on Location and the Web*, 11. LOCWEB ’09. New York, NY, USA: ACM.

- Ferreira, Pedro, Pedro Sanches, Kristina Höök, and Tove Jaensson. 2008. “License to Chill!: How to Empower Users to Cope with Stress.” In *Nordic Forum for Human-Computer Interaction Research (NordiCHI)*, 123–32. Lund, Sweden: ACM. doi:10.1145/1463160.1463174.

- Bylund, Markus, Jussi Karlgren, Fredrik Olsson, Pedro Sanches, and Carl-Henrik Arvidsson. 2008. “Mirroring Your Web Presence.” In *Proceedings of the 2008 ACM Workshop on Search in Social Media*, 87–90.

Contents

1 Introduction

An increasing amount of data on human behavior is collected and processed by information and communication technology (ICT) systems. The relative ease of data collection coupled with efficient storage and the construction of increasingly complex algorithms produces a myriad of globe-spanning data flows (Castells 2011). This data is created from a variety of sources: it may be a by-product of an existing ICT infrastructure, and repurposed—as is the case with many so-called Big Data applications (Manyika et al. 2011)—or purposefully generated to serve a given information need. It is said (Data et al. 2013) that exploitation of such data can have great potential, like to inform and guide policy, to be commoditized and sold for commercial purposes, and it is said to hold potential for sociological research. Such exploitation, however, has its drawbacks. The most acknowledged one is the issue of privacy. Privacy is also connected with data ownership and intellectual property, security, and liability (Manyika et al. 2011). Producing knowledge with attention to ethics entails, in addition, systematic attention to be given to data collection and analysis, accounting for complex methodological issues of analyzing massive datasets, and giving attention to context when interpreting data. These concerns are voiced by a number of researchers (boyd[1] & Crawford, 2012; Floridi, 2012), and constitute arguments against the notion that massive data can speak for itself. Another potentially problematic issue with the increasing exploitation of data is that data is not fairly distributed: there is a divide between those who control or can process the data, and those who are implicated or affected by it.

In the domain of health, ICT is made to produce data about individuals for the purposes of managing, visualizing, or improving the health of an individual, or of a population. The work included in this thesis concerns production of health data as well as the **c**onsequences of design decisions made by software engineers, such as myself, who contribute to the production of ICT-based surveillance. In particular, I show that the production of health data, and the technological characteristics of databases and algorithms that produce it depend on factors present at the site of ICT production. These come from the intersection of two paradigms that define which entities should be represented, and how data should be produced. These paradigms have developed from (i) technological and (ii) health and wellbeing issues. I describe some of the social codes, (historically situated) concepts, and networks of humans and objects involved in the production of health data, based on two case studies. I look at the work of producing and processing health data about individuals and groups of individuals, both in the context of the actual development of two ICT

[1] The researcher danah boyd prefers her name not to be capitalized.

systems that produce information about humans, and in the context of use where that information is interpreted. Although I have a stronger focus on how data is created, i.e., the design decisions, both contexts are important to get a wide angle picture of how the decisions of ICT designers influence the lives of individuals. Finally, I show that producing data implies to take a stance on what is represented (and how), and what is made invisible.

1.1 ICT-based health surveillance

In the domain of health, science and technology have produced a multitude of categorizations and methods that describe and organize several aspects of individuals and populations (Webster 2002; French 2009; Lupton 2012). I will not provide a general survey but focus solely on the aspects of ICT-based surveillance that intersect with health. In this intersection, engineers and software designers conceptualize, design, and deploy systems that render bodies and populations visible and available to policy and management. Recent developments in ICT for health have introduced ICT in aspects of prevention and care, often taking the form of surveillance systems (Boman 2014).

Surveillance is a practice that is often colloquially associated with negative, dystopian meanings. Arguably, these meanings may distract from the everyday practice of individuals watching over individuals, which can be also called surveillance. This does not necessarily signify an evil practice *per se*. The rationalizing processes of monitoring and record-keeping that characterize surveillance can be considered to constitute the processes and structures that make and maintain modern society (Giddens 1987). In fact, one can conceptualize this practice as ranging from the ordinary act of making lists of individuals, to CCTV cameras, personal identification, and credit card numbers (Norris, McCahill, and Wood 2002; Lyon 2007). In addition, ICT-based surveillance systems have been presented as a solution to pressures on the public health system, and promises changes in how medical care is to be delivered (Hood and Auffray 2013). According to proponents of increasing use of ICT in health (Felt, Gugglberger, and Mager 2009; Webster 2002; Savel, et al. 2012), modern healthcare systems suffer pressures from demographic changes in populations (Commission 2013; Boman 2014). Economic changes in how healthcare costs are to be distributed, such as the emergence of several public/private configurations in state welfare, insurance companies, private and public hospitals and clinics, etc. to an increasing extent define what data should be collected, how it should be processed, who should be under surveillance, and who should be the ones looking. ICT-based solutions may enable efficient management of processes, which may lead to lower costs of maintaining healthcare, although this has not been conclusively shown (Chaudhry et al. 2006; Rudin, Motala, Goldzweig & Shekelle 2014). The progressive introduction of surveillance in health arguably contributes to a shift in medical intervention from “treatment and cure” to “management and care” (Webster 2002). As ICT tends to become more relevant to decision making in healthcare, the role of a discipline like software engineering becomes more prominent. I argue that it therefore matters to describe exactly what ICT systems and their means of representation are transforming.

1.2 Aim and research questions

My research aim is to make technology design a socially responsible enterprise. This allows for ICT for health to benefit the widest population possible. One can then rightly say that technology is a factor of change in society. On this note, it is necessary to acknowledge the role of technology in defining social order if one is to talk about social issues by tackling ICT design. This is different from a technological deterministic stance, where social order is defined only by technology, conceived to be on a path to increased efficiency (Smith and Marx 1994). Contrary to this idea, historical accounts of technological development have shown how heterogeneous the task of an engineer can be, having to simultaneously enact social and economical changes in order to develop successful technological innovations (Mayntz and Hughes 1988; Law 1987). Algorithms can have effects on social order if they are produced in order to include certain phenomena, and people, and exclude or de-prioritize others (Friedman and Nissenbaum 1996). This justifies looking at how existing infrastructures and systems are connected to each other, and how data is produced and processed by algorithms. I pursue my aim through two research questions:

1. How is health data produced by ICT-based surveillance?
2. Which entities are made (in)visible in ICT-based health representations?

The formulation of the first research question assumes that health data is dependent on specific configurations and design decisions taken during the process of ICT production. The question points to how these may influence the design of ICT and the data produced by ICT systems. The second research question stems from Ball's (2009) problematization of how to jointly analyze (i) the need to monitor, as created by institutions and the practices of surveillance with (ii) positions created for the individuals implicated by surveillance. The data structures and flows of information constructed by ICT designers embed ideas on how to represent individuals, or how data analysts are to interact with these representations. These configure specific visibilities and ways of managing them; for example, they define the boundaries of what is private and what is public, configure scripted ways in which individuals are to comply or not with the system, or how data is to be interpreted and put to use. Looking at the possibilities enabled or constrained by certain design decisions was, to me, an instrumentalization: a way of grounding the more abstract first research question in the second. That said, the answers to the two questions were pursued in parallel, and they were intertwined in my own work on design of ICT-based surveillance.

1.3 Two types of health surveillance

Self-surveillance produces data from individuals in order to be seen chiefly by the same individuals. The unit of analysis is micro. *Mass-surveillance* systems produce data from individuals, or groups of individuals, in order to produce knowledge about a population. The unit of analysis at this level is a population, the macro level. In addition, the meso level (families, work places, dwellings) is often in focus, albeit indirectly (Boman and Holm 2004). The goal is to model features of a population that can be used to inform public health authorities. Common to both types of surveillance is care,

risk management, and prevention. Models often embody a benevolent intent of caring for the wellbeing of individuals, and achieving that by managing risk.

1.3.1 Self-surveillance

Self-surveillance relies on personal tracking devices. These are typically used by one individual in order to produce data about the human body. Through these devices, behavior or particular characteristics of the human body are measured and made intelligible, in some form, usually to the same individuals that the data is collected from. I focus on two purposes of self-surveillance: fitness and stress management.

There is presently a wide proliferation of wellness and fitness apps. It is estimated that 485 million wearable computing devices will be shipped by 2018 (ABI Research, 2013). Activity tracking is a sensor-driven method for monitoring physical activities of individuals. It pertains to a group of technologies that use wearable sensors to collect information about how individuals move. They convert the traces of the activities performed by the individual to information about those activities into the form of calories spent, steps taken, distance covered, etc. (Kranz et al. 2013). With this conversion, these technologies promise to help individuals remember the amount of movement they perform, often with the intent to motivate them to move more (Barkhuus 2006). An activity tracker has two components: a wearable sensor for movement and a user interface that displays the metrics related to that movement. The sensors are typically accelerometers that measure acceleration and orientation of the body. The user interface can display the statistics generated from the monitored activities, and allows individuals to manipulate the collected data, for example to share with others, to set goals for themselves, or to follow exercise programs.

The growing popularity and availability of these technological devices is not only pushed forward by increasing miniaturization, but also by communities such as the ‘quantified self’ (or life-logging) community (Quantified Self Labs 2013). In this loosely connected community, individuals build on technological and scientific knowledge for measuring, analyzing, and controlling different parts of their own life. The focus of this dissertation is on devices specially customized for measuring specific aspects of the human body. There is vast amount of health and wellbeing applications that rely on devices for gathering data (Swan 2009). Some specially purposed commercial products measure aspects of behavior, such as physical activity through accelerometers (Fitbit 2013). Other products, like W/Me (Phyode 2013), rely on optical heart rate sensors to represent emotional states. There is also a large body of research that focuses on continuous, biosensor-based, stress detection for work, home, and general everyday settings (Fessl et al. 2011; Bakker, Pechenizkiy, and Sidorova 2011; Vidyarthi and Riecke 2013). Most of the systems are either not mobile, which limits recording time, or record data on a minute basis, which makes it difficult to identify stress patterns in data (Sharma and Gedeon 2012). Other studies have used various sensors attached to different places in the body, claiming to be able to determine emotional stress in short term measurements (de Santos Sierra et al. 2011). A number of them aim to represent data related to emotional states, such as a heart rate, together with mechanisms and interaction techniques, such as gentle reminders, or backgrounds, that aim to promote long-term reflection on

stress. Diedrich and Song (2014) compile a wealth of systems that attempt at automatically diagnosing mental health disorders, including stress, using bio-data, behavioral data, speech analysis, and image analysis. For stress-related data, I focused on how cardiovascular activity or increased electrical conductivity of the skin can be related to stress, and the ways that computers can represent stress and emotional states.

1.3.2 Mass-surveillance

Lane and colleagues (2010) argue that using data collected from millions of mobile phones provides a number of societal benefits, and that the risks of location-based attacks are generally understood due to many years of research. Some research into privacy intrusions is driven by corporations that have access to data, such as mobile operators or Internet providers (Tech in Asia 2011). Other reports are academic, with researcher access to released data in restricted forms or researchers creating their own datasets (Nanni et al. 2010; Becker et al. 2013). They make use of a variety of mobility data from multiple sources in order to produce a vast range of location-based services. Some are primarily data-driven, and aimed at predicting correlations between demographics, semantic annotations, and location data. Others consider specific use-cases such as public safety, emergency relief, or commercial services. Many aim at producing high-level concepts related to population behaviors, such as the purpose of trips, home-work commuting patterns, or systematic vs. occasional movement, from large datasets of individual mobility. The focus is not on the individual, but rather on how the aggregated population-wide indicators and metrics are produced, and what the implications are for the individuals implicated in those aggregated metrics. I have analyzed privacy of individuals implicated by surveillance, and how the construction of higher-level representations came about.

Syndromic surveillance is a type of surveillance that uses already existing health-data that pertains to diagnoses — ideally, preceding it— with the goal of predicting outbreaks or situations that will warrant a public health response (Eysenbach 2006). The goal is to accelerate the response by speeding up the diagnosis, in the so-called race to trace. One particular type of indicator data is the collection of records of ambulance dispatches, which contain information about geographical location and the medical reason for the dispatch of an ambulance. Such data directly precedes diagnosis of a disease and, when aggregated at a regional level, it can be tentatively used to predict population-wide health threats. Mobility patterns can also be seen as a valuable addition to syndromic surveillance systems (Anema et al. 2014). Their usefulness is connected to indicators of how disease spreads through a population. Particularly, mobility data derived in quasi real-time could reduce latency in spread indicators and achieve functional real-time disease surveillance, therefore improving public health response times in case of an outbreak.

1.4 Types of contributions included in the thesis

The contributions of this thesis are described in:

- Chapter 4 – Contributions, where I detail the research contribution of each individual paper, and
- Chapter 5 – Discussion, where I answer the research questions, and reflect on the overall contribution of this thesis as a whole.

Here, I provide a concise overview of the types of contributions that this thesis provides. This thesis rests on experience as a practitioner in several academic/industrial projects at the Swedish Institute of Computer Science (SICS) and Mobile Life VINN Excellence Centre (Mobile Life). From 2009 until late 2012, I worked in the roles of a privacy researcher with the task of evaluating ICT systems, and of a software engineer, designing and implementing technological solutions. I produced code and a number of project deliverables, consisting of non-functional requirements[2] for the technology we developed, analysis of legal regulations, and privacy guidelines. Another side of the work was academic, pursuing as a graduate student at KTH the goal of describing the consequences of producing data from human behavior, and the goal of describing the socio-technical milieu in which design decisions were made. This thesis presents the results of the academic side, as shown in publications. It does not include all the academic work completed. Instead, a relevant set of publications was selected, forming a coherent whole. The academic work here presented spans several communities of practice, such as engineering, interaction design, and technology studies. Cooperation with colleagues in those communities was essential in order to achieve methodological rigor. The work included in this thesis focuses on two separate but related case studies:

- The production of a system that aims at visualizing bio-data, addressing how to process and visualize specific types of data pertaining to physiological processes related to stress management.
- The production of a system that aims at creating mobility patterns of a population, with a focus on methods that ensure that personal data is not disclosed, while providing value for customers that wish to make use of aggregated datasets for population health and wellbeing.

