Surveillance & Society

Article | 'Depends on Who's Got the Data': Public Understandings of Personal Digital Dataveillance

Deborah Lupton

News & Media Research Centre
University of Canberra, Australia.
deborah.lupton@canberra.edu.au

Mike Michael

Department of Sociology, Philosophy and Anthropology
University of Exeter, UK.
M.Michael@exeter.ac.uk

Abstract

Post-Snowden, several highly-publicised events and scandals have drawn attention to the use of people's personal data by other actors and agencies, both legally and illicitly. In this article, we report the findings of a project in which we used cultural probes to generate discussion about personal digital dataveillance. Our findings suggest the prevailing dominance of tacit assumptions about the uses and benefits of dataveillance as well as fears and anxieties about its possible misuse. Participants were able to identify a range of ways in which dataveillance is conducted, but were more aware of obvious commercial and some government actors. There was very little identification of the types of dataveillance that are used by national security and policing agencies or of illegal access by hackers and cybercriminals. We found that the participants recognised the value of both personal data and the big aggregated data sets that their own data may be part of, particularly for their own convenience. However, they expressed concern or suspicion about how these data may be used by others, often founded on a lack of knowledge about what happens to their data. The major question for our participants was where the line should be drawn. When does personal dataveillance become too intrusive, scary or creepy? What are its drawbacks and risks? Our findings suggest that experimenting with innovative approaches to elicit practices and understandings of personal digital data offers further possibilities for greater depth and breadth of social research with all types of social groups.

Introduction

In the contemporary global digital data economy, personal information about individuals is constantly generated by their use of digital technologies and their movements around sensor-equipped physical environments. Digital data are beginning to influence people's concepts of themselves, their bodies and their social relations. As yet, little research has been published that identifies what members of the public are making of the ways in which the personal details generated about them by digital technologies are collected and used by other actors. In this article, we have two aims: 1) to introduce some innovative methods to conduct empirical research in critical data studies; and 2) to report findings from a study using these methods to generate discussion about personal digital dataveillance.

In what follows, we begin by providing an overview of the literature on dataveillance and public understanding of personal data issues. We then outline the methods and findings of our research project, involving conducting six focus groups in Sydney in which the participants were asked to work as a group to respond to some novel tasks that we had devised using the cultural probes approach.

Lupton, Deborah, and Mike Michael. 2017. 'Depends on Who's Got the Data': Public Understandings of Personal Digital Dataveillance. *Surveillance & Society* 15(2): 254-268.
http://library.queensu.ca/ojs/index.php/surveillance-and-society/index | ISSN: 1477-7487

Dataveillance and Personal Data

Our conceptual perspective starts with the idea that when people interact with digital technologies, digital data assemblages are configured. These assemblages are heterogeneous entanglements of humans, technologies and data that are constantly changing as users' new encounters with digital technologies occur and as different data sets come together and interact and are taken up for a range of purposes (Lupton 2016a). A diverse array of actors and agencies use this information. Sometimes people make personal use of their data assemblages. More often, however, they are employed for commercial, research, governmental or managerial purposes by other actors and agencies, including hackers and cybercriminals (Andrejevic 2013; van Dijck 2013; Fuchs 2011; Zuboff 2015; van Dijck 2014).

The use of personal digital data in surveillance activities ('dataveillance') (van Dijck 2014; Esposti 2014) has become a controversial topic. Dataveillance is undertaken at both the personal or interpersonal level, involving self-surveillance or surveillance of others in one's life, and the organisational level, conducted by agencies such as commercial enterprises, security and policing agencies, transport organisations, workplaces, schools and so on. It is important to note that dataveillance may be voluntary and self-imposed. This kind of dataveillance occurs when people use digital devices to monitor aspects of their lives as part of self-surveillance efforts (Klauser and Albrechtslund 2014; Lupton 2016b, 2016c) or to conduct 'intimate surveillance' of loved ones (Levy 2015; Albrechtslund and Lauritsen 2013) and when they participate in 'social surveillance', or the consensual watching of each other on social media sites (Marwick 2012; Trottier and Lyon 2012; Albrechtslund and Lauritsen 2013). The use of digital devices such as smartphones to record and watch others in public spaces has been dubbed 'sousveillance' (literally, watching from below) (Mann and Ferenbok 2013). This type of dataveillance is often voluntarily taken up by people to challenge powerful authorities and draw attention to their wrongdoing.

Alternatively, dataveillance may be conducted on people, either overtly or covertly, by other actors who access the digital traces left by people's online activities or their movements in public spaces as monitored by CCTV cameras or digital sensors (Klauser and Albrechtslund 2014; Albrechtslund and Lauritsen 2013). These actors include policing, security and other government agencies, educational institutions, employers, internet corporations, and advertising and insurance companies. Given the dispersed and constant nature of digital dataveillance, it has become impossible to avoid becoming a subject of this type of monitoring (Lyon and Bauman 2013). The people who are being watched often have little power over or even knowledge of who is conducting dataveillance and how their personal data are used. This type of dataveillance conforms to the classic model of surveillance, or watching from above.

The third-party use of people's personal digital data is beginning to have significant implications for their life chances. Predictive algorithms that draw on personal digital data are used now in many social and economic domains to construct scores that are used to determine whether individuals should be provided with access to special offers, goods and services, or whether they pose risks such as the possibility of engaging in criminal acts and terrorism. Concerns have been consequently raised by privacy and ethics organisations and legal scholars about invasions of personal privacy, and the implications for social justice and civil rights (Crawford and Schultz 2014; Polonetsky and Tene 2013; Nuffield Council on Bioethics 2015). Critics have pointed out that established social and economic disadvantage and marginalisation can be exacerbated by government and corporate dataveillance strategies (Crawford and Schultz 2014; Gangadharan 2015; Rosenblat et al. 2014). Other commentators have discussed the commercialisation of personal digital data and critiqued the ways in which this information is used for the financial benefit of others. They contend that a new 'digital divide' is emerging, in which powerful institutions and organisations such as the internet empires (the likes of Facebook, Google, Apple, Microsoft and Amazon) have control over people's digital data while others are excluded from access (Andrejevic and Burdon 2015; Andrejevic 2014; boyd and Crawford 2012; Zuboff 2015).