I also worked in a third project, which dealt with smart homes, with a focus on privacy, trust, and safety (Cantarell and Sanches 2013). However, I chose to include only the two case studies described above because I was able to discern more complex and critical issues related to surveillance. Additionally, the two case studies have several characteristics in common, which make

[2] Functional requirements are the ones concerning the essential components necessary for its usage and specify what a system must do. An example of a functional requirement is “when an individual presses the power button, the computer should turn on”. Non-functional requirements concern how the system should function. “Visually impaired individuals should be able to localize the power button” is one such requirement.

them suitable for a coupled examination. They were both situated within health surveillance, although not meant to be used within a clinical setting. Both systems were developed within a market logic, in which the goal was to produce innovative products that would satisfy the needs of particular markets. Due to their innovativeness, there were also not a lot of other realized systems that we could build on. This meant that the compromises we made when developing our systems had to be thoroughly justified, with careful examining of each technological component. Thus, they provided good case studies to explore the production and consequences of health data. Together, they provided me with a rich material pertaining to the development of ICT-based surveillance, and the consequences of design decisions in engineering projects allowing me to:

1. Develop a system (case study) that generates representations of metrics from one individual or many individuals (population), describing how they were achieved through a compromise between several technological and social requirements (Papers I and II).
2. Problematize an aspect of these systems, with special attention given to the social and technological aspects related to the production and maintenance of ICT and how these affect the data being generated (Papers III and IV).
3. Return to the system I designed earlier, and reflect on the factors present in its construction (Papers V and VI).

The papers are structured in three types of contributions, summarized in Table 1.

	Technological	**Analysis**	**Reflection**
Self-surveillance	Paper I	Paper III	Paper V
Mass-surveillance	Paper II	Paper IV	Paper VI

Table 1 – Contributions of this thesis, by paper.

In the role of software engineer, I created systems that reflected my own understanding of the world. Through that process, I learned how ICT can be made to represent individuals and populations in certain ways, and gained domain knowledge related to technical difficulties of achieving those representations. That work is summarized in the technological production part of my dissertation, in Papers I and II. As I distanced myself from the initial work, I adopted a critical stance, and started looking at other systems around me, related to the ones I produced. Two systems, Fitbit and SIDARTHa, I considered representative of the two types of surveillance, at the micro and macro levels. These were analyzed, discursively and in the context of use. In Papers III and IV my colleagues and I discuss aspects related to how they represent individuals and populations, and what the implications of those representations are. Finally, towards the end of my studies, I revisited my initial work, with the goal of reflecting scientifically on my initial results and experiences. I asked the question of whether the systems I built could have been done otherwise,

and what the consequences of the decisions we took at the time were. This work is described in Papers V and VI.

The four papers pertaining to the Technological and Analytical part of my work are published, and are included in this dissertation in the format determined by the publishers. These were written in collaboration with other researchers. In Chapter 4, I detail the extent of my contribution to each paper. The two reflecting papers are authored solely by me, and currently in the process of being published.

2 Theory

In this chapter, I outline the theoretical terms and advances that form a basis for my own research, and assist in my answering of the research questions (in Chapter 5).

2.1 Power and knowledge

Definitions of health and wellbeing are sometimes intertwined with techno-scientific endeavors (Rose 2007) with health defined as a fragile state requiring constant surveillance. From the literature, it is obvious that there is no clear line separating care and power (Vaz and Bruno 2002). Those who establish themselves as authorities, often do so under the premise of caring.

Foucault defined power as a productive capacity that can modify the behaviors of others and of ourselves, form knowledge, or produce discourses (Heller 1996; Foucault 1973; Foucault 1995; Foucault 1978). Foucault conceives of history as a conflict between multiple discourses, each trying to constitute itself as a representation of truth. According to Foucault, power and knowledge are indissociable. There is no objective or neutral knowledge, only partial perspectives, and discourses which may be considered, at a particular historical moment, as truths. Truth is a function of power. Power is not something that an individual, or institution, has. It can only be seen in a relation between individuals, when it is put into action. It is also not inherently repressive. Instead, it is more like a productive force field that runs over the whole social body, and it is a part of everything. It can be located in habits, tasks, activities, or conventions. Foucault gives the example of how medicine adopted a scientific language that granted it the power of seeing objectively, creating a "medical gaze". That allowed for the structuring of illnesses in medicine, creating norms of health and deviances, by which all individuals were to be classified (Foucault 1973). That re-structured, for example, the encounter between doctors and patients: which tasks were to be performed and which habits were recommended. Instead of seeing this development as a progression towards a more enlightened truth about bodies and pathologies, Foucault analyses the knowledge/power duality: how and to which extent knowledge constitutes power, and power constitutes knowledge, and what are the implications.

In order to problematize the management of health by ICT-based surveillance in modern society, particularly, I rest on his concept of 'biopower' (Foucault 1998). Biopower refers to the knowledge and disciplinary apparatus that defines and manages aspects of life and wellness. An apparatus is a structure that produces and maintains knowledge. More specifically, an apparatus of biopower defines vital characteristics of humans, including aspects of prevention, illness, and death. Biopower can be used to understand the production of knowledge about health, and is thus a

suitable theoretical framework from which one can depart to understand the wider implications of ICT-based health surveillance, both in its focus on individual bodies (self-surveillance) and populations (mass-surveillance). I do not use biopower directly in any of my papers included in this thesis. However, Foucault's notion of power has been central to most of my work: both of the research questions were derived from it, and I found it important for casting a light on this thesis as a whole. Foucault himself has written extensively about it (Foucault 1998; Foucault 1980), and Rabinow and Rose have recently summarized the concept in three facets (2006). The first facet of biopower is:

> "One or more truth discourses about the 'vital' character of living human beings, and an array of authorities considered competent to speak that truth. These truth discourses may not themselves be 'biological' in the contemporary sense of the discipline, for instance they may hybridize biological and demographic or even sociological styles of thought." (Rabinow and Rose 2006, 197)

A requirement for biopower is the existence of truth-carrying discourses. The apparatus of biopower is composed of disciplinary institutions—such as schools, hospitals, and prisons, which produce useful and healthy individuals—and sciences that construct medical, biological, and sociological discourses. The second facet of biopower it its reactive function, particularly at the population level:

> "Strategies for intervention upon collective existence in the name of life and health, initially addressed to populations that may or may not be territorialized upon the nation, society or pre-given communities, but may also be specified in terms of emergent biosocial collectivities, sometimes specified in terms of categories of race, ethnicity, gender or religion, as in the emerging forms of genetic or biological citizenship." (Rabinow and Rose 2006, 197)

Although history shows that particular institutional configurations have produced contemptible practices based on population management discourses, e.g., forced sterilization of certain groups based on the discourse of eugenics, this apparatus is not necessarily oppressive; rather, it produces modern society, including all of its institutions. These institutions include those that maintain social welfare and those that uphold the law. However, because these institutions can have repressive effects, they should not be left unquestioned. Therefore, it matters to enquire into how truth-carrying discourses are produced and disseminated, and what effects they have. I argue that the contribution of ICT-based surveillance to knowledge about human health constitutes one area that deserves attention, given that a large number of datasets about it are being created. I am also concerned about the effects of applying the knowledge created from these datasets to decision making in crucial areas, such as public health. In particular, one should pay attention to problematic modes of subjectification, which are collected into a third facet of biopower:

> "Modes of subjectification, through which individuals are brought to work on themselves, under certain forms of authority, in relation truth discourses, by means of practices of the self, in the name of their own life or health, that of their family or some other collectivity, or indeed in the name of the life or health of the population as a whole." (Rabinow and Rose 2006, 197)

Subject positions refer to how individuals tend to behave according to historically constructed discourses. Institutions with authority can create what is considered normal, or desirable, and pathologize the deviances from the norm (Yapa 1996). These can refer to categories such as: sinner, criminal, underdeveloped, pervert, etc. Specific discourses on health may, for example, normalize notions of what it means to be fit, and how to behave in order to achieve fitness (Rich and Evans 2005). In this context, ICT-based self-surveillance, with its reliance on sensors worn or carried by individuals and the production of metrics about the human body, can be seen as one of the means to produce health (Vaz and Bruno 2002; Williamson 2014). Individuals allow themselves to be measured and compared to each other, and put under surveillance. This has the potential to affect individuals' experience of embodiment, self, and social relationships (Lupton, 2012). Bowker and Star define *torque* as: "the twisting that occurs when a formal classification system is mismatched with an individual's biographical trajectory, memberships, or location." (Bowker and Star 2000, 223).

For instance, many mobile health applications maintain that obesity can be overcome solely by changing one's behavior (Lupton 2014). This vision may fail to consider the broader social and political dimensions of illness, as well as the medical dimension (e.g., genetics). Unlike techno-scientific endeavors medical gaze that cast the patient as a passive object in constant need of observation, self-surveillance is often operating under the umbrella of games and play (Whitson 2013; Albrechtslund and Dubbeld 2002) Such endeavors are mainly driven by commercial enterprises, with health promotion realized in the same way as when commodities are to be sold. Rather than passively providing information, the technologies employed act on individuals by constructing different types of embodiments and subjectivities in the name of health promotion and wellbeing (Haggerty and Ericson 2000)These technologies cast the individual as responsible for her own health, with the obligation to self-observe and self-correct undesirable behaviors. They do not appeal merely to the desire of an ideal body: discourses about getting fit and healthy appeal also to a sense of self-development (Heyes, 2006), where the practice of knowing oneself is seen as an essential component for progress. Representations of progress are thus made possible by increased self-surveillance and by allowing one's data to be compared with other individuals and past data from oneself. Only then can progress be quantified and made visible via data, relating the present to the past. Quantitative data is presented as more objective and reliable than human memory. These types of statements are often seen in marketing materials related to self-surveillance applications. Similarly, large amounts of data pertaining to a population are also often deemed as more objective and reliable than qualitative survey methods. By looking at the history of quantitative knowledge (Espeland & Stevens, 2009; Porter, 1997), one can see that quantifications are often granted with an objectivity that provides them with exclusive access to the real. This realism, which is achieved by deleting the work and assumptions present in their production, makes quantifications particularly influential. For this reason, they can become problematic if they reify problematic categorizations. Like other types of quantification, ICT-based algorithms producing quantifications may also contain biases and hidden values (Mager, 2011).

2.2 Information and communication technologies

Foucault himself wrote extensively about how knowledge is connected to procedures and techniques of discipline and power, but very little about ICT. However, technological development led him to the question:

> *"*How can the growth of capabilities be disconnected from the intensification of power relations?" (Foucault 1984, 48)

The production of knowledge, and its establishment as truth, is dependent on power relations, and the growing capabilities, including ICT, of institutions who produce it are likely to intensify them (Lyotard 1984). This intensification may be problematic and contribute to the creation of subject positions that lean towards conditions of less freedom. If healthcare becomes health management through constant surveillance, and that surveillance relies on standardized classification schemes carrying truth, there is a risk that individuals are placed in subject positions where they are not able to make their own definition of wellbeing. While ICTs do not, by themselves, produce knowledge, quantification and numerical methods can be seen as playing a major role in how knowledge gets inscribed, mobilized, and made durable:

> "If inventions are made that transform numbers, images, and texts from all over the world into the same binary code inside computers, then indeed, the handling, the combination, the mobility, the conservation and the display of the traces will all be fantastically facilitated." (Latour 1978, 228)

The ability to inscribe and transport information, in an immutable and combinable form, is a powerful way to gain control over the phenomenon under study, and according to (Latour 1986) one of the key processes of knowledge production. Latour has introduced the concept of immutable mobiles to refer to imposable and combinable representations, which remain stable irrespectively of time and place. A digital map is an example of such a representation. Immutable mobiles are capable of exerting power over great distances, since they translate complex local phenomena, into intelligible and stable objects. Structures that maintain an object, such as users, databases, and algorithms must remain stable and constant across places and contexts: immutability can only be accomplished by mobilizing people and resources. Returning to the example of a digital map, one can say that for it to be readable in different places, a number of standards must be set, such having a compatible version of the visualization software.

There are few studies that pay attention to the practice of engineers who produce data Both the techno-positivists that proclaim data as the end of theory, and the critics that proclaim the coming age of Big Brother are flawed. A different discourse is needed, one that acknowledges both the subjective nature of engineering practice and the implications of ICT-based surveillance. In the midst of increasing interest in the use of large datasets for creation of knowledge on how humans behave, I argue that a careful analysis of how knowledge is produced through ICT is needed. The need for an analysis of how ICT is made to represent information about humans comes from the considerable influence that this knowledge may have on decision making in public health.

2.3 The effects of ICT-based surveillance

The definition of surveillance used in this thesis is quite broad: *"*the focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction"(Lyon, 2007, p. 14). It ranges from the ordinary act of making lists of persons to CCTV cameras, personal identification, and credit cards numbers (Dubbeld 2003; Norris, McCahill, and Wood 2002). Instead of seeing it as a negative practice, I see it as a societal phenomenon that has advantages and disadvantages. Giddens (1987) argues that the rationalizing processes of monitoring and record-keeping that characterize surveillance are also features that constitute the processes and structures we know as modern society. These characterizations suggest that paranoid and entirely negative responses to surveillance are unsuitable since the same systems that may be a cause for concern for their power to track personal lives, are established to promote security, justice, and participation in public life. That said, they may also exert a form of control. The reactions and responses to surveillance, as well as what motivates it, vary greatly and its specificities are dependent on culture, gender, ethnicity, class, and income (Lyon 2007). Understanding and explaining surveillance practices and processes therefore requires diverse cultural lenses and methodologies ranging from empirical analysis to interpretative theory.

An aspect of surveillance that has received a lot of attention, both in public (non-scientific) discourse and within computer science is the issue of privacy. Privacy is a multifaceted concept that has both originated and evolved within multiple disciplines. It has remained pervasive in the discourse of resistance against surveillance (Bennett 2011). Sociologists such as Altman have contributed to an understanding of privacy as a boundary regulation mechanism, constitutive of one's personality and agency (Altman 1975). Individualistic definitions of privacy arise from issues of personal territory and protection against governmental intrusion. In many countries, the first legal definition of a private space is the home as a sanctuary. This is the basis for the data-centric approaches that define the limits of what can be done with data, and that has been placed under the umbrella of Fair Information Practices (Cavoukian 2012). These perspectives emerge from the belief that every person has the right to maintain some degree of control of her own data, independent of its use and dissemination. In Europe and the United States at least, practices have materialized in laws, guidelines, and technologies for privacy-enhancing technologies, especially during the last part of the twentieth century. Although many definitions of privacy have an individualistic starting point, recent work on privacy frames it as a social value, "a social construction that we create as we negotiate our relations with others on a daily basis" (Steeves 2008, 193). If an individualistic privacy discourse is taken without considerations of representation, it may be serving only a privileged class of individuals. Coll argues for a view of privacy that considers a more systematic view on how power configures knowledge about individuals, including how it creates individualistic subject positions that must learn to protect themselves. Instead, he argues, we should have a wider conception of privacy as a collective good:

> "Privacy as an individual resource, which every individual should 'learn' to protect thanks to the self-determination principle, cannot compete with political concerns such as the wealth

and security of the state. Only a conception of privacy oriented in terms of a collective good can possibly balance measures meant to serve these overwhelming interests." (Coll 2014)

Some (Palen and Dourish 2003; Barkhuus 2012) call for work that addresses not only aspects of data management but also individual and collective behaviors, choices, and the relation of privacy to other values. Privacy, as a discourse seen through different lenses and multiple definitions, can be made to act upon laws and norms in order to regulate and control information flows in several ways. Privacy then becomes a social and relational value with contextual complexity (Nissenbaum 2009). It is therefore important to realize that the computer engineer making decisions about how to represent data pertaining to individuals has to make implicit or explicit choices (especially when laws are not clear) reflecting an understanding about the social order. In computer science, design methodologies have been developed that include privacy as a central design criterion right at the beginning of the development process. It is argued that it is much more effective than to fit data-centric privacy protection into an already existing system (Cavoukian 2012). Therefore, compliance checking or data protection audit processes that ensure that the system being developed follows a policy guided by privacy principles are often included already in the planning and requirements phase of the project. However, despite its visibility, privacy does not capture all the implications that may come from the act of surveillance (Bennett 2011; Gilliom 2011; Regan 2011; Coll 2014). One aspect that deserves attention is the act of using personal data representing aspects of human behavior to place persons into classes of income, attributes, or offences. This practice, also called "social sorting" (Lyon 2002), is central to understanding surveillance as a tool for influencing, managing, or controlling a population. Through surveillance, law enforcement can look at categories of suspicious groups, credit rating agencies can identify beforehand who are likely to be defaulters, and marketers can influence and attract costumers based on habits and behaviors (Van Brakel and De Hert 2011; Lyon 2007).

French (2009), writing about public health surveillance, shows that the foundations of some of the large-scale systems being built in the area of health are based on an understanding of data as holding a promise of control over the complexity of a chaotic world, a birds-eye view, and the possibilities of transcending time and space. Pursuing these goals, however, while ignoring that ICT systems have physical and social consequences may be problematic in the sense that it tends to underestimate the costs of producing standardized data, and ignore the heterogeneity of doctors, patients, and computer systems. One of the effects of circulating data through a heterogeneous network is friction. Data friction occurs in the points where data moves between "data surfaces" (Edwards et al. 2011; Edwards 2010), as when transferring data from a spreadsheet to a language like XML, or from one institution to another. Friction impedes movement of data. Unlike physical friction, which is manifested in heat, data friction is manifested in data loss, misinterpretation, or costs in time, energy, and human attention. Establishing a network of surveillance systems, ICT-based or otherwise, will necessarily involve an amount of maintenance of the processes of data production and transfer. A conception of data that ignores the friction that occurs with heterogeneity, may contribute to shifts in resources allocated to health, privileging the types of knowledge required to produce and maintain ICT, and end up marginalizing doctors and patients.

Additionally, in health, surveillance seems to be suffering a shift towards increased market-driven solutions, led by companies selling consumers ideas of health and how to manage risk (French and Smith 2013). This shift calls for analyses of how truth discourses are produced and what subject positions they create for individuals.

2.4 The production of ICT

The object of study in this thesis, the process of producing data from ICT-based surveillance, crosses over several disciplines. For example, my work in design and software engineering was not set to describe phenomena, but rather to construct ICT systems that collect, process, and store data about humans. By contrast, my role within technology studies was to produce rich descriptions of the conditions of how an ICT system came to be produced in the way it was, as well as critically reviewing other systems in the same domain. The research described in this thesis is therefore socio-technical. The heterogeneous nature of ICT required, from me, to occupy a trans-disciplinary position, in order understand the different constituents. The knowledge produced through this kind of research is obtained from methods from disciplines as diverse as design, sociology, and engineering, each tangled up in different scientific traditions.

One idea that has been prevalent for centuries is that technology is applied science. The idea stems from a linear model of technological innovation that starts with basic research, which creates a type of factual knowledge materialized in text, which is then taken by engineers and materialized into technological systems. This idea has been contested by many historians of science (Grandin, Wormbs, and Widmalm 2004). The historical view proposes that the work of scientists and engineers is interrelated and does not necessarily respect disciplinary boundaries. Technological design is often almost oblivious to basic scientific theories, and creates both artifacts and knowledge framed within engineering language. Although I do not necessarily make a distinction between scientists and engineers, as they can be roles that the same individual may take, I do use the two terms to distinguish between the role of theory production (made by the scientist), which often takes the form of a literary artifact, and the role of material production (made by the engineer), which in turn takes the form of a (blueprint for) a physical tool. Many sociological descriptions of scientific activity portray scientific practice as an articulation of arguments and experiments carefully devised to convince the audiences, using the rules of a specific language game (Latour 1978; Latour and Woolgar 2013). Scientists themselves engage in boundary work: delineating boundaries between what counts as valid scientific knowledge and other types of knowledge:

> "...scientific knowledge is at once theoretical and empirical, pure and applied, objective and subjective, exact and estimative, democratic (open for all to confirm) and elitist (experts alone confirm), limitless and limited (to certain domains of knowledge)." (Gieryn 1983, 792)

A technological system is evaluated neither through such dichotomies, nor through a process of peer review similar to an academic paper. Instead, it is evaluated based on whether it works or is commercially successful (Latour 1987). In academic settings however, the theorization of the practical work and the formulation of scientific knowledge is as important as the developed system

itself. That knowledge can take many forms: it can be about the system itself, the process of designing the system, or the usage of it (Feast and Melles 2010; Cross 2001). One can place the practice of engineering between diametrically opposing stances: positivism and constructivism (de Figueiredo and da Cunha 2007) . On the one hand, the engineer uses analytical modeling to determine the workings of social and technological systems. On the other, engineering devises creative solutions in a dynamic and often corporate world of limited information and resources, and conflicting stakeholder requirements. Within computer science, areas like theory of computation and theory of algorithms use mathematical language to formulate theories. Other disciplines, such as software engineering, apply methods and processes to develop software that satisfies certain requirements (*IEEE Std 830-1998* 1998). Engineers often need not only to focus on technical work, but on creating interest in the technology they are developing, assembling raw materials, knowledge and capital; what Law (1987) has called heterogeneous engineering.

My work departs from the constructivist stance that sees technology as a political entity defined by the social actors that construct it and use it. A technological system can be perceived very differently by different social groups: it has interpretative flexibility (Pinch and Bijker 2012). The way that each social group, including the engineers themselves, perceives a technological system may generate new problems. For example, a CCTV camera may represent security for some, danger for others. Social constructivist models of technology development usually show that even seemingly purely technological components are always dependent on the social groups that give them importance, or see them as a problem. Technology is a political entity. As such, it is amenable to change, and that change can be localized in the social actors or the processes that shape it. Alternative solutions can be sought by, for example, giving priority to different stakeholders and allowing the design process to be permeated by different values and points of view. It is also compatible with the constructivist end of engineering practice. The flavors of engineering that function within positivist stances are less adequate for tackling complex, "wicked problems" (Rittel and Webber 1973) that appear in the design of technological systems.

2.5 A theoretically informed research path

The first part of this section refers to the notion that 'data is out there': it is as real as the real worldly thing that it refers to. This is a common notion among ICT designers and engineers, but a notion that I will replace by the concept of apparatus. Foucault defines it clearly in one of his later interviews:

> "What I'm trying to pick out with this term is, firstly, a thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions–in short, the said as much as the unsaid. Such are the elements of the apparatus. The apparatus itself is the system of relations that can be established between these elements." (Foucault 1980)

ICT-based health surveillance is an apparatus of biopower. It contains certain discourses that define what can be measured, and what is supposed to be measured. The health surveillance apparatus is an ensemble, a network, of many things: discourses, ICT systems, institutions, and human actors, e.g. epidemiologists. Data is one of the things produced by the apparatus. An apparatus of biopower, composed by institutional, material mechanisms and knowledge structures, produces truth discourses about wellness, life, and death. Individuals and populations are described and classified by it. If one is to design socially responsible ICT, an important question becomes how this apparatus comes to exist, who produces it, and why. I approach these questions from the perspective of a software engineer, who produces ICT (a part of the apparatus):

> "A technological artifact can be seen to be a set of features bundled together into an identifiable and bounded package. Such features may be material (as in a tool such as a hammer), conceptual (as in a technique such as TQM), or both (as in a computer imaging device and its prescribed radiographic procedures). Artifacts are usually designed and developed under the auspices of explicit objectives and to some particular specifications." (Orlikowski 1995, 2)

Analyzing how these objectives and specifications become part of ICT, and how ICT is made to produce health data, is the goal of the first research question. My choice to focus on ICT systems that produce health data does not mean that I can ignore all other components of the surveillance apparatus, since ICT systems bring together all those things. They are more than material objects. They are artifacts: networks of things. Fitbit brings together the notion that health can be individually pursued through a collective game where people achieve fitness goals by sharing data taken by a wrist-worn device, for instance. The specific data that Fitbit collects, how it is represented, and who is given access to it, is all brought together by propositions on how to achieve a fit body, how to combine reliable measurements of data with a stylish bracelet, etc. And here is where we can start to see that data is not something that is just out there. Context and perspective of the different aspects of production of data can be gained through historical research (Shapin, Schaffer, and Hobbes 1985), or ethnography of laboratory work (Latour and Woolgar 1979). Data is produced through the interaction between ICT and the real. If we take an organ, a notion of periodicity, technological characteristics of devices such as stethoscopes, Holter monitors, or wrist watches, and the assertion "Listen to your body" (Polar 2011), one may end up with heart rate data measurement as a common practice. More specifically, we will end up with a Polar Wrist Watch (Pantzar and Ruckenstein 2014). The same description could be made for maps, with their different projections, different choices (or politics) when inscribing borders, geographical landmarks, roads, etc. I chose the examples of maps and heart rates because they exemplify the ground truth in the case studies I describe in four of the papers included in this thesis (Papers I, II, V, and VI). My ontological stance does not deny realism. But it entails that our ways of understanding the world necessary need to go through many layers of meaning-making, which define what and how can we understand the world. They are acquired through apparati. I propose the visual metaphor of a grid, to think about apparati. This metaphor does not intend to replace or even complement other

metaphors such as assemblages, rhizomes (Deleuze and Guattari 1987), or diffractions (Barad 2003), which other researchers have produced in order to model or describe representations. Unlike them, I do not build on the metaphor in order to produce theory. Instead, I use the metaphor of a grid merely as a visual aid to pedagogically describe the work in this thesis.

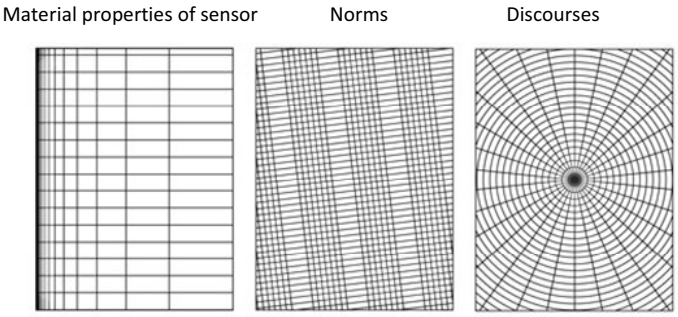

Figure 1 – Examples of elements of data production, adapted from (GeoMedia 2014).

One can imagine that each grid represents an element of an apparatus such as health surveillance, as seen in Figure 1. Each grid can be a discourse, or a technological artifact like a bio-sensor, an application that records cell-ids, an institution, a scientific statement, or a law that regulates how is data to be shared with third parties. Each grid defines possibilities of measuring the world, and limits others. The combination of grids in Figure 2 shows the white plane as the world, and the grids representing several methods of accessing it. Intersection of the grids represents what is possible, or desirable to obtain: the data. The white spaces between the grids are blind spots. They do not produce data, and are invisible, or unmarkable, in this apparatus. Here it is important to remark that to combine is not the same as to superimpose. Combining the discourses, material properties, etc. requires work and that is what we, as engineers, do when we produce programming interfaces, protocols, or algorithms that compose ICT systems, as well what the individuals do when making sense of it. In the grid metaphor, that looks like a projection.