The public has been exposed to a number of events and scandals in the past few years which have drawn attention to dataveillance. In mid-2013, whistle-blower Edward Snowden's revelations about national security agencies' digital surveillance of their citizens were first publicised, and his leaked documents and pronouncements have received a high level of media attention. A continuing number of data breach scandals and hacking events have also emerged for public discussion. For example, news reporting of the Facebook and OKCupid user manipulation experiments as well as the hacking of nude celebrity photos on iCloud, personal details of Sony Entertainment employees and highly intimate information about users' sexual preferences and practices on adult dating platforms, have all publicised the ways in which people's personal data may be used by other actors both legally and illicitly.

In response, there is evidence of a growing unease among members of the public about how their personal data are generated and used by other actors and agencies. Members of the public recognise the value of big data for such public goods as maintaining national security, controlling crime, promoting public health, improving healthcare and so on. However, many people realise that their personal data have become commercially valuable and are hostile to the idea that government agencies should sell big data for profit rather than use it for the public good, as evidenced by research conducted the Wellcome Trust in the UK (Wellcome Trust 2013). The public outcry created by the British government's attempt to capitalise on data from the National Health Service (care.data) is another example (Public Administration Select Committee 2014).

Two Pew Research Center reports outlining the findings of surveys about Americans' attitudes to data privacy (Madden 2014; Madden and Rainie 2015) found that most respondents felt that their online privacy was challenged. They were largely aware of national security agencies' dataveillance of citizens and the use of their personal information by commercial companies. The second report (Madden and Rainie 2015) identified a significant element of personal data insecurity that had begun to affect people's attitudes towards dataveillance and data privacy. Very few respondents felt they had much control over the types of data collected on them and how these data are used. They expressed strong views about the importance of preserving personal data privacy and security. Smaller-scale academic research provides some further insights into how the public is responding to digital dataveillance. In a project involving a survey, interviews and focus groups conducted pre-Snowden, Andrejevic (2014) found that his Australian respondents did express concern about privacy issues. His survey revealed that there was strong support for do-not-track legislation, for example. However, only a little over half of the participants in his survey agreed that websites collected too much information about users. Andrejevic also found a strong sense of powerlessness among respondents about the extent to which they could protect their personal data from being collected and used by other parties, and a lack of knowledge about how to do this.

Using Cultural Probes for Critical Data Studies

In our project, we set out to experiment with some innovative approaches to eliciting people's understandings of digital data—particularly those data that are generated by or about themselves. We were not just interested in people's understandings of social media mining but their attitudes to personal digital dataveillance across the diverse domains in which it takes place and in the wake of recent scandals and revelations about dataveillance of the public. Rather than focusing specifically on data privacy and security issues, we were interested in how people conceptualised the generation of personal digital data and the routes and circuits through which their data pass, as well as their possible endpoints (Michael and Lupton 2016).

As a first attempt to use novel methods to explore these issues, we conducted six focus groups with Sydney residents, each with eight members (a total of 48 participants). We wanted to explore ways to invite people to think and work together creatively and, therefore, decided to employ cultural probes to stimulate thought, discussion and debate, involving asking people to work together as a group or in pairs to generate material

(Michael and Lupton 2016). Cultural probes are objects or tasks that are designed to be playful and provocative so as to encourage people to think in new ways. They are particularly valued for their ability to address intimate or controversial issues, to act as 'irritants' to engage people's responses. Cultural probes have frequently been used in speculative design (Gaver et al. 2004; Boehner, Gaver, and Boucher 2012). They have since moved from this original context into human-computer interaction research, particularly studies directed at developing and testing prototypes (Boehner et al. 2007). The standard approach for using cultural probes involves providing research participants with objects that invite them to perform specific tasks. These may include disposable cameras to take images, digital memo recorders, maps, diaries to jot down notes or drawings and custom-made postcards with questions on the back for participants to respond to in writing. Participants are typically requested to take these items home and to engage with them over a period (often several days or weeks). Sometimes these probes are left in domestic settings for couples or members of families to complete (Boehner, Gaver, and Boucher 2012). Probes have also been used in focus groups and workshop settings (for example, Vetere et al. 2005).

The emphasis of the use of cultural probes in these contexts is on generating inspiration for design ideas rather than research information. However, the original creators of the cultural probes approach have begun to outline how probes can be included in existing research tools such as questionnaires and self-documentation tools to provide a more playful dimension, or used in intervention research to invite participants to think about their beliefs and practices more explicitly or in unfamiliar ways (Boehner, Gaver, and Boucher 2012). We think that the use of probes in research on public understanding of personal data has something to offer regarding new ways of inviting people to think about the impact of digital devices, the gathering of personal data that they facilitate and the politics of these data. Rather than asking people directly or in abstract terms about personal dataveillance and data privacy issues, this approach encourages directions of thought and discussion that go beyond obvious responses by locating the issues in the context of everyday lives and practices.

In the focus groups we ran, the participants were asked to engage in collaborative activities or tasks, and then to discuss their experiences of these tasks within the group. We developed a set of original tasks to do so, as follows:

1. *The Daily Big Data Task*. This task asked participants to work together to draw a timeline of a typical person's day and to add the ways in which data (digital or otherwise) may be collected on that person. We provided some initial ideas to start off the task. Participants were asked to write on a large sheet of butcher's paper in constructing the timeline, and also to add post-it notes with further details if they wished. The rationale behind this task is to derive material on how people perceive the sites, occasions and activities through which data about them is gathered, but also how they situate digital data within the wider data ecology. They were called on to reflect not only on what they knew about how digital data may be gathered about people by themselves or others but also what they did not know: the gaps, absences or ambiguities in their knowledge. Figure 1 provides an example of one of the timelines made in this task.

2. *The Digital Profile Card Game*. This involved a task designed as a 'game'. Small groups of participants together chose a card from each of four stacks. Each stack contained a particular class of information about individuals: demographic, employment, cultural consumption, and health. There was also a stack of blank cards which participants were asked to draw on and fill out with other details as they chose (for example, financial status, travel habits, criminal record). We asked participants to discuss how this profile might have been digitally derived, the inferences that could be drawn about the person from the profile and the uses to which such a profile can be put. The rationale here is to enable participants to explore how particular databases might be combined, how they are generative of

cases/caricatures, and the problematic issues that might arise from these. Figure 2 shows an example of the set of cards with which one group worked.

3. *The Personal Data Machine*. In this activity, participants were asked to work in pairs to design two data-gathering devices. This task came in two parts. In the first, participants took five minutes to draw or describe a design to collect digital data on themselves and to reflect on such questions as: What would it do—what data would you collect? What would the device look like? What would you do with the information about yourself? How might others—family, colleagues, friends, professionals, other participants—respond to these digital data machines? For the second task, participants were asked to design a personal data machine to collect data on *other people* and reflect on the application of that design. We encouraged participants to be as innovative or personal or surreal as they wish. With this in mind, we ran a slideshow that presented a range of innovative or speculative technologies for data gathering and display. Figure 3 is an image of notes that one pair put together in response to this task.