In conclusion, data is not just out there, it is constructed through articulation (in the grid metaphor, articulation is projection) of discourses, material properties of ICT, interests of actors that define what should be considered data, or how can it be measured, and historical context.[3] Such articulation is an exercise of power.

[3] There is a quote from Heisenberg that illustrates this point quite well: “We have to remember that what we observe is not nature herself, but nature exposed to our method of questioning.” (Heisenberg 1958, 44)

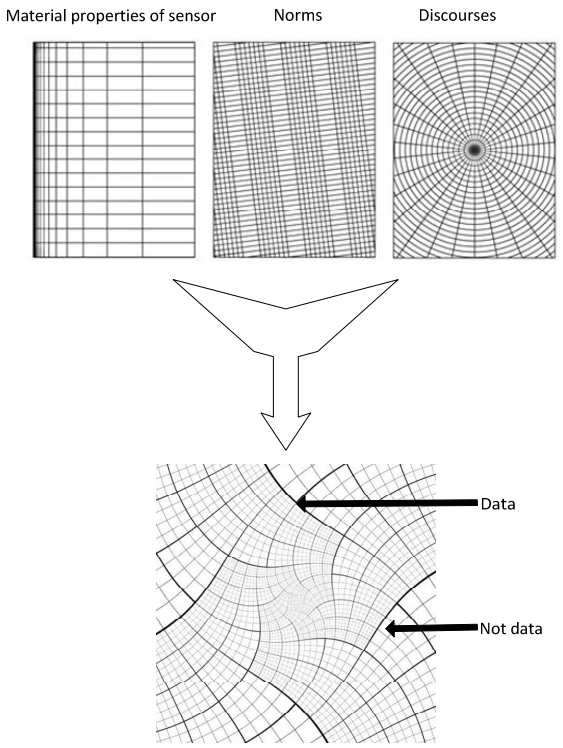

Figure 2 – Through the practice of software engineering, each element is combined, and possibilities of representation are defined.[4]

2.6 Justifying the problem and research questions

One could have the intent of eliminating all the blind spots, which is obviously impossible, but nevertheless attempts are made[5]. But that is not what I want to do. Instead, I follow a call for action that takes the stance of an individual, standing in the intersection of networks, and the consequences that come from that intersection, or lack of it. Ball writes about the concept of *exposure* (2009), which she proposes be used to define the experience of the subject of surveillance. Here, the individual is depicted as unambiguously under surveillance, even if unaware of it, and enrolled in a network that extracts and processes data from her. That surveillance is justified and constructed by specific configurations. She writes:

[4] The grid images are adapted from (GeoMedia 2014) and (De Smit et al. 2003).

[5] The US National Security Agency appears to be certainly doing so, bringing together a global-spanning variety of actors such as ICT-based systems, private and governmental institutions brought together by, at least, an anti-terror discourse.

> "the body interior of the surveilled subject is more open to division, classification and scrutiny, either because it is seen as a source of truth, the target of public revelation or fetish…[and]…the nature and character of exposure is a product of institutional configurations which have consequences at the level of the individual." (Ball 2009, 640)

Among other things, Ball calls for studies of "the origin, nature and coherence of institutional values, technologies and practices which produce the political economy of interiority" (Ball 2009, 653): what causes the need to see certain aspects of the surveilled subject? In a different argument, Star writes about the experience of being allergic to onions (1990) and her invisibility to a standardized network of food production. Here, the experience can be seen as complementary to the experience described by Ball. The enrollment in a certain network may or may not be voluntary, and one individual may be part of multiple networks. She goes on to describe how an individual negotiates (un-)enrollment in one or multiple networks is a question of power. According to Star, two main approaches in science have been used to observe the issue of power in technology development: analyses that focus on how different actors translate other actors' interests into their own, and how actor-networks are kept disciplined and stable (Latour 1978), and analyses that focus on effects of technology upon specific groups (Woolgar and Cooper 2011). The latter includes aspects related to access (or lack thereof) and discrimination of groups such as women of all colors, ethnic minorities, the poor, the physically disabled, etc. These approaches are not easily combined, since the first one typically tells the story of engineers competitively assembling different interests in materials in order to accomplish their goals, a story which omits social groups which in turn are often underrepresented (or made completely invisible) in that process. There are also methodological issues in combining descriptions from multiple heterogeneous viewpoints. As a way to overcome the divide, Star calls for analytical approaches that account for the work of keeping the network of humans and physical objects stable, a work that usually produces loss of multiplicity in order to create standard representations.

Ball and Star identify similar problems, different facets of the same prism. These facets are: visibility, invisibility, selected visibility, reification, and enforcement of visible characteristics. In sum, the prism describes one individual negotiating how she is to enroll in a network that attempts to see her. In the work included in my thesis, this is a network where software engineers have a particularly preeminent role. When engineers produce and maintain a network of actors, ICT, and materials, they create a specific artifact. Using my metaphor, this becomes part of a combination of grids (an apparatus), from which individuals are looked at, and classified. Problems like privacy intrusions, social sorting, and invisibility of minorities can be reduced to consequences of that classification system. This reduction entails loss of complexity, but there are gains in simplicity:

- Privacy: the need to make visible certain characteristics of an individual can be justified by a multitude of discourses, ICT systems, and things that create the need to see certain things. These can be, for example, marketing, anti-terror discourse, or an IP logging system. Other things, like Data Protection, personal integrity, PGP, and k-anonymity are examples of elements can be either juxtaposed to the apparatus of surveillance, or combined with it.

Consider8 (the project that is described in papers II and VI of my thesis) is an example of a combination of controlled seeing: the need for seeing groups without seeing individuals.

- Social sorting: classification systems that reify problematic categories. Individuals are classified in specific ways for a specific purpose. The data collected about individuals reflects that purpose. Examples include race and gender. Sorting can contribute to problematic classifications, or in extreme cases reproduce social inequalities that limit the opportunities of individuals: they may be trapped in the grids that were made to classify them, or may be able resist them.
- Invisibility: an apparatus that is made to see health threats as sudden unexpected events is not able to see other kinds of health threats, or individuals suffering from them. These are the often unacknowledged blind spots of the discourse of syndromic surveillance (Mandl et al. 2004).

I seek to describe the grids of the projects I worked in, in order to be able to explain (this is a secondary goal, in the second research question) the implications of why or how some of the representations we choose are likely to make more or less visible some social groups and problems. In order to describe these issues, I describe also the work of combining the grids from the point of view of a software engineer.

3 Methodology

There are a number of critical studies in sociology that explore dimensions in ICT-based health surveillance (French, 2009; Lupton, 2012) albeit without particularities on the work of software engineers in this area. The immaturity of the literature on ICT-based health surveillance did not allow for a completely deductive research approach. This justifies the exploratory aim, and the necessity to cross disciplinary borders into history, and science and technology studies. The meta-analysis of the papers is deductive, however, and is reported on in Chapter 5.

Crucial questions for knowledge production concern reliability and validity. Which epistemological values are required to validate the knowledge produced in this thesis? Longino's (1990) discussion on objectivity in science is illuminating, both to the aim of this thesis, and to my methodology. She argues that the social character of the scientific practice, through a constant evaluation from peers, both in the process of making an experiment, peer-review, and subsequent uptake or rejection by the scientific community, produces a form of objectivity that is quite different than the one espoused by logical positivists. Objectivity to her is defined as a social process that enables scientific communities to critically evaluate and modify the work of others. It is through this process of community response that background assumptions can be questioned and re-evaluated. In order to make my work as available for scrutiny as possible, I reflect on my assumptions, and expose the processes, as I see them, of knowledge and technological production that I participated in. I aim for a strong form of objectivity (Harding 2002) and situated knowledge (Haraway 1988). Lyotard's conception of the post-modern society is one where knowledge is produced by those who control vast and complex computer networks (Lyotard 1984). In this type of society, knowledge can be legitimated if it produces new types of language games, going against a established rationality. Influenced by Foucault and Lyotard, my goal is not to produce consensus, but dissensus. I aim for a critical approach, one that provides alternative views of approaching the domain of ICT-based health surveillance.

3.1 Research environments

The work included in this thesis has been conducted in several institutions, and my committee includes people with different disciplinary backgrounds. I benefited from external funding from FCT, which allowed me relative freedom to select projects and pursue independent lines of work. However, because I had earlier completed my masters degree in KTH based on work done at Mobile Life and SICS, I remained in the same institutions and followed up on earlier work.

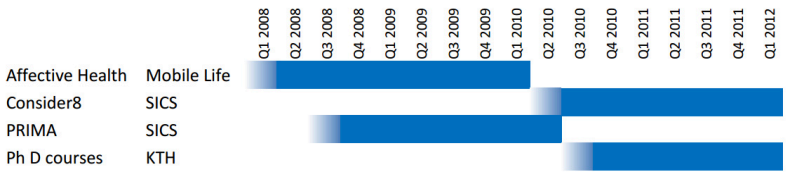

Figure 3 – Timeline of projects included in this thesis.

The earliest work reported on here was conducted in Mobile Life, an academic and industrial center with a focus on mobile services and ubiquitous computing. In Mobile Life I worked on the Affective Health project, which had as principal investigator Kristina Höök and later, Marie Sjölinder. This part of the work was influenced by the disciplines of interaction design and software engineering. Later, I conducted work in the Userware research lab, led by my main advisor Magnus Boman, at SICS; a non-profit research center for applied ICT. In the SICS project Privacy in the Making, with principal investigator and industrial co-advisor Markus Bylund, I realized the first explorations of the mass-tracking system, then within the Privacy in the Making (PRIMA) project. Consider8 was done in collaboration with Ericsson Research. Here I adopted a more traditional role of software engineer, combined with the role of a privacy analyst. Throughout, I took graduate courses at several different departments at KTH. Most of the analytical work was influenced by co-advisor Nina Wormbs and colleagues from the Division of History of Science, Technology and Environment, a post-disciplinary research division with a focus on historical studies in science and technology. The practical analytical work was done in collaboration with co-advisor Louise Barkhuus, from Stockholm University in affiliation with Mobile Life, with an orientation towards human-computer interaction.

3.2 Research trajectory

	Technological	Analytical	Reflective
Self-Surveillance	Paper I	Paper III	Paper V
Mass-Surveillance	Paper II	Paper IV	Paper VI

Table 1 – Contributions of this thesis (repeated here for reasons of pedagogy).

The trajectory was guided by the exploration of the research questions. The first question, “How is health data produced by ICT-based surveillance?” was initially pursued from the perspective of software engineering. The aim of this research was to produce systems, and the reported academic results work as a positioning of these systems. However, I did not find these to be satisfactory, as some facets of mass-surveillance and self-surveillance were not considered. This led me to pursue analytical research, assisted by secondary material and theories from technology studies and surveillance. The analytical work aimed at answering both research questions and was aimed at

systems in both self- and mass-surveillance. Finally, the reflection work relied on historical sources. It was theoretically informed and deductive. The reflection work uses technology studies methods to look at the results reported in the technological production part of my thesis, and considers both research questions simultaneously. Together, these methods contribute to the production of descriptions of the data and ICT systems produced during my thesis. Finally, I produce an additional analysis at the end of Part I of this thesis. That analysis draws on theory and examples of current work by other practitioners to bring together the previously independently pursued domains of self- and mass-surveillance.

The questions asked in each project were: "How can one represent stress?" and "How can one represent privacy-sensitive mobility patterns?" These questions have a direct link to the first research question. This part of the work is guided by the research question and practice-led, rather than informed by theory. It followed an inductive research method. Both projects rest on other research and development projects that had been concluded (Bylund et al. 2010; Ståhl et al. 2009), which generated primary sources in quantitative (algorithms, requirements) and qualitative (reports and exploratory inquiry). Both projects were anchored in literature studies that were conducted deductively, attempting to integrate views from psychology (Paper I) and sociology (Paper II).

There were two main differences between the two projects, which resulted in slightly different methodological choices. The first is that the work reported on Paper I was purely academic, and did not, during my participation, culminate in a product[6]. This contrasts with Paper II, which had an initial focus on innovation, and a production of a patent application, and ultimately aimed at a product. The second distinction lies in the disciplines and traditions of the two projects. The team of Affective Health (Paper I) was composed mostly by software engineers, psychologists, and interaction designers. Our iterative design methodology was based on the tradition of interaction design (Zimmerman, Forlizzi, and Evenson 2007). In Consider8 (Paper II), the team was composed of software engineers, and later, a business analyst.

Affective Health raised questions about how individuals make sense of data on their own bodies, and how this is affected by design choices. The Consider8 team had raised questions regarding the necessity of considering other factors than privacy when producing surveillance systems, and how these could be located in the design process. The aim was still exploratory, and the methods were inductive. This part of the work was done to pursue simultaneously the two research questions: aspects of the production of ICT and how it is made to represent phenomena were considered, and possible implications were proposed. The implications were theoretically deduced. Two methods were used: content analysis (Paper III), and discourse analysis (Paper IV). Both are based on publicly available text that was considered adequate for exploring both the production and the implications of ICT-based health surveillance. Different analytical methods were chosen because the text differed significantly: one was structured in the form of online posts and threads, and the other was unstructured.

[6] The project did result in a spin-off company, based on an updated version of the system (Biosync 2014).

First, I detail Paper III. I focus only on the method that applies to my contribution, refraining from describing the foresighting part that was done solely by Boman[7]. Fitbit was chosen as a case because it is a widely known system, and the most commercially successful fitness tracker, at the time of the study. The approach to the research was theoretically informed. Our aim was to explore how data was produced in the usage of the fitness tracker. This guided the choice of material, as the forum was found to be rich in instances where users of Fitbit helped each other solve a variety of problems. We focused on publicly visible posts posted within the period of February and March of 2014, which formed a total of 656 threads. We used conventional content analysis (Hsieh and Shannon 2005), which is an interpretive approach to the analysis of text and consists on the following initial steps: 1) initial immersion in a text corpus, in order to get acquainted with the material, 2) inductive category generation, and 3) classification of text. More steps could have been considered, such as refinement of categories, and re-classification. However, we found that not a lot of variability was found within categories. After two researchers had categorized approximately 50 threads according two agreed upon categories, one researcher categorized the rest with sample testing by the second researcher. A few threads contained topics of more than one category. After the text was categorized, we described and related each category to the others, as well as to torque.

In Paper IV, we analyzed the production of truth in the discourse of the developers of SIDARTHa. We relied on a definition of discourse by Foucault: "a regulated practice that accounts for a number of statements" (Foucault 1972, p. 80, cited in Mills 2004, p. 6). Seven project reports were taken as the unit of discourse: the text. Following a method of discourse analysis espoused by (Jørgensen and Phillips 2002), we marked sentences that made truth claims about entities that existed outside of the text, or referred to individuals or groups of people, in order to define subject positions. These were annotated, interpreted, and written down. My co-author Baki Cakici had the necessary prior knowledge in syndromic surveillance to bootstrap the analysis. The analysis was based on inductive and deductive reasoning. The deductive part relied on previous knowledge of the domain, and was guided by the question "What is considered to be true?", for each sentence. The inductive part consisted on grouping the marked sentences into common themes, based on the topics found in the reports.

In the final stage of my thesis, I returned to the systems I made, with the knowledge from the systems I analyzed. In this part of the thesis, I positioned the work[8] I made earlier in papers III and IV: what contributed to the problem formulations, why we chose to look at that particular data, and why data was processed in those particular ways. I tried to show how it could have been done otherwise.

In paper V, I re-visited the design process of Affective Health. I did so by building on a historical review of stress and emotion, using secondary sources of historical research. The research was purely deductive. The final outcome of this research was an essay, a personal reflection assisted by secondary sources of data. Future work is planned for transforming the essay into an overview

[7] The methodology for this part is described in (Boman 2014).

[8] Positioning the work is in itself a kind of work: an articulation. But it is not a type of work that aims to produce a system to fulfill a function. This is transversal work, based on critique.

paper of the area of quantified self, in the case of stress management, for an HCI audience. The planned steps are: doing a systematic literature review on the last ten years of HCI themed conferences—this would include all from ACM SIGCHI, and possibly some additional papers from IEEE—for systems on the topic of stress and stress management, relate them to relevant literature in sociology of health, as well as relate them to the historical lines of stress and emotions in HCI.

In paper VI, I re-visited the early stages of the design process of what then became the Consider8 project, described in Paper IV. This phase was previously unreported, and consisted of setting the stage for the production of the algorithms that followed in later stages. I departed from the technology studies perspectives on ICT construction, particularly the observation that scientists and engineers create abstract immaterial knowledge by deleting the localized material context through diverse material-discursive forms. I also had in mind the second research question in the thesis: What are the consequences of the work of engineers for how individuals are represented, and what are the ethical implications of these? Therefore, I chose to focus on the work ideally done in early stages of applied research projects that aim at defining the problem, and specific requirements. The analysis of these objects is intermingled with a narrative that connects them. The supporting materials for producing this narrative are the over 700 emails between project members sent around the period, many containing attachments of code, visualizations, notes summarizing meetings, and slideshow presentations. Project members were also contacted during the analyses in order to clarify issues. Some of them have also reviewed early versions of the paper. The choice of objects to analyze, and the process of producing the narrative is inherently subjective and situated. My participation is fully acknowledged, by using the pronoun “we”, as I describe the actions that were taken. Depicting the project members as a unit and the analyzed objects as a product of collective action reduces the complexity of the analysis and is sufficient, since my focus is on the objects, rather than the team dynamics. I follow the objects as they are made to encode knowledge, built through a process of inscription, and as they are anchored in order to represent a human mobility. A limitation of this paper is that it presents an *ex post facto* analysis. A richer description could have been made, and perhaps interventive action could have been taken had I done it as the design process was ongoing. However, this limitation is related to my progression as a student, as my theoretically informed enquiries on ICT production were formulated later in the process of writing this thesis.

3.3 Ethical considerations

During the research conducted in this thesis, efforts were made to ensure fair principles in collecting and disseminating data related to individuals, by following the CODEX rules and guidelines for the humanities and social sciences (Bioethics 2014). Two issues deserve special mention. One ethical issue was related with development of Affective Health (Paper I). Particularly, during the design process, we questioned whether representing stress to an individual would cause the individual to be more stressed. No practical ethical guidelines were employed in the project. Following the advice of the psychologists who were consulting with the project, we conducted a user study, n=10, (Ferreira et al. 2008) that aimed at investigating the performance of the interface, as well as whether the

system itself causes stress. General conclusions cannot be made from such a small sample, but we did not find any evidence of the device causing stress. In addition, special care was also taken for not including, in any study, people who reported to have suffered from a stress-related illness.

Another issue was whether the communication in places like forums should be considered public or private. Doing text analysis on online communication has been an object of discussion on ethics (Kozinets 2001; Langer and Beckman 2005; Whiteman 2012; Im and Chee 2006). The paper on Fitbit (Paper IV) deserved special attention, when it comes to ethics, because it consisted of a secondary use of information obtained without explicit consent. Consent from each participant was not applicable due to the volume of the posts analyzed. However, even if they were accessible, the people posting them may not be aware of the visibility that their posts have, and most certainly did not have researchers as intended audience. In attention to that, special care was given to selecting quotes from forum participants to figure in the paper, as representative of each category: 1) all user names were anonymized, 2) all the quotes referred to the relationship between an individual and the fitness tracker, 3) they were not unique, since other quotes from different participants could have been selected with very similar content, 4) none of the quotes include information related to location of the person, or related to a personal health issues. According to Frankel and Siang (1999), Internet research should balance benefits and risks of collecting data from individuals. This study was motivated to directly benefit the individuals who participate in activity tracking, not Fitbit or any other public or private entity.

4 Contributions

This chapter details the research contribution of each paper included in this thesis. It also details how I, contributed to each paper.

4.1 Paper I - Mind the body!

Here was presented the work of designing representations of stress from bio-data collected in everyday settings while combining two sets of requirements: medical descriptions of physiological reactions to stress, and technical difficulties of measuring bio-data data outside of a controlled environment. We positioned ourselves in contrast with a worldview that would attempt at gathering data, passing it through classification algorithms and deriving a level of stress. We argued against that position using arguments from our design exploration, which entailed considerations related to sensor reliability, and medical understanding of stress. Building on an embodied interaction design stance, which aims at producing systems with a focus on the duality of mind/body, subject/object, or knower/thing, we took the constructivist side of the controversy related to physiological mapping of human emotions. As a result, we presented several representations of bio-data and proposed that when designing self-tracking systems for everyday life, one should include, as parts of the representations: openness for interpretation, fluency, and interactive history, to facilitate understanding of bio-data and foment self-reflection. I wrote my M Sc thesis in the Affective Health project, based on a literature review of methods for acquiring stress measurements using biosensors (Sanches 2008). After the writing was concluded, my participation continued during the initial phase of my doctoral studies. My own work at that stage consisted specifically on producing algorithms that translate sensor data into representations that, through the work of the interaction designers, make the sensor data available to users' own interpretation. This work produced a number of publications not included in this thesis, e.g. (Silvăşan et al. 2010). Most of the text in Paper I was written by Kristina Höök and me; the other authors have contributed with some text, corrections, and proof-reading.

4.2 Paper II - Knowing Your Population

Here, we presented the rationale for the design of a system that represents mobility patterns from mobile phone data. We provided an analysis of: 1) possible application areas for using network data (including a syndromic surveillance system), 2) technical characteristics of network data, and 3) the legal landscape regarding data protection and privacy considerations in EU. We showed that using time geography as an ontological framework for modeling mobility patterns from network data

allows for a compromise between the three sets of requirements above. Our conceptual solution materialized in particular technical solutions, also struck a balance between privacy sensitivity and utility. Utility was demonstrated for the particular case of transportation planning, and proposed for syndromic surveillance. The work on the Consider8 project was initiated in 2008 at SICS. During the period up to 2012, I took several roles: mobile developer, back-end developer, and systems administrator. I also accompanied the development of the project as a technical advisor. This paper details the project contributions during the period when collaboration had been established with Ericsson Research. My main role, other than participating in the development of the technological contributions, was to elicit requirements and articulate them in terms of privacy vocabulary. The paper was first drafted by Svee and me, with contributions from Bylund, and later re-structured by Boman. The other authors also contributed with text and proof-reading.

4.3 Paper III - Sensemaking in Intelligent Health Data Analytics

Fitbit design documents bring together lifestyle statements intertwined with medical statements on how to achieve a fit body. It produces a game where individuals share their data with each other for the purposes of motivation. Data is collected and processed for that purpose. In this study, we looked at how users encounter such data about themselves, and the work they do in order to interpret or appropriate the data. This paper produced evidence of the complexity of the negotiations of users in the network of meanings that Fitbit produces. Through an analysis of forum posts, we showed that the practice of self-tracking acts on the lives of individuals in ways that are not necessarily the same as those advertised by the fitness trackers. For example, using Fitbit as a tracker to achieve a healthy lifestyle requires a specific set of actions to be performed (constituting a form of disciplining). Activity trackers divide and categorize physical activity of individuals in non-obvious ways that require an effort in interpretation from the part of individuals, and individuals were found to subvert the functionality of the tracker in order to create alternative information, co-constructing through the online forums new definitions of correct and accurate data. Additionally, we discussed the implications of producing aggregated data from Fitbit users, based on the visions of digital health analytics as suggested in a foresight exercise by EIT ICT Labs. Baki Cakici helped me to frame the initial problematization, and contributed with background and ideas. The collection of material for the Fitbit case study was done by me, as well as most of the analysis. Louise Barkhuus assisted in the whole process, and did the initial inductive coding with me. Finally, Boman and I edited the paper together.

4.4 Paper IV - Detecting the Visible

We analyzed documents detailing aspects of the design of a European syndromic surveillance system. The discourse of SIDARTa brings together a variety of actors (International Classification of Disease codes, ambulance dispatches, syndromes considered relevant, an event of a volcano eruption) in order to construct a concept of a health threat. We describe the work necessary to construct a health threat from data from multiple sources, particularly from ambulance dispatches. We showed that the concept of a health threat, as it is discursively constructed in the analyzed

reports, refers to a sudden and unexpected event. Furthermore, the described system requires a specific kind of work of maintaining material information systems, and information flows. We argued that attention should be directed towards how information systems are to be maintained (the material and social aspects of systems) and the blind spots of this particular construction of health threat. Not doing so can render invisible important aspects of health care. The textual material was collected by Cakici. He did most of the analysis, and I helped him in later stages. The first draft of the text was written by Cakici, with text contributions from me, and I also contributed to subsequent drafts.

4.5 Paper V - Positions in technology design

I showed, based on secondary historical research, that stress can be seen as social problem, located in structures at home and work and all the places in between, or it can be located in the inability of the individual to adapt to causes of stress. I described some paradigms in how scholars in the area of Human Computer Interaction have dealt with the issue of representing emotions, and how our representation of stress is framed for those lenses. Finally, I discussed some implications of the decisions we made in the design process of Affective Health, based on a secondary critical sociological analysis of mobile health applications. I argued that design decisions and assumptions reify stress as an individual problem to be solved by the individual. The belief that stress can be tackled individually may penalize individuals who do not have control over their life circumstances, and preclude other ways of tackling stress and foment deeper societal change. I did all the secondary research, analysis, and writing. This paper is included in this thesis in the format of a standalone essay.

4.6 Paper VI - Data-driven knowledge production in software engineering

In this paper, I reflected on the implementation of Consider8, particularly on the initial phase of defining requirements. This analysis is based on the social and material aspects of the mobility aspects we have constructed. Special attention was given to how application areas were defined for mobility patterns, and the intersection of material and social aspects of network data (particularly its dependency on mobile phone usage, or existing infrastructure). In order to transform mobile phone data into mobility patterns, a number of steps are taken where the data is manipulated in specific ways: that transformation entails the deletion of contextual aspects related to mobile phone communication, physical (material) aspects of the mobile network infrastructure, and identifiable individual information, in order to gain the ability of making general statements about mobility. This deletion contains specific assumptions about what constitutes valuable mobility patterns. Not considering these can sort and render invisible parts of the population. I did all the secondary research, analysis, and writing. This paper has been submitted to a journal, and is currently under review.

5 Discussion

My answers to the two research questions can now be given, based on the papers included herein, and guided by the theoretical concepts I presented in Chapter 2.

5.1 How is health data produced by ICT-based surveillance?

A starting point for answering the first research question is the observation that the production of health data is neither fully located in purely discursive elements nor in particular technological constraints. As an example, based on Paper III, analyses of discourses of obesity within health can give insights into how particular subject positions are created for individuals, by for example having to control calories by limiting food intake and doing exercise. These discourses have been found to be problematic (Rich and Evans 2005). Similarly, analyses of technology such as biosensors can only answer questions related to accuracy, data storage, and similar matters. Even if one applies a discourse of privacy on data storage, or one of its materializations (e.g. cryptographic algorithms), one would still not be fully equipped to explore the implications of a biosensor-based system that aims to tackle obesity. One needs to explore how the discourse shapes materiality and how materiality shapes discourse, through the practice of software engineering.