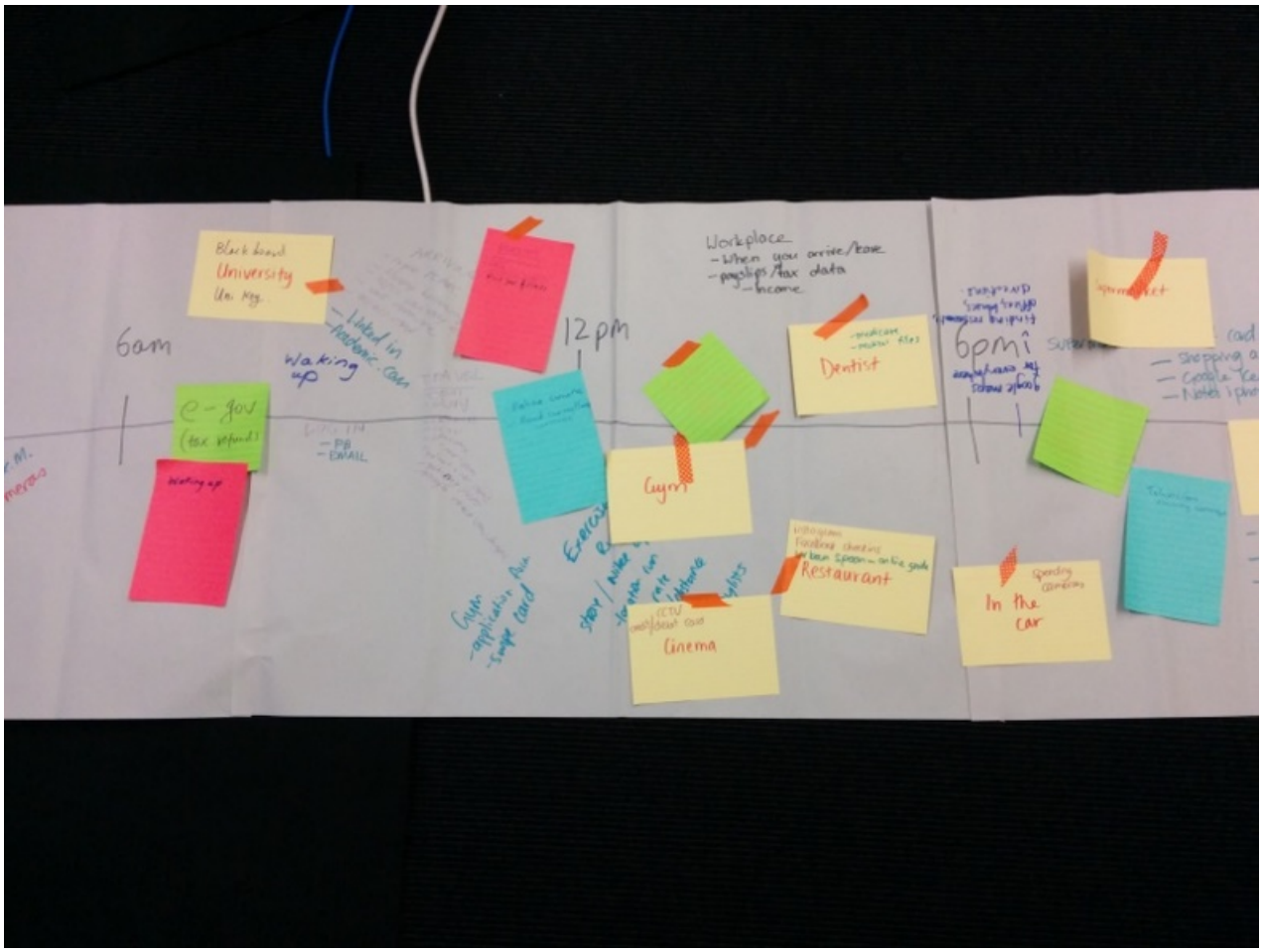

***Figure 1**: Example of a timeline produced from the Daily Big Data Task*

We note here that none of these tasks was designed to elicit the participants' understandings of specific modes of dataveillance. As part of our experimental method, we did not seek to pre-empt what material might be generated by the tasks. As it happened, however, each task worked to inspire thought and discussion on different forms of dataveillance, as we outline below. Following the completion of each task,

we asked the group to reflect on it and discuss its implications for individuals' experiences of being data subjects. Final questions at the end of the discussions further invited the participants to provide feedback on their experiences of being part of the group and carrying out the tasks. All the focus groups were facilitated by our research assistant, who had been instructed by us in the ways to use the cultural probes with the groups. The project was approved by the human ethics committee of the University of Canberra.

Our research data comprised of the material artefacts that were generated from these tasks (the timelines generated by each group and the notes and drawings that they made when responding to the other two tasks, examples of which are shown in Figures 1-3) as well as the transcriptions of the audiotapes of the group discussions during and following the completion of each task. We both read through the discussion transcripts and viewed the material artefacts generated from the tasks to identify recurring themes. We discussed our interpretations with each other as part of conducting joint analysis of the data.