The engineering papers of this thesis represent the work of bringing together different elements: medical documents, legal documents, bio-sensors, mobile phones, privacy considerations, stress definitions and time geography, into two systems with specific characteristics. In papers I and II, work was done to assemble these elements into a coherent argument and to position our systems in relation to legal requirements or wearability requirements, of some kind or another. Through both the technological production and the academic work, stress data and mobility patterns were produced in ways that combine discursive and non-discursive aspects. These were then problematized in Papers V and VI. Paper V showed how knowledge about stress is evolving through conflicting discourses: it is simultaneously a phenomenon located in the body, and a societal problem. Each discourse has the potential to produce different subjectivities. Looking at stress purely as an individual phenomenon produces a subject that should take care of herself, whereas a societal phenomenon could produce subjects that aim to attack the causes of stress. In Affective Health, the first definition was adopted, and a system was designed for individuals to reflect on their own bio-data. Designing a system like Affective Health is an exercise of power that entails taking specific positions on controversies about how to represent emotions. These controversies are not only guided by epistemic values, such as accuracy and truth. They are also guided by cultural values, not only on what can be represented but also what should be represented.

Similarly, mobility patterns can also be seen as objects constituted through ICT artifacts, as a network of discursive and non-discursive elements. Paper VI provided a description of how they were constructed in a system that produces mobility patterns. I showed in this paper how privacy-sensitive mobility patterns were produced through an alignment of inscriptions that began with annotating mobility through GPS, hand-written notes, and network data, and how these annotations were then connected to wider discourses on privacy, and usefulness of mobility patterns. By considering the legislation related to fair information practices, I showed that privacy through individual anonymity does not adequately address the ethical implications of the production of this type of data. This resonates with the concerns of other researchers, such as Coll (2014), who argues that individualistic definitions of privacy may in fact be allies of surveillance, contributing to legitimize its expansion without consideration for wider democratic values, such as representation. Through the processes of describing how the objects of stress data and mobility patterns were produced, I also showed how heterogeneous the work of a software engineer (Law 1987) can be. Producing data depends on factors that are not only within the realm of technical puzzle solving, and calls for engineers, and those involved in the process of producing ICT, to be aware of the wider implications of the work of producing knowledge about humans.

A characterization of the artifactual nature of ICT, and the process of inscribing phenomena in algorithms, was for me a way of addressing the implications of the way that data is brought into existence. Espeland and Stevens (2009) argue that, for ethical dimensions of quantifications and data to became visible, the process of production must also become visible, because it is in that process that one can find problematic assumptions. This point is also raised in Ball's (2009) problematization of exposure of an individual to an apparatus of surveillance. An observation related to this point is that health data can neither be fully located in the design of ICT-based surveillance, nor limited to the moment of when an individual is exposed to it. Data is produced both in the process of ICT production, and in the process of exposure of individuals to ICT. It is an ongoing process of production and interpretation. Although I have not focused so much on the process of production through use of ICT, in Paper III we provided evidence of that process, and how it can be facilitated or complicated by the design of ICT.

5.2 Which entities are made (in)visible in ICT-based health representations?

The question of invisibility cannot be divorced from the question of who benefits from a particular representation. Here, I side with the individuals implicated by the data, the subjects of surveillance. During my research, I investigated the potential of users and subjects of surveillance to negotiate exposure to ICT-based surveillance. When I address invisibility and problems of representation, I am interested in the problematic aspects that can affect the individuals under surveillance. It is important to acknowledge this in order to understand the answers below, which otherwise could have focused on e.g. how aspects of data production would affect a possible plan for monetization of data produced by users of Fitbit. According to my stance, I have shown evidence of three types of things that can be made invisible in ICT-based health representations. These are similar to what

French (2009) calls "local orders of concern": the concerns of individuals—including doctors and patients—implicated by health surveillance. However, these are derived from my analyses in Papers III through VI:

1. Broad determinants of health – By prioritizing certain scenarios and definitions of health and wellbeing over others, or by following data-driven approaches, one runs the risk of not considering broader determinants of health. The latter are not so easily quantifiable, and can have a more positive impact on the lives of individuals implicated by surveillance.
2. Individuals' own conception of self and wellbeing – The reliance on schemes to classify individuals' emotions, fitness, or other phenomena into particular metrics may marginalize conceptions of self among individuals.
3. Population strata – Through the work of those implicated in the design of ICT for surveillance, discourses and materials are assembled in ways that can render invisible parts of the population, while prioritizing others.

The first item refers to the discursive work done by engineers, or truth-preserving discourses used as resources to design or engineer ICT. When data is produced, part of the work of engineers is to produce scenarios of usage, produce ideas of health and wellbeing (such as health threat = unexpected event), and design towards achieving them. That work was shown, in Paper V, to exclude societal determinants of stress, in Paper III to exclude broader determinants of the health of a population, and in Paper VI to exclude alternative ways of representing mobility. Producing scenarios is part of the work of constructing accuracy (MacKenzie 1987). The process of producing data starts with attention to previously defined goals. When data is finally produced, e.g. when a sensor is tested, or a data packet is stored, it will be matched against these. Even so called data-driven approaches that merely intend to find patterns in a sea of data will also be influenced by the context of the raw data that is used as input, and by the assumption that what can be measured is what should be measured.

The second item refers to the impact that surveillance can have on individuals' conception of self. Classifications of stress and emotions (Papers I and V), or classifications of bodies that depend on static classification schemes (Paper IV) can cause torque (Bowker and Star 2000) in individuals' lives by rendering invisible parts of them. The concept of data friction (Edwards 2010), which is also an implication of data production in its own right, can be seen as a lens for viewing implications for representation. First, data friction can refer to the work of carrying a device, calibrating it, maintaining battery, synchronizing it, or configuring it in order to achieve the goals of measurement. Engineers and designers can find solutions to diminish friction or facilitate this type of work, either by working out new kinds of sensor design, new types of batteries, inventing easier ways of interacting with the system, etc. But a conception of health that is dependent on subject-positions of individuals managing their own body through ICT-based self-surveillance creates subject-positions of individuals that are also now managers of data friction. How proficient they are at managing it will have consequences for how they are represented. For example, if individuals are

able to understand how a device works and are able to appropriate it, they will be able to control how they are seen. In Paper III, we provided evidence of individuals actively appropriating the measurements done on them. The technological artifact of Fitbit has interpretative flexibility (Pinch and Bijker 1984). In Paper I, we designed specifically for those kinds of appropriations to take place, albeit within the conception of the subject of surveillance as a reflective individual who is able to take steps to overcome her condition of being under stress. Other individuals though, may find the task too daunting, or incomprehensible, and be affected by it.

Regarding the third item, data friction is also useful to understand wider consequences of mass-surveillance. In Paper III, we problematized that aspect, and proposed that syndromic surveillance may add to the list of professional skills necessary in public health, technological work, and data management. That may have consequences for how economical costs are distributed. In the mass-surveillance systems I focused on (Papers II, III, and VI), aspects of how individuals are to negotiate their exposure to the systems were not considered. This was because these systems were making use of already existing data, which in turn was a by-product of mobile phone communication (Papers II and VI), or things like ambulance dispatches (Paper III). Making use of already produced data, in a responsible way, entails paying attention to how and why was that data produced in the first place. There are three reasons to do this. First, it can violate contextual privacy (Nissenbaum 2009): by re-purposing data, one must pay consider that data considered acceptable to use for the purposes of the primary collection will not necessarily be acceptable when re-purposed. We addressed this problem in Paper II, by proposing data anonymization. Second, because the data used as raw may be skewed towards particular strata of population. This fact alone can contribute to social sorting (Lyon 2002). Third, because re-purposing data that is collected for other purposes complicates how individuals manage their degree of exposure to surveillance. I illustrate the latter two problems in Paper VI, particularly by showing how managing exposure would require dramatically changing mobility habits, or mobile phone usage habits, in order to control the flow of data that the devices produce.

5.3 Implications for the design of ICT-based surveillance

I do not argue that the combined results of the included papers reflect a unified rationality of surveillance. Nor do I argue that there is a conspiratorial surveillance apparatus rising under the guise of benevolent healthcare, in order to monitor all individuals. However, the combined results of my thesis offer evidence that the production of data through ICT-based surveillance systems, and its application to health discourses, have implications that should be addressed. My original contributions are the small-scale analyses that illuminate aspects of the production and effects of ICT-based surveillance.

> "Before we can see whether some general political rationality is emerging, the task of analysis is to articulate some preliminary diagnoses at a smaller scale. Placing the evidence from such analyses in the framework of biopower, we think we can begin to identify and

analyze elements of such a domain, though it is neither stable nor homogeneous, nor does merely repeat patterns familiar from history." (Rabinow and Rose 2006, 205)

The quote above is a recommendation on the usage of the theoretical framework of biopower. In my thesis, I argue that truth discourses about health are produced through apparati where ICT is centrally positioned. This type of power is not exerted by one particular organization. Instead, it is exerted by a constellation of private industries and public entities that produce ICT, or produce data. For the purpose of surveillance, ICT engineers become more and more involved in the process of producing truths. Part of the process of producing truth-carrying discourses about health is located within the practice of software engineering. Being a software engineer in this field is therefore quite political. It is biopolitical, because we are implicated in the process of producing truth about life processes, by the production of health data and representations of individual bodies, and of populations. That implies that the processes of ICT production should be subject to a type of scrutiny similar to any other political enterprise.

The emerging field of Information Quality (Stvilia et al. 2007; Madnick et al. 2009) implements some of the processes needed to keep track of how data is produced and re-produced as it flows through various networks. But it no longer suffices to talk about accuracy and information privacy. We need to look beyond the technical problem solving and question the broader power structures that we are a part of. A way of making the process of engineering more situated is to make it more trans-disciplinary and more open to scrutiny, and to perspectives from humanities and social sciences. In our work, we have addressed this by exploring angles of investigation for future foresight activities that aim to include higher representation of different perspectives and critical, or disruptive, angles (Moore, Sanches, and Boman 2014). We have also proposed activities of critical thinking with engineers and designers of emerging integrated emergency response systems, in order to imagine alternative scenarios (Buscher, Bylund, Sanches, Ramirez & Wood 2013).

In the same line of thought, embracing a practice of situated software engineering can also create conditions for individuals to better negotiate their inclusion or exclusion in mass-surveillance, even when it relies on sources of secondary data, such as mobile phone records. This call for situatedness is not new. Dourish and Mainwaring (2012), for example, called for a shift in ubiquitous computing towards polyvocality and situatedness, while opposing grand projects of knowledge production, which are controlled by a few actors. Such a perspective could be accompanied by a commitment to be open about the methods we use and the algorithms we produce. Open source has also been described as a type of situated knowledge (Truscello 2003). Openness, not only in data (as in Open Data[9]), but also in open processes, and open algorithms, can be a first step towards a more situated, and accountable knowledge production through software. In fact, just having open data, but closed processes of knowledge production, may only contribute to a greater divide between those who can process the data, which are typically the big organizations and corporations, and those who cannot.

[9] https://okfn.org/opendata/

What positions does ICT-based health surveillance allow for individuals? In self-surveillance, it does seem that there is space for appropriation, where individuals can resist particular forms of representation. The tech-savvy can and do find alternative ways of representing themselves. Forums, like the one we observed in Paper III, provide means for this type of knowledge to be shared with others. This type of practice can be conceptualized perhaps as a type of cyborg politics (Haraway 1991). Cyborg politics is associated with biopower[10] but, as a manifesto, it breaks away from it. It is a metaphor that blurs the ontological distinction between human and non-human, and instead proposes that, in order to resist what Haraway calls the "informatics of domination", and pre-defined categorizations; the cyborg appropriates them for its own means. The metaphor of the cyborg has been associated with health surveillance, particularly the one that relies on mobile devices, called m-health:

> "A space is opened up here for researchers to identify and explore the experiences of individuals as they take up (or indeed, resist) the potentialities of mobile digital devices and the new social media as they are adopted into the 'toolbox' of health promotion. Questions that have yet to be answered include: What are the implications for subjectivities and embodiment in the world of m-health – how are the assemblages of technologies/practices/flesh enacted, re-enacted and lived? What are the political dimensions and power relations inherent in the use of m-health technologies? How will privacy (or loss of privacy) be defined and experienced in the context of these media? What are the implications for how people conduct their everyday lives and intimate relationships?" (Lupton 2012, 19)

These questions should indeed be answered. For example, Nafus and Sherman (2014) encouraged attention to be paid to individuals in the quantified self community, as their practices can be seen as a form of resistance to uniform aggregated data. They show that individuals quantify themselves to serve their own needs, mixing data and subtly subverting the logic of data aggregation. As software engineers, we can create conditions for cyborg individuals for self-expression, and allow them to participate in big data on their own terms, or to reject it entirely.

5.4 Reflection and future work

In the thesis I asked *how* health data is produced, and then I raised questions about the consequences of the circumstances of its production. I could have also asked *why* is it produced, and by *whom*. This would have taken me in different directions of research. My reason for asking *how* was that I want my results to be useful for the community of software engineers, so therefore I focused on the practice of production of data. Asking *why* could have taken me into analyses of the discourse of legitimating of ICT-based health surveillance. And the question of *who* could have taken me into analyses of the actors involved in producing that discourse.

[10] Haraway, in her Cyborg Manifesto writes that Foucault's biopolitics is "a flaccid premonition of cyborg politics".

Due to my research trajectory and learning process did not allow me to observe the process as it was ongoing, and I had to rely on the traces of research, like emails and pieces of code. If I could go back in time, my process could have involved an ethnographic study of the practice of software engineering. This could have provided greater richness in my analysis. Another possible research process, instead of doing ethnography, could have been action research: to propose workshops and activities that could intervene in the design process, and observe how these actions would influence the design process.