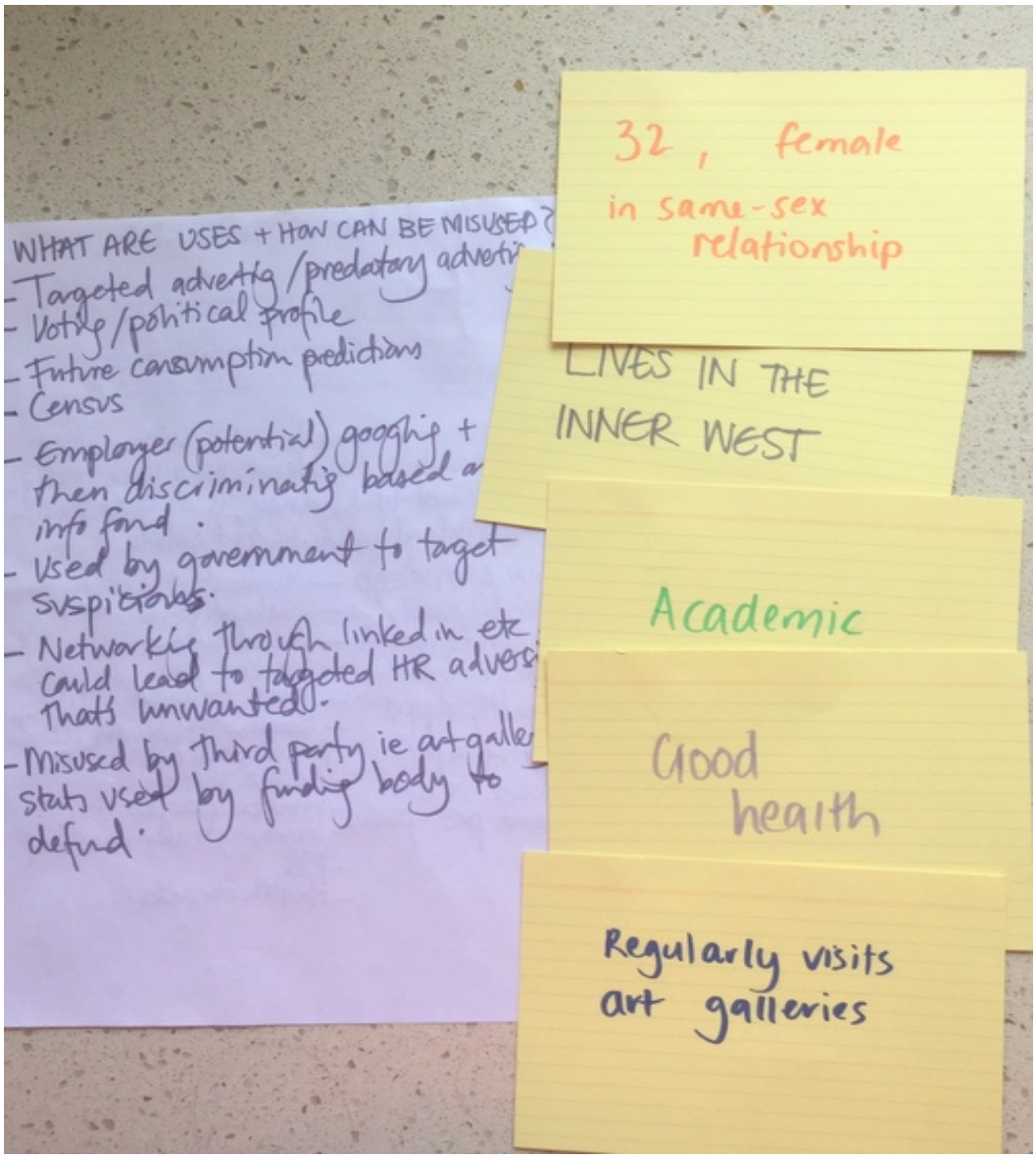

***Figure 2**: Example of cards used and participants' notes from the Digital Profile Card Game*

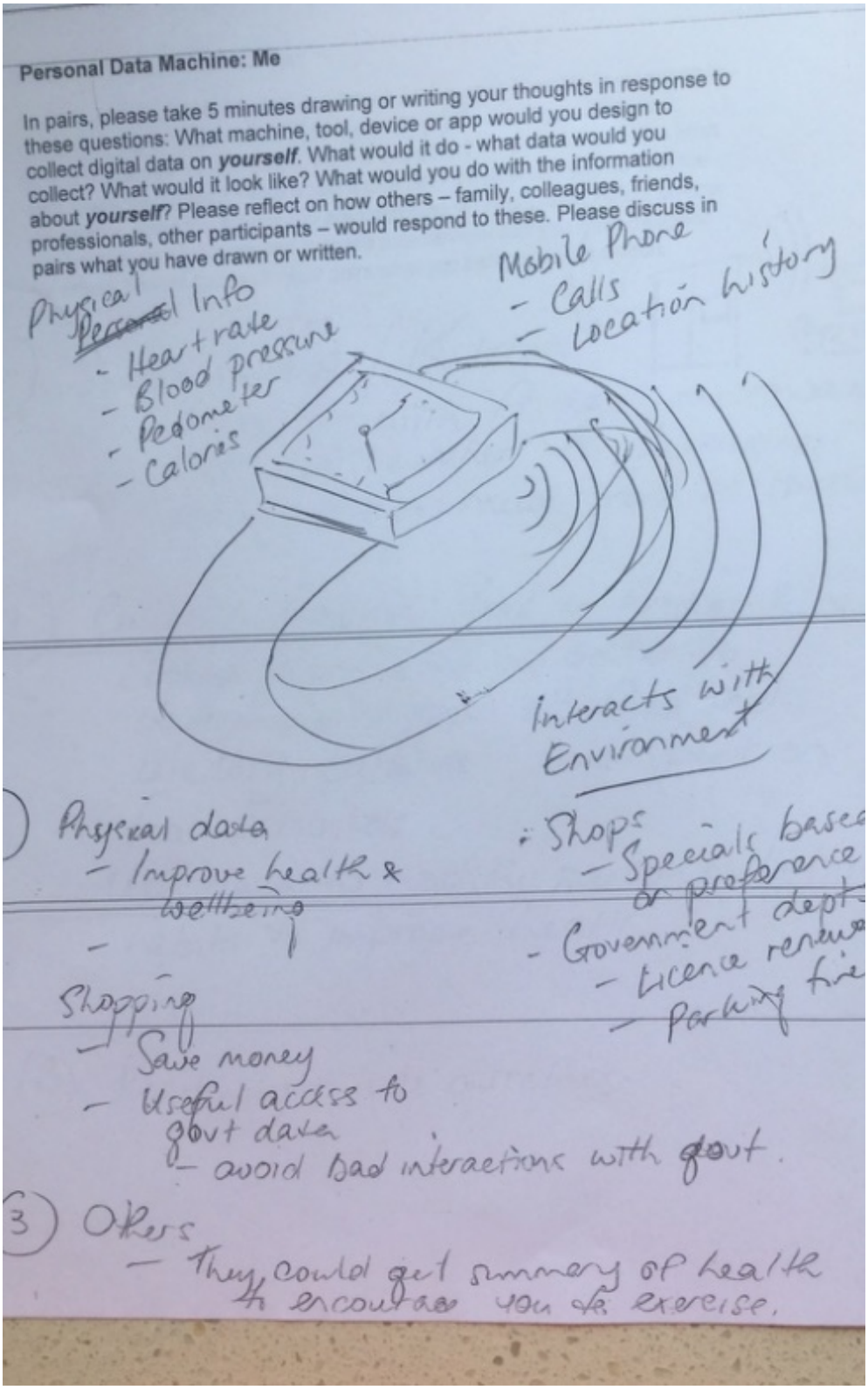

***Figure 3**: Example of a Personal Data Machine created by one pair*

Details of Participants

We first held a pilot focus group with university students in Sydney to test our methods, using personal contacts for recruitment. After running this pilot group, we modified our methods slightly and provided more explanation for the ensuing focus groups. The participants for the remaining five focus group participants were recruited by a Sydney-based social market research company that specialises in focus group research. This company was briefed to recruit participants from its existing panels (comprised of people who volunteer to take part in social research). We asked the company to recruit people living in Sydney, equal numbers of men and women and from a diverse range of social backgrounds. The focus groups were all held at a central Sydney location in late 2014. All participants (including those in our pilot group) were remunerated with a gift card for their participation.