Another limitation is that my assumption on biopower throughout the analysis conducted here in Chapter 5 has been that the institutions that produce data about individuals are benevolent. This can be seen as a limitation in the way that it does not account for harmful usage of the data that is produced. One of the main concerns, posed by many researchers, of the increasing production of data about humans is the "function creep":

> "...where data flows from one setting to another and information gathered for one purpose finds uses in new setting and for new purposes." (Wood et al. 2006, 11)

In fact, producing data about individuals for the purposes of health, to their own benefit, can quickly go wrong if the same data is used against them, perhaps by insurance companies that would use it for increasing premiums, or by a non-benevolent government who may want to find dissidents. Additional analyses would have to be conducted that can focus more on privacy aspects. However, I would advocate for accounting for privacy in a way that does not ignore the wider power relations between individuals and institutions, and the reasons for surveillance.

Another issue related to privacy is that in this thesis I did not pay attention to human communication mediated by technology. In fact, health data may be generated by individuals, or groups, and communicated to other individuals or groups. That occurs in communities of self-tracking, such as the Fitbit community. Studies that would pay attention to aspects of interpersonal privacy, in particular, privacy boundary regulations (Altman 1975) would be needed to illuminate these issues.

5.5 Conclusion

Through an analysis of the practice of software engineering, I have shown that health data is an object produced through an articulation of discursive and material aspects. Designing ICT systems to fulfill the purpose of surveillance is to articulate these aspects in a particular way. As an element of biopower, ICT-based surveillance plays a role in producing knowledge about individuals and populations. Star (1990) called for analyses of power and technologies, of the work of creating standards and how standardized representations can be problematic for some. In this thesis, I presented descriptions of the work of producing health data, by analyzing how ICT is produced, and how it is used. My descriptions of ICT within the frame of health surveillance sought to problematize certain aspects of its production.

Through descriptions, and subsequent reflections, of two main case studies and two related analyses, I showed that the work of producing health data is neither abstract, nor linear: there is no universal way of doing it. I showed that the way that discursive and material aspects are articulated can contribute to potentially problematic renderings of parts of the population as invisible. Moreover, I argued that to achieve fair and less myopic health data systems, one must pay attention to all the social codes, concepts, technologies, and people involved in the design process. One must also consider how the data can be interpreted once the system is assembled and put to use, and what the consequences are for the data produced.

The current practice of software engineering does not seem to be taking these into account. I proposed a situated practice of ICT development that exposes the rationalities inherent to the practice, in order to allow individuals and others implicated by surveillance to easier discern the subject positions created for them. Situating data production entails both acknowledging the situatedness of engineering practice, and being aware that there is a relationship between knowledge and power, namely that producing data is an exercise in power. The enterprise of producing health data requires accountability from the ones involved in the process.

References

Albrechtslund, Anders, and Lynsey Dubbeld. 2002. “The Plays and Arts of Surveillance: Studying Surveillance as Entertainment.” *Surveillance & Society* 3 (2/3).

Altman, Irwin. 1975. *The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding*. Brooks/Cole Pub. Co. http://books.google.se/books?id=GLBPAAAAMAAJ.

Anema, Aranka, Sheryl Kluberg, Kumanan Wilson, Robert S Hogg, Kamran Khan, Simon I Hay, Andrew J Tatem, and John S Brownstein. 2014. “Digital Surveillance for Enhanced Detection and Response to Outbreaks.” *The Lancet. Infectious Diseases* 14 (11): 1035–37. doi:10.1016/S1473-3099(14)70953-3.

Bakker, Jorn, Mykola Pechenizkiy, and Natalia Sidorova. 2011. “What’s Your Current Stress Level? Detection of Stress Patterns from GSR Sensor Data.” In *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*, 573–80.

Ball, Kirstie. 2009. “Exposure: Exploring the Subject of Surveillance Information.” *Information, Communication and Society*. http://oro.open.ac.uk/12233/.

Barad, Karen. 2003. “Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter.” *Signs: Journal of Women in Culture and Society*. doi:10.1086/345321.

Barkhuus, L. 2012. “The Mismeasurement of Privacy: Using Contextual Integrity to Reconsider Privacy in HCI.” In , 367–76. http://dl.acm.org/citation.cfm?id=2207727.

Barkhuus, Louise. 2006. “Designing Ubiquitous Computing Technologies to Motivate Fitness and Health.” In *Grace Hopper Celebration of Women in Computing*. Vol. 6.

Becker, Richard, Chris Volinsky, Ramón Cáceres, Karrie Hanson, Sibren Isaacman, Ji Meng Loh, Margaret Martonosi, James Rowland, Simon Urbanek, and Alexander Varshavsky. 2013. “Human Mobility Characterization from Cellular Network Data.” *Communications of the ACM* 56 (1): 74. doi:10.1145/2398356.2398375.

Bennett, Colin J. 2011. “In Defense of Privacy: The Concept and the Regime.” *Surveillance & Society* 8 (4): 485–96.

Bioethics, Center for Research Ethics and. 2014. “CODEX - Rules and Guidelines for Research.” http://www.codex.uu.se/en/forskninghumsam.shtml.

Biosync. 2014. “BIOSYNC Technology - BRINGING LIFE-CHANGING IMPACT THROUGH WEARABLE BIO-SENSORS.” http://signup.biosynctechnology.com/.

Boman, Magnus. 2014. *Active Healthy Ageing. White Paper, TR2014-004*.

Boman, Magnus, and Einar Holm. 2004. “Multi-Agent Systems , Time Geography , and Microsimulations,” 95–118.

Bowker, Geoffrey C, and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: The {MIT} Press.

boyd, danah, and Kate Crawford. 2012. “Critical Questions for Big Data.” *Information, Communication & Society* 15 (5). Routledge: 662–79. doi:10.1080/1369118X.2012.678878.

Buscher, Monika, Markus Bylund, Pedro Sanches, Leonardo Ramirez, and Lisa Wood. 2013. “A New Manhattan Project? Interoperability and Ethics in Emergency Response Systems of Systems.” In *10th International Conference on Information Systems for Crisis Response and Management*. Baden, Germany.

Bylund, Markus, Mikael Johnson, Asko Lehmuskallio, Peter Seipel, and Sakari Tamminen. 2010. *PRIMA—Privacy Research through the Perspective of a Multidisciplinary Mash up*.

Cantarell, José Félix Guerrero, and Pedro Sanches. 2013. *D6.4 Evaluation of Privacy, Trust, and Safety Requirements*. Stockholm. https://onedrive.live.com/view.aspx?resid=E9323F309EF8909C!224&ithint=file,.pdf&app=WordPdf&authkey=!ACKCbaNH-4oY_rk.

Castells, Manuel. 2011. “The Rise of the Network Society: The Information Age: Economy, Society, and Culture.” In , 1:500–509. John Wiley & Sons.

Cavoukian, A. 2012. “Privacy by Design.” *Report of the Information & Privacy Commissioner Ontario, Canada*. http://privacybydesign.ca/content/uploads/2012/04/Privacy-by-Design-Origins-Meaning-and-Prospects.pdf.

Chaudhry, Basit, Jerome Wang, Shinyi Wu, Margaret Maglione, Walter Mojica, Elizabeth Roth, Sally C Morton, and Paul G Shekelle. 2006. “Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care.” *Annals of Internal Medicine* 144 (10). Am Coll Physicians: 742–52.

Coll, Sami. 2014. “Power, Knowledge, and the Subjects of Privacy: Understanding Privacy as the Ally of Surveillance.” *Information, Communication & Society*, no. ahead-of-print. Taylor & Francis: 1–14.

Commission, Oxford Martin Commission. 2013. “Now for the Long Term.” Oxford, Report of the Oxford Martin Commission for Future Generations.

Cross, Nigel. 2001. “Designerly Ways of Knowing: Design Discipline Versus Design Science.” *Design Issues* 17 (3): 49–55. doi:10.1162/074793601750357196.

Data, Committee on the Analysis of Massive Data, Committee on Applied and Theoretical Statistics Statistics, Board on Mathematical Sciences and Their Applications, Division on Engineering and Physical Sciences Sciences, and National Research Council Council. 2013. *Frontiers in Massive Data Analysis*. National Academies Press. https://books.google.com/books?id=lQB1AgAAQBAJ&pgis=1.

De Figueiredo, António Dias, and Paulo Rupino da Cunha. 2007. “Action Research and Design in Information Systems.” In *Information Systems Action Research*, 61–96. Springer.

De Santos Sierra, Alberto, Carmen Sánchez Ávila, Javier Guerra Casanova, and Gonzalo Bailador del Pozo. 2011. “A Stress-Detection System Based on Physiological Signals and Fuzzy Logic.” *Industrial Electronics, IEEE Transactions on* 58 (10). IEEE: 4857–65.

De Smit, B, H W Lenstra, D Dunham, and R Sarhangi. 2003. “Artful Mathematics: The Heritage of MC Escher.” *NOTICES-AMERICAN MATHEMATICAL SOCIETY* 50 (4). AMERICAN MATHEMATICAL SOCIETY: 446–57.

Deleuze, Gilles, and Félix Guattari. 1987. “Introduction: Rhizome.” *A Thousand Plateaus: Capitalism and Schizophrenia*, 3–25.

Diederich, Joachim, and Insu Song. 2014. “Mental Health Informatics: Current Approaches.” In *Mental Health Informatics*, 1–16. Springer.

Dourish, Paul, and Scott D Mainwaring. 2012. “Ubicomp’s Colonial Impulse.” In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 133–42.

Dubbeld, Lynsey. 2003. “Observing Bodies. Camera Surveillance and the Significance of the Body.” *Ethics and Inf. Technol.* 5 (3): 151–62. http://portal.acm.org/citation.cfm?id=957962.

Edwards, Paul. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Mit Press.

Edwards, Paul, Matthew S Mayernik, Archer Batcheller, Geoffrey Bowker, and Christine Borgman. 2011. “Science Friction: Data, Metadata, and Collaboration.” *Social Studies of Science*. Sage Publications, 0306312711413314.

Espeland, Wendy Nelson, and Mitchell L. Stevens. 2009. “A Sociology of Quantification.” *European Journal of Sociology* 49 (03). Cambridge University Press: 401. doi:10.1017/S0003975609000150.

Eysenbach, Gunther. 2006. “Infodemiology: Tracking Flu-Related Searches on the Web for Syndromic Surveillance.” In *AMIA Annual Symposium Proceedings*, 2006:244.

Feast, Luke, and Gavin Melles. 2010. “Epistemological Positions in Design Research : A Brief Review of the Literature,” no. July: 1–5.

Felt, Ulrike, Lisa Gugglberger, and Astrid Mager. 2009. “Shaping the Future E-Patient : The Citizen-Patient in Public Discourse on” 22 (1): 24–43.

Ferreira, Pedro, Pedro Sanches, Kristina Höök, and Tove Jaensson. 2008. “License to Chill!: How to Empower Users to Cope with Stress.” In *Nordic Forum for Human-Computer Interaction Research (NordiCHI)*, 123–32. Lund, Sweden: ACM. doi:10.1145/1463160.1463174.

Fessl, Angela, Verónica Rivera-Pelayo, Lars Müller, Viktoria Pammer, and Stefanie Lindstaedt. 2011. “Motivation and User Acceptance of Using Physiological Data to Support Individual Reflection.” In *2nd MATEL Workshop at European Conference for Technology Enhanced Learning (ECTEL 2011)*.

Fitbit. 2013. “Fitbit.” https://www.fitbit.com/uk.

Floridi, Luciano. 2012. “Big Data and Their Epistemological Challenge.” *Philosophy & Technology*. Springer, 1–3.

Foucault, Michel. 1973. “The Birth of the Clinic.” *London: Tavistock*.

Foucault, Michel. 1978. “The History of Sexuality, Volume 1: An Introduction (R. Hurley, Trans.).” *New York: Vintage*.

Foucault, Michel. 1980. *Power/knowledge: Selected Interviews and Other Writings, 1972-1977*. Random House LLC.

Foucault, Michel. 1984. *The Foucault Reader*. Edited by Paul Rabinow. Random House LLC.

Foucault, Michel. 1995. *Discipline & Punish: The Birth of the Prison*. 2nd Vintag. New York: Vintage.

Foucault, Michel. 1998. *The Will to Knowledge: The History of Sexuality (volume 1)*. London: Penguin Books.

Frankel, Mark S, and Sanyin Siang. 1999. “Ethical and Legal Aspects of Human Subjects Research on the Internet.” *Published by AAAS Online*.

French, Martin. 2009. “Picturing Public Health Surveillance: Tracing the Material Dimensions of Information in Ontario’s Public Health System.” Canadian theses.

French, Martin A. 2009. “Woven of War-Time Fabrics: The Globalization of Public Health Surveillance.” *Surveillance & Society* 6 (2): 101–15.

French, Martin, and Gavin Smith. 2013. “‘Health’surveillance: New Modes of Monitoring Bodies, Populations, and Polities.” *Critical Public Health* 23 (4). Taylor & Francis: 383–92.

Friedman, Batya, and Helen Nissenbaum. 1996. “Bias in Computer Systems.” *ACM Transactions on Information Systems (TOIS)* 14 (3). ACM: 330–47.

GeoMedia. 2014. “Become a Master Designer: Rule Four: Spacing Is Your Friend.” http://s3.gomedia.us/wp-content/uploads/2008/01/spacing_sample_grids.jpg.

Giddens, Anthony. 1987. *The Nation-State and Violence: Volume 2 of A Contemporary Critique of Historical Materialism*. University of California Press.

Gieryn, Thomas F. 1983. “Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists.” *American Sociological Review*. JSTOR, 781–95.

Gilliom, John. 2011. “A Response to Bennett’s ‘In Defence of Privacy.’” *Surveillance & Society* 8 (4): 500–504. http://www.surveillance-and-society.org/ojs/index.php/journal/article/view/363.

Grandin, Karl, Nina Wormbs, and Sven Widmalm. 2004. *The Science-Industry Nexus: History, Policy, Implications*. New York, NY, USA: Science History Publications.

Haggerty, Kevin D, and Richard V Ericson. 2000. “The Surveillant Assemblage.” *The British Journal of Sociology* 51 (4). Wiley Online Library: 605–22.

Haraway, Donna. 1988. “Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective.” *Feminist Studies*. JSTOR, 575–99.

Harding, Sandra. 2002. “Rethinking Standpoint Epistemology: What Is―strong Objectivity‖.” *Knowledge and Inquiry: Readings in Epistemology*, 352–84.

Harraway, Donna. 1991. “A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century.” *Simians, Cyborgs and Women: The Reinvention of Nature*, 149–82.

Heisenberg, Werner. 1958. *Physics and Philosophy: The Revolution in Modern Science*. Prometheus Books.

Heller, Kevin Jon. 1996. “Power, Subjectification and Resistance in Foucault.” *SubStance*. JSTOR, 78–110.

Hood, Leroy, and Charles Auffray. 2013. “Participatory Medicine: A Driving Force for Revolutionizing Healthcare.” *Genome Med.* 5: 110–14.

Hsieh, Hsiu-Fang, and Sarah E Shannon. 2005. “Three Approaches to Qualitative Content Analysis.” *Qualitative Health Research* 15 (9). Sage Publications: 1277–88.

IEEE Std 830-1998. 1998. “IEEE Recommended Practice for Software Requirements Specifications,” 1–40. doi:10.1109/IEEESTD.1998.88286.

Im, Eun-Ok, and Wonshik Chee. 2006. “An Online Forum as a Qualitative Research Method: Practical Issues.” *Nursing Research* 55 (4). NIH Public Access: 267.

Jørgensen, Marianne W, and Louise J Phillips. 2002. *Discourse Analysis as Theory and Method*. Sage.

Kozinets, Robert V. 2001. “The Field Behind the Screen: Using Netnography For Marketing Research in Online Communities.”

Kranz, Matthias, Andreas Möller, Nils Hammerla, Stefan Diewald, Thomas Plötz, Patrick Olivier, and Luis Roalter. 2013. “The Mobile Fitness Coach: Towards Individualized Skill Assessment Using Personalized Mobile Devices.” *Pervasive and Mobile Computing* 9 (2). Elsevier: 203–15.

Lane, Nicholas D, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T Campbell. 2010. “A Survey of Mobile Phone Sensing.” *Communications Magazine, IEEE* 48 (9). IEEE: 140–50.

Langer, Roy, and Suzanne C Beckman. 2005. “Sensitive Research Topics: Netnography Revisited.” *Qualitative Market Research: An International Journal* 8 (2). Emerald Group Publishing Limited: 189–203.

Latour, Bruno. 1978. *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press. http://www.amazon.com/Science-Action-Scientists-Engineers-Through/dp/0674792912.

Latour, Bruno. 1986. “Visualization and Cognition: Drawing Things Together.” *Knowledge and Society* 6: 1–40.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Harvard university press.

Latour, Bruno, and Steve Woolgar. 1979. "Laboratory Life: The Social Construction Ofscientific Facts." *Beverly Hills: Sage*.

Latour, Bruno, and Steve Woolgar. 2013. *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press.

Law, John. 1987. "Technology and Heterogeneous Engineering: The Case of Portuguese Expansion." *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* 1. MIT Press Cambridge, MA: 1–134.

Longino, Helen E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press.

Lupton, Deborah. 2012. "M-Health and Health Promotion: The Digital Cyborg and Surveillance Society." *Social Theory & Health* 10 (3): 229–44.

Lupton, Deborah. 2014. "Apps as Artefacts: Towards a Critical Perspective on Mobile Health and Medical Apps." *Societies* 4 (4). Multidisciplinary Digital Publishing Institute: 606–22. doi:10.3390/soc4040606.

Lyon, David. 2002. "Surveillance as Social Sorting." *Surveillance as Social Sorting: Privacy, Risk and Automated Discrimination*. Routledge, 13.

Lyon, David. 2007. *Surveillance Studies: An Overview*. 1st ed. Polity Press.

Lyotard, Jean-Francois. 1984. *The Postmodern Condition: A Report on Knowledge (Theory and History of Literature, Volume 10)*. University Of Minnesota Press. http://www.amazon.com/The-Postmodern-Condition-Knowledge-Literature/dp/0816611734.

MacKenzie, Donald. 1987. "Missile Accuracy: A Case Study in the Social Processes of Technological Change." *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. MIT Press, 195–222.

Madnick, Stuart E, Richard Y Wang, Yang W Lee, and Hongwei Zhu. 2009. "Overview and Framework for Data and Information Quality Research." *Journal of Data and Information Quality (JDIQ)* 1 (1). ACM: 2.

Mandl, Kenneth D, J Marc Overhage, Michael M Wagner, William B Lober, Paola Sebastiani, Farzad Mostashari, Julie A Pavlin, et al. 2004. "Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience." *Journal of the American Medical Informatics Association* 11 (2): 141–50.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. 2011. "Big Data: The next Frontier for Innovation, Competition, and Productivity." *McKinsey Global Institute*, 1–137.

Mayntz, Renate, and Thomas Parke Hughes. 1988. *The Development of Large Technical Systems*. Campus Verlag Frankfurt am Main.

Moore, Heather, Pedro Sanches, and Magnus Boman. 2014. “Communities of Vision, Practice, and Foresight.” In *5th International Conference on Future - Oriented Technology Analysis (FTA) - Engage Today to Shape Tomorrow*. Brussels.

Nafus, Dawn, and Jamie Sherman. 2014. “Big Data, Big Questions| This One Does Not Go Up To 11: The Quantified Self Movement as an Alternative Big Data Practice.” *International Journal of Communication* 8: 11.

Nanni, M., R. Trasarti, C. Renso, F. Giannotti, and D. Pedreschi. 2010. “Advanced Knowledge Discovery on Movement Data with the GeoPKDD System.” In , 693–96.

Nissenbaum, Helen. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford Law Books.

Norris, Clive, Mike McCahill, and David Wood. 2002. “The Growth of CCTV: A Global Perspective on the International Diffusion of Video Surveillance in Publicly Accessible Space.” *Surveillance & Society*. http://library.queensu.ca/ojs/index.php/surveillance-and-society/article/view/3369.

Orlikowski, Wanda Janina. 1995. “Action and Artifact: The Structuring of Technologies-in-Use.” Sloan School of Management, Massachusetts Institute of Technology].

Palen, L., and P. Dourish. 2003. “Unpacking Privacy for a Networked World.” In , 129–36. http://dl.acm.org/citation.cfm?id=642635.

Pantzar, Mika, and Minna Ruckenstein. 2014. “The Heart of Everyday Analytics: Emotional, Material and Practical Extensions in Self-Tracking Market.” *Consumption Markets & Culture*, no. ahead-of-print. Taylor & Francis: 1–18.

Phyode. 2013. “PHYODE | Home.” http://www.phyode.com/.

Pinch, Trevor J, and Wiebe E Bijker. 1984. “The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other.” *Social Studies of Science*. JSTOR, 399–441.

Pinch, Trevor J, and Wiebe E Bijker. 2012. “The Social Construction of Facts and Artefacts : Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other The Social Construction of Facts and Artefacts : Or How the Sociology of Science and the Sociology of Technology Migh” 14 (3): 399–441.

Polar. 2011. “User Manual for Polar RS300X in English | Polar USA.” http://www.polar.com/us-en/support/User_manual_for_Polar_RS300X_in_English.

Quantified Self Labs. 2013. “Self Knowledge Through Numbers.” http://quantifiedself.com.

Rabinow, Paul, and Nikolas Rose. 2006. “Biopower Today.” *BioSocieties* 1 (2): 195–217.

Regan, Priscilla M. 2011. “Debate Response to Bennett : Also in Defence of Privacy” 8 (4): 497–99.

Rich, Emma, and John Evans. 2005. "'Fat Ethics'--the Obesity Discourse and Body Politics." *Social Theory & Health* 3 (4). Palgrave Macmillan: 341–58.

Rittel, Horst W J, and Melvin M Webber. 1973. "Dilemmas in a General Theory of Planning *" 4 (December 1969): 155–69.

Rose, Nikolas. 2007. *The Politics of Life Itself: Biomedicine, Power, and Subjectivity in the Twenty-First Century*. Princeton University Press.

Rudin, Robert S, Aneesa Motala, Caroline L Goldzweig, and Paul G Shekelle. 2014. "Usage and Effect of Health Information Exchange: A Systematic Review." *Annals of Internal Medicine* 161 (11). Am Coll Physicians: 803–11.

Sanches, Pedro. 2008. "Supporting Self-Reflection in Everyday Life." *Master's Thesis, Swedish Inst. of Comp. Sci*.

Savel, Thomas G, Seth Foldy, and others. 2012. "The Role of Public Health Informatics in Enhancing Public Health Surveillance." *MMWR Surveill Summ* 61: 20–24.

Shapin, Steven, Simon Schaffer, and Thomas Hobbes. 1985. *Leviathan and the Air-Pump*. Br Soc Philosophy Sci.

Sharma, Nandita, and Tom Gedeon. 2012. "Objective Measures, Sensors and Computational Techniques for Stress Recognition and Classification: A Survey." *Computer Methods and Programs in Biomedicine* 108 (3). Elsevier: 1287–1301.

Silvăşan, Iuliana, Per Kreuger, Pedro Sanches, Elsa Vaara, and Marie Sjölinder. 2010. "Movement Identification in Affective Health--a Mobile Biofeedback Monitoring System." In *18th Telecommunications Forum TELFOR*, 340–43. Belgrade, Serbia.

Smith, Merritt Roe, and Leo Marx. 1994. *Does Technology Drive History?: The Dilemma of Technological Determinism*. MIT Press.

Star, Susan Leigh. 1990. "Power, Technology and the Phenomenology of Conventions: On Being Allergic to Onions." *The Sociological Review* 38 (S1). Wiley Online Library: 26–56.

Steeves, V. 2008. "Reclaiming the Social Value of Privacy." *Lessons from the Identity Trail*.

Stvilia, Besiki, Les Gasser, Michael B Twidale, and Linda C Smith. 2007. "A Framework for Information Quality Assessment." *Journal of the American Society for Information Science and Technology* 58 (12). Wiley Online Library: 1720–33.

Ståhl, Anna, Kristina Höök, Martin Svensson, Alex S. Taylor, and Marco Combetto. 2009. "Experiencing the Affective Diary." *Personal Ubiquitous Comput.* 13 (5): 365–78. http://portal.acm.org/citation.cfm?id=1536073.

Swan, Melanie. 2009. "Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking." *International Journal of Environmental Research and Public Health* 6 (2). Molecular Diversity Preservation International: 492–525.

Tech in Asia. 2011. “How Docomo Plans To Use Mobile Data For City Planning & Earthquake Preparation.” http://www.techinasia.com/docomo-japan-mobile-spatial/.

Truscello, Michael. 2003. “The Architecture of Information: Open Source Software and Tactical Poststructuralist Anarchism.” *Postmodern Culture* 13 (3). The Johns Hopkins University Press.

Van Brakel, Rosamunde, and Paul De Hert. 2011. “Policing, Surveillance and Law in a Pre-Crime Society: Understanding the Consequences of Technology Based Strategies.” *Journal of Police Studies* 20 (20): 163–92. http://www.vub.ac.be/LSTS/pub/Dehert/378.pdf.

Vaz, Paulo, and Fernanda Bruno. 2002. “Paulo Vaz and Fernanda Bruno—Types of Self-Surveillance: From Abnormality to Individuals ‘at Risk’.” *Surveillance & Society* 1 (3): 272–91.

Webster, Andrew. 2002. “Innovative Health Technologies and the Social: Redefining Health, Medicine and the Body.” *Current Sociology* 50 (3). Sage Publications: 443–57.

Whiteman, Natasha. 2012. *Undoing Ethics*. Boston, MA: Springer US. doi:10.1007/978-1-4614-1827-6.

Whitson, Jennifer R. 2013. “Gaming the Quantified Self” 11: 163–76.

Vidyarthi, Jay, and Bernhard E Riecke. 2013. “Mediated Meditation: Cultivating Mindfulness with Sonic Cradle.” In *CHI’13 Extended Abstracts on Human Factors in Computing Systems*, 2305–14.

Williamson, Ben. 2014. “Algorithmic Skin: Health-Tracking Technologies, Personal Analytics and the Biopedagogies of Digitized Health and Physical Education.” *Sport, Education and Society* 20 (1). Routledge: 133–51. doi:10.1080/13573322.2014.962494.

Wood, David Murakami, Kirstie Ball, David Lyon, Clive Norris, and Charles Raab. 2006. “A Report on the Surveillance Society.” *Surveillance Studies Network, UK*.

Woolgar, Steve, and Geoff Cooper. 2011. “Do Artefacts Have Ambivalence ? Moses ’ Bridges , Winner ’ S Bridges and Other Urban Legends in S & TS” 29 (3): 433–49.

Yapa, Lakshman. 1996. “What Causes Poverty?: A Postmodern View.” *Annals of the Association of American Geographers* 86 (4). Taylor & Francis: 707–28.

Zimmerman, John, Jodi Forlizzi, and Shelley Evenson. 2007. “Research through Design as a Method for Interaction Design Research in HCI.” In , 493–502. San Jose, California, USA: ACM. doi:10.1145/1240624.1240704.

TRITA-ICT/ECS AVH 14:17
ISSN 1653-6363
ISRN KTH/ICT/ECS/AVH-14/17-SE
ISBN: 978-91-7595-403-5