This strategy of recruitment resulted in the following sociodemographics of the total of 48 participants across the six focus groups. Apart from members of the pilot focus group, some of whom knew each other, none of the participants had met before. We achieved recruitment of equal numbers of women and men, but despite attempts to recruit participants of diverse ages, the age range of participants was predominantly 18 to 35 years. Fifteen people were aged between 18 and 25 years, 20 people were aged between 26 and 35 years, eight people were aged between 36 and 45 years and five people were aged between 46 and 55 years. No participants were aged over 55 years. Participants were asked to nominate their ethnic background. Twenty-two people said that they were of British/Irish ancestry, 11 were Asian-Australian, 11 were of continental European ancestry, four nominated themselves as having Middle-Eastern ancestry, and one each said South American and South African ancestry (two people nominated two of these ethnic backgrounds). Regarding their education level, 18 participants had either completed a postgraduate university degree or had engaged in uncompleted postgraduate study, 11 had completed one or more undergraduate degrees while a further six were still studying at university. A further 13 participants had experienced no university-level education (seven had some technical training; six had completed secondary school education). As these details demonstrate, the participant group was relatively young and highly educated, although their ethnic background was quite diverse (as is the population of Sydney).

Findings

The Daily Big Data Task

It was clear from this task and ensuing discussions that many of the participants had some awareness of the diverse ways in which surveillance technologies, mobile devices, search engines and social media sites collect information about people's activities. When they were constructing their timelines outlining how personal data may be collected on an average person, for the most part, discussion centred on such modes of dataveillance as CCTV cameras in public spaces, government institutions, transport systems, customer loyalty programs or online shopping databases and workplaces or healthcare facilities keeping digital records of people. The activities of commercial enterprises such as Google and Facebook were also routinely mentioned as tracking personal details of users.

In Group 1, for example, as the group members were completing their timeline, they talked about credit cards, customer loyalty programs, social media, power and water usage, taxation records and medical and pharmacy records and how these were sources of different forms of data about users. Group 3 participants nominated CCTV cameras in public spaces, public transport smart cards, smartphones, digital tracking of postal deliveries by Australia Post, digital payment systems, smart mattresses and wearable devices for tracking sleep patterns and other biometrics, Wi-Fi providers having access to geolocation and internet browsing practices, Facebook tracking users' likes and preferences, Google tracking browsing and searches, credit card use, gyms tracking use of their facilities by members and satellites tracking people. As well as mentioning some of these technologies for generating personal digital data, Group 4 participants nominated swipe cards to enter buildings, digital road toll collection systems, online shopping, restaurant bookings,

work computer software monitoring employees, car GPS systems, road traffic monitoring systems, water and electric usage, monitoring of ticket sales and attendance at sporting events and electronic diaries.

The dual nature of some forms of collecting personal data on people as they moved around in public spaces, including the use of a transport 'smartcard' (the Opal card) for travelling on the Sydney metropolitan public transport system was discussed in Group 5. As members of this group noted, they are using the Opal card or moving around in sensor-embedded public spaces for their own particular purposes. As people use these objects or travel through these spaces, data about people's movements are automatically generated about them in ways in which people may be unaware.

This mode of description of dataveillance was also evident in the following discussion as part of Group 3's observations on how people are monitored in public spaces:

> Male: But yeah, there are cameras everywhere—shopping centres, post office, the airports, hotels, public transport, uni, school yard, office, gym I guess.
> Female: Even phones—people can hack phones and turn on cameras and computers now at the push of a button.
> Facilitator: What else might be collected on you in those places other than video footage?
> Female: Calls and messages.
> Female: Yeah, so, for example if you're at the pub and you access their Wi-Fi they can get all your details and they can track what sites you've been looking at—they have filters too.
> Male: The other thing is that your phone. Smartphones use Wi-Fi for location as well.

As was evident in this discussion excerpt, the use of the vague term 'they' about actors or agencies collecting personal data and then using the data for their own purposes was common. People realise that 'they' are collecting the information, but specific details of who 'they' may be or how these actors may be using the information were not forthcoming. Later in the discussion in Group 3, participants talked about their concerns about where their personal data ended up:

> Male: Yeah, I suppose the big concern is the different agencies collect different sorts of data on the consumer. Then how they pass it on—what sort of agreements they have with their partners—is a concern.
> Facilitator: So not necessarily the data itself, but where it's sent to.
> Male: That's correct. For example Facebook. Facebook would be collecting what time you log in, what sort of friends you have; what sort of activities you do, et cetera, et cetera. What's your buying pattern when it comes to the pages, the different pages that you look at. Based on that, they would recommend other similar products, so that's advertising. Obviously, they are passing on the activity that I do to someone else, because someone else is contacting me.

The research material from this task suggested that the participants were particularly aware of 'top-down' dataveillance activities. Participants in Group 2 made reference to 'the government' using surveillance cameras and 'spying on everyone'. However, neither these participants nor people in any other of the groups directly discussed Snowden's revelations or made reference to national security agencies' dataveillance of their citizens. Nor was there much raising of hacking and cyber-criminal activities, although identity theft, frauds and scams were discussed as a potential misuse of personal digital data in some group discussions. Furthermore, there was very little recognition of 'sousveillance' or 'social surveillance': the watching of each other that takes place on social media platforms or in public places. Dataveillance was almost exclusively understood as undertaken for commercial, security or government-related purposes.

The Digital Profile Card Game

This cultural probe incited greater reflection about how certain details about a person gathered from digital interactions and sensors could be revealed to other actors and agencies. A discussion of this task by Group 3 members demonstrates their understandings of this type of data collection. When discussing the profile of a 63 year old single woman who is a full-time student, doesn't drive, enjoys listening to the rock group Queen and has a chronic health condition, they reflected on how these details were elicited. Healthcare and pharmacy records, downloading music habits or Spotify preferences, a digital public transport card, her online browsing and searching habits and her university student online platform were identified as possible sources of information about this woman. When discussing how this information could be misused, the group identified frauds and scams, identity theft, the possibility for thieves to break into her house if they knew that she was in hospital, dubious advertising of alternative medical treatments or targeting by Big Pharma for advertising of drugs to help her medical condition.

People in Group 6 discussed their profile of a 32 year old female academic in a same-sex relationship who lives in the inner west of Sydney, is in good health and regularly visits art galleries (the materials and their written notes in response are shown in Figure 2). They surmised that these details could have been collected by this woman's credit card and other online purchasing records, a survey at an art gallery, online subscriptions to magazines or journals, her professional profiles on Academia.edu, Google Scholar or LinkedIn, her use of dating apps, ticket purchases, gym membership details, her personal university website page, her library membership, her use of Facebook and her electronic health records. They identified the uses or misuses of such information as including the conduct of targeted advertising, to construct a political preference profile on the woman, to predict her future consumption habits, use by prospective employers seeking information about this woman, government use for security surveillance purposes and networking through LinkedIn leading to possible targeting by unwanted advances from head-hunters.

When discussing this task, one person noted that geolocation data could be used both to incriminate people accused of a crime and to provide an alibi if they were innocent. Another participant provided the example of a person with severe depression (based on the cards provided to his group), noting that 'it depends on who's got the data' how helpful or damaging such dataveillance could be:

> If his depression was known among his friends, for example, they could ask how he's going and prevent him—say, that he hasn't been in contact for a while, they could pick up the phone and say, 'Hey, are you going through a bad time?' or whatever. But on the other hand, it could be negative if his employer or somebody else finds out about it. There's lots of prejudice against mental illness.

As this account suggests, engaging in this task drew out more nuanced accounts of the various forms of dataveillance that take place on online forums. This participant acknowledged that users of digital media platforms such as social media may have access to each other's personal details as part of their engagement on these sites, and may even be able to use this knowledge productively to support people who may be experiencing negative life events. This task, therefore, worked well to draw participants' attention to voluntary and intimate forms of dataveillance and raise discussion of their more consensual and positive aspects, as well as highlighting top-down organisational dataveillance.

The Personal Data Machine

Further insights came to light when the participants were asked to design and explain to the group their 'Personal Data Machine'. As part of this task, a willingness to use digital devices to participate in ever-more intrusive forms of surveillance of oneself or others was apparent.

Devices that were able to monitor closely users' bodily functions were a popular choice, such as one that involved analysing the user's sweat to determine whether they were eating a nutritious diet. One pair devised a heart-shaped pendant for continually measuring heart rate, designed to look like conventional jewellery so that it was unobtrusive. In the same group, another pair had designed an app that could accurately measure the nutritional content of food. One described it thus:

> A smartphone app that you can use to scan food objects in front of you: like that can of Sprite. It would be able to recognise what it was and then display all the information. If you had existing data, it could tell you, 'No you've had too much sugar today! Don't eat that.' Help you to make a plan and keep on that plan.

This device was one of several designed by other pairs that monitored biometrics or food intake and then provided advice for how people should live to achieve good health, weight loss or physical fitness. One imagined device was designed to be placed into the user's mouth as they ate so that the nutritional content of the food they were eating could be monitored and assessed, and simultaneously their kilojoule expenditure could be monitored. One pair described a device similar to the new Apple Watch, with even greater capabilities than the Watch to monitor and record many aspects of their lives. This device, for example, would know the user's preferences for shopping, and would alert them when the user is walking in a shopping district that sales for items that they might want to buy are on. The device would also beep when the user was near a government department and notify them of their overdue parking fine or the requirement to renew their driving licence.

Another invention engaged with the Internet of Things, interacting with all the machines in the user's kitchen (such as fridges, stoves, ovens, waste disposal and dishwasher) and eliciting information about food consumed and ready to eat, dishes washed and so on. The associated app would know exactly what food users were preparing and eating and would then provide dietary and nutrition advice, or advise people what food they needed to purchase. It would upload information to users' mobile devices so that they would always have this information to hand. One invention involved a 'smart home' app. This worked to ensure that lights and heating were switched on and off at appropriate times, monitored how much money was spent on household bills, the dog was fed and the groceries purchased.

The participants were perhaps even more inventive when they were designing personal data devices that would collect information on other people. Thus, for example, one pair designed a dream-recording app that would allow them to remember their dreams the next day. They went on to describe how this could be linked to a dating app so that prospective couples could share each other's dreams and perhaps work out how compatible they were. Another pair discussed a data machine that could monitor the social interactions of people's partners so that the user could determine if too great a level of attention was being paid by their (possibly unfaithful) partners to other people. Further devices that were invented included a lie-detecting device, one that could track commercial competitors' activities, another that revealed the salary of workmates (so that the user could know if they were fairly remunerated) and a dating device that could scan a prospective partner's hand or face and reveal their financial assets and criminal record details.

A device for measuring the geolocation and productivity of workers, in the form of a wearable wristband, was another suggestion for a personal data machine that could be used to collect information on other people. A similar tool suggested was software that allowed employers to monitor the websites that their workers visited. Other people described software installed across individuals' digital devices that could track people's use of time spent on different tasks, to ensure that they used their time productively. Devices for keeping a watchful digital eye on one's children were also frequently suggested, including features that could let parents know the location of their children, record their biometrics, monitor their food intake and activities and check that they were doing their homework. One pair, for example, discussed their invention of a microchip that could be placed somewhere on a child so that parents would know their location at all

times. Another pair invented software that would track children's computer use to determine how much time and effort they were putting into their homework.

This task, therefore, worked to elicit the participants' reflections on how the personal data potentially collected by dataveillance devices could be useful in their own lives. It served to uncover some of the elements of their lives and social relationships that these devices could reveal; and in so doing, some of the seductive features of dataveillance were revealed.

Reflections on the tasks

Our participants observed that they enjoyed the tasks that we set them. For many of them, the tasks had made them consider digital data understandings and practices in a different or new light, as they had not been called on to reflect on these issues previously. It was common for participants to remark on how much they had learnt about the ways in which digital data were generated about the public. As a woman in Group 1 noted, the Daily Big Data timeline task was 'eye-opening' for her, 'because I never thought how many people were collecting data about me. I just noticed a few things I never thought about'. For another female participant, the knowledge that 'some people out there know as much about you as you know about yourself' is 'scary'. She observed that 'there is a lot going on that we don't know' regarding how other actors are accessing people's personal data. This discussion eventuated in Group 5 at the end of the focus group:

> Facilitator: Does this discussion raise any ideas or issues with you about data collection? With yourselves?
> Female: There's a lot of it that I had forgotten about or was ignoring.
> Facilitator: Okay.
> Female: I think I probably give out my personal information maybe a little too much. Because I sign up for …
> Male: Yeah, sign up for a lot of things…
> Female: Even just bars and pubs and shopping.
> Female: Makes it more obvious how dependent we are on ourselves, not just being a person, but being a presence that's not just affecting ourselves. Online in databases and clubs that we join and that sort of thing that we take lightly but create a huge picture of ourselves.
> Female: We don't really have a choice in what we give, either. If you want an app, you have to let them use all your data, as well. But the app could be really useful to you, so …
> Facilitator: So there's positives and negatives to it.
> Female: Yeah. I don't think I'd think anything too negative unless something bad happened to me as a result of giving away too much data.

As this exchange suggests, one of the issues that people are coming to terms with is the extent to which information about themselves—even what may seem to them mundane dimensions of their lives—may have a commercial (or research, or managerial) value for others. The ways in which their digital data are monetised were not well understood. It is also clear that many people do not consider the implications of signing up to apps or other software in terms of how their personal data may be used by others. They feel that they have little choice in relinquishing control over the data.

For some people, it is not necessarily important how much control they have over their own data and how they are used which is important, but the ways in which people's personal data are used is the important issue. In one of the discussions following the Daily Big Data Task, for example, a male participant had this to say:

> It's how the information is controlled and accessed. To me, that's the biggest problem about this whole Big Brother thing … I don't care what they collect on me. They can collect any information they want on me. As long as only the people who need to use that information, for my benefit, whether that be the community's benefit or individually. And that's where I think we have a problem at the moment.

Discussion

We found that our adapted use of the cultural probes technique worked well in eliciting people's responses. Each of the tasks that we set our participants enabled them to think through and comment on aspects of personal digital data assemblages from different perspectives, thus eliciting details of greater depth for our research purposes. What emerged from our focus groups is a somewhat diffuse but quite extensive understanding on the part of the participants of the ways in which data may be gathered about them and the uses to which these data may be put. It was evident that although many participants were aware of these issues, they were rather uncertain about the specific details of how their personal data became part of big data sets and for what this information was used.

Our participants were often highly aware that companies such as Facebook and Google are tracking their preferences, habits and the content that they upload to social media. Evidence of this monitoring is obvious from the targeted advertising that is delivered to people when they are using these sites. What was unclear for most of the participants is the detail of how this data exchange takes place, or which other parties may have access to their data. While several people used the term 'scary' when describing the extent of data collection in which they are implicated and the knowledge that other people may have about them from their online interactions and transactions, the participants struggled to articulate more specifically what the implications of such collection were. Similar findings were evident in Kennedy et al.'s focus groups held in the UK, Spain and Sweden (2015) and in another focus group study with British teenagers (Pybus, Coté, and Blanke 2015).

Like the participants in Nafus' study interviewing people in the UK and the USA who used home energy-monitoring systems, some of our participants were aware, as Nafus (2014: 217) puts it, of 'data's capacity for betrayal'; its potential to reveal private details about people to others in ways that may be unanticipated or unwanted. In the participants' accounts, it was evident that the context of the use of personal data was vital to judging how 'scary' the implications of digital data surveillance may be. We found that the participants tended to veer between recognising the value of both personal data and the big aggregated data sets that their personal data may be part of, particularly for their own convenience, and expressing concern or suspicion about how these data may be used by others. As the man we quoted above commented; 'in a lot of ways information is very useful'. He acknowledged that he may have little control over or even knowledge about what information is collected on him. However, for this participant, and for some others, the ways in which this information is repurposed is also important.

Yet, as evidenced by some of the devices and discussions in relation to the Personal Data Machine, the lure and promise of generating reams of personal information about the self or others in one's life appears in some ways to obviate the knowledges that the participants articulated about 'Big Brother'-type or purely commercial surveillance. Particularly in their response to this cultural probe, the participants drew on sociocultural scripts that valorise dataveillance. They could easily conceptualise the possible benefits of using digital technologies to engage in self-surveillance or intimate surveillance of others, often in the interests of disciplining their lives. They foresaw the importance of collecting personal data to promote their health, wellbeing and productivity. So too, the ability to conduct intimate surveillance on others in one's lives—partners, children, workmates and so on—offers tempting opportunities to exert greater control over these others. These types of devices afford a type of watching that is not so much social surveillance as frank surveillance of others. Unlike social surveillance, there is no expectation that people are watching

each other. Rather, such devices invite users to position themselves in the 'Big Brother' role as they use digital technologies to pry into the lives, activities and even dreams of others. It is interesting that most of these imagined devices assumed that the device inventor would be watching 'from above', or even covertly, to obtain information to benefit the inventor rather than those they were watching.

Our findings suggest the prevailing dominance of tacit assumptions about the uses and benefits of dataveillance as well as fears and anxieties about its possible misuse. The major question for our participants was where the line should be drawn. When does personal dataveillance become too intrusive, scary or creepy? What are its drawbacks and risks? These are the issues with which our participants are grappling. Members of the public and regulatory agencies alike are attempting to make sense of the interactions and intersections between small and big data and the rapidly growing ways in which these data are purposed and repurposed in a constantly shifting context in which some people's life chances and opportunities are enhanced, and those of other people are limited by the digital data that are generated about them.

Conclusion

Our findings suggest that experimenting with innovative approaches to eliciting public understandings of personal digital data offers further possibilities for greater depth and breadth of social research with all types of social groups Our project has only just scraped the surface regarding eliciting and identifying the manifold ways in which members of the public are responding to and making sense of their personal digital data assemblages. Future research needs to delve into other aspects of the complex nature of these lively assemblages and how people gain purchase on them: including the ways in which they contribute to and use their digital data assemblages, how they live alongside them, domesticate them and co-evolve with them (Lupton 2016a).

As we have acknowledged, our participants were relatively young and highly educated. People who are already marginalised, excluded from education and employment opportunities and access to digital technologies or knowledge about how to use these technologies are far more vulnerable to data privacy breaches or further discrimination based on the information that they may contribute when interacting with digital technologies (Gangadharan 2015; Rosenblat et al. 2014). Much more work needs to be carried out on this topic with people who are less educated or otherwise socioeconomically-disadvantaged and in older age groups. Building on this initial research, we are developing methods that will attempt to investigate not only how people understand and make sense of the social lives of their personal digital data assemblages, but also how they incorporate these assemblages into their lives.

Acknowledgements

This research was funded by personal research funds provided by the University of Canberra to Deborah Lupton. We thank Sophie Johnson for conducting the focus groups and the participants for agreeing to take part.

References

Albrechtslund, Anders, and Peter Lauritsen. 2013. Spaces of everyday surveillance: Unfolding an analytical concept of participation. *Geoforum* 49:310-316. doi: http://dx.doi.org/10.1016/j.geoforum.2013.04.016.

Andrejevic, Mark. 2013. *Infoglut: How Too Much Information is Changing the Way We Think and Know*. New York: Routledge.

Andrejevic, Mark. 2014. The big data divide. *International Journal of Communication* 8:1673-1689.

Andrejevic, Mark, and Mark Burdon. 2015. Defining the sensor society. *Television & New Media* 16 (1):19-36.

Boehner, Kirsten, William Gaver, and Andy Boucher. 2012. Probes. In *Inventive Methods: The Happening of the Social*, edited by Celia Lury and Nina Wakeford, 185-201. London: Routledge.

Boehner, Kirsten, Janet Vertesi, Phoebe Sengers, and Paul Dourish. 2007. How HCI interprets the probes. Proceedings of CHI 2007, San Jose. ACM Press, 1077-1088.

boyd, danah, and Kate Crawford. 2012. Critical questions for Big Data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15 (5):662-679.

Crawford, Kate, and Jason Schultz. 2014. Big data and due process: toward a framework to redress predictive privacy harms. *Boston College Law Review* 55 (1):93-128.

Esposti, Sara Degli. 2014. When big data meets dataveillance: the hidden side of analytics. *Surveillance & Society* 12 (2):209-225.
Fuchs, Christian. 2011. Web 2.0, prosumption, and surveillance. *Surveillance & Society* 8 (3):288-309.
Gangadharan, Seeta Peña. 2015. The downside of digital inclusion: Expectations and experiences of privacy and surveillance among marginal Internet users. *New Media & Society*. DOI: 10.1177/1461444815614053.
Gaver, William, Andrew Boucher, Sarah Pennington, and Brendan Walker. 2004. Cultural probes and the value of uncertainty. *Interactions* 11(5):53-56.
Kennedy, Helen, Dag Elgesem, and Cristina Miguel. 2015. On fairness: User perspectives on social media data mining. *Convergence* online first. DOI: 10.1177/1354856515592507.
Klauser, Francisco R., and Anders Albrechtslund. 2014. From self-tracking to smart urban infrastructures: towards an interdisciplinary research agenda on Big Data. *Surveillance & Society* 12 (2):273-286.
Levy, Karen. 2015. Intimate surveillance. *Idaho Law Review* 51:679-693.
Lupton, Deborah. 2016a. Digital companion species and eating data: Implications for theorising digital data–human assemblages. *Big Data & Society* 3 (1). Available at: http://bds.sagepub.com/spbds/3/1/2053951715619947.full.pdf. Accessed 7 January 2016.
Lupton, Deborah. 2016b. The diverse domains of quantified selves: self-tracking modes and dataveillance. *Economy and Society* 45(1):101-122.
Lupton, Deborah. 2016c. *The Quantified Self: A Sociology of Self-Tracking Cultures*. Cambridge: Polity Press.
Lyon, David, and Zygmunt Bauman. 2013. *Liquid Surveillance: A Conversation*. Oxford: Wiley.
Madden, Mary. 2014. *Public Perceptions of Privacy and Security in the post-Snowden Era*. Pew Research Center.
Madden, Mary, and Lee Rainie. 2015. *Americans' Attitudes about Privacy, Security and Surveillance*. Pew Research Center.
Mann, Steve, and Joseph Ferenbok. 2013. New media and the power politics of sousveillance in a surveillance-dominated world. *Surveillance & Society* 11 (1/2):18-34.
Marwick, Alice. 2012. The public domain: social surveillance in everyday life. *Surveillance & Society* 9 (4):378-393.
Michael, Mike, and Deborah Lupton. 2016. Toward a manifesto for the 'public understanding of big data'. *Public Understanding of Science* 25 (1):104-116.
Nafus, Dawn. 2014. Stuck data, dead data, and disloyal data: the stops and starts in making numbers into social practices. *Distinktion* 15 (2):208-222.
Nuffield Council on Bioethics. 2015. *The Collection, Linking and Use of Data in Biomedical Research and Health Care: Ethical Issues*.
Polonetsky, Jules, and Omer Tene. 2013. Privacy and big data: making ends meet. *Stanford Law Review Online* 65. Available at: http://www.stanfordlawreview.org/online/privacy-and-big-data/privacy-and-big-data. Accessed 4 September 2013.
Public Administration Select Committee. 2014. *Statistics and Open Data: Harvesting Unused Knowledge, Empowering Citizens and Improving Public Services*. London: The House of Commons.
Pybus, Jennifer, Mark Coté, and Tobias Blanke. 2015. Hacking the social life of Big Data. *Big Data & Society* 2 (2). Available at: http://journals.sagepub.com/doi/full/10.1177/2053951715616649. Accessed 12 May 2016.
Rosenblat, Alex, Kate Wikelius, danah boyd, Seeta Peña Gangadharan, and Corrine Yu. 2014. *Data & Civil Rights: Health Primer*. Data & Society Research Institute. Available at: http://www.datacivilrights.org/pubs/2014-1030/Health.pdf. Accessed 16 December 2014.
Trottier, Daniel, and David Lyon. 2012. Key features of social media surveillance. In *Internet and Surveillance: The Challenges of Web 2.0 and Social Media*, edited by Christian Fuchs, Kees Boersma, Anders Albrechtslund and Marisol Sandoval, 89-105. New York: Routledge.
van Dijck, José. 2013. *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford University Press.
van Dijck, José. 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society* 12 (2):197-208.
Vetere, Frank, Martin R. Gibbs, Jesper Kjeldskov, Steve Howard, Florian 'Floyd' Mueller, Sonja Pedell, Karen Mecoles, and Marcus Bunyan. 2005. Mediating intimacy: designing technologies to support strong-tie relationships. Proceedings of CHI 2005, Portland, 471-480.
Wellcome Trust. 2013. *Summary Report of Qualitative Research into Public Attitudes to Personal Data and Linking Personal Data*. http://wellcomelibrary.org/item/b20997358#?c=0&m=0&s=0&cv=0.
Zuboff, Shoshana. 2015. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology* 30 (1):75-89.