

- Powell, A. 2012a. 'Democratizing Production through Open Source Knowledge: From Open Software to Open Hardware'. *Media, Culture & Society* 34(6): 691-708.
- Powell, A. 2012b. 'Critical Making, Teaching & Politics'. In G. Hertz (ed.), *Critical Making*. www.conceptlab.com/criticalmaking/ [accessed 20 February 2018].
- Powell, A. 2015. 'Open Culture and Innovation: Integrating Knowledge across Boundaries'. *Media, Culture & Society* 37(3): 376-93.
- Powell, A. 2016. 'Coding Alternative Modes of Governance'. In *Code and the City*. Edited by R. Kitchin and S.Y. Peung. London: Routledge, 178-95.
- Ratto, M. and Boler, M. (eds). 2014. *DIY Citizenship: Critical Making and Social Media*. Cambridge, MA: MIT Press.
- Smith, P. 2016. 'Walking and Subjectivities'. Accompanying presentation, Asylum Topographies, Bethlem Gallery, Kent.
- Stengers, I. 2010. *Cosmopolitics*. Minneapolis: University of Minnesota Press.
- Stiegler, B. 1998. *Technics and Time, 1. The Fault of Epimetheus*. Stanford: Stanford University Press.
- Thrift, Nigel. 2014. 'The "Sentient" City and What It May Portend'. *Big Data & Society* 1(1): 1-21.
- Tironi, M. and Criado, T.S. 2015. 'Of Sensors and Sensitivities. Towards a Cosmopolitics of "Smart Cities"? *TECNOSCIENZA: Italian Journal of Science & Technology Studies* 6(1): 89-108.
- Van Dijk, J. 2014. 'Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology'. *Surveillance & Society* 12(2): 197.

10

Working ethnographically
with sensor data

Dawn Nafus

This chapter is primarily about methods. I work in Intel Labs, the research and development organisation at Intel. Since 2007, I have been asking research participants to collect digital data about themselves, and giving it back to them in forms designed to stimulate conversation. I invite participants to reflect on data as matters of concern, not matters of fact (Latour 2004), and they largely respond in this spirit. Much like the chapter from Powell (Chapter 9 above), and in the spirit of the broader turn towards critical geographic information systems (GIS) (Wilmott 2016) and community-based participatory research (Corburn 2005), this approach evolved as a way to reclaim the situatedness of data, and create space for more participatory, humanistic ways of engaging with data. This space necessarily sits alongside positivist modalities of using datasets, and occasionally creates room for dialogue between positivists and those who value situated knowledge. Even those who value generalisability as a goal, and treat situatedness a bias to be avoided, often can agree that best person to explore what a dataset means is often the very person to whom it refers. Perhaps more significantly for anthropology, this approach offers ethnographic opportunities that are hard to come by in other ways. Data evokes the past without fully predefining it, and that opens up conversations between researcher and the participants who understand the context behind it.

My first goal for this chapter, then, is to describe in a relatively straightforward way what this research practice involves for those who

might be curious about adapting it for their own research situation. The considerations for setting up the research and working with these forms of data are non-trivial, but the approach can be done by those of us who do not possess engineering or data science skills. This was not always the case. Five or ten years ago, working with sensors used to be much more difficult. Even today, sensing systems can be tricky to set up, and data visualisations can fall flat. Working with data as a qualitatively trained person always produces a certain amount of hesitance about where one's capacities reach their limits, and when it is necessary to seek help from others. Therefore, I hope to share some of the pitfalls and possibilities I have found.

These research practices have evolved in a context where the material infrastructures available have been transforming. In a sensor-rich world, the same tools that create massive data stockpiles at technology companies and scientific laboratories also create smaller, but still richly analysable data streams that have a different sort of value. Sensor data creates circumstances where analysis can grow numerically complex even when that data refers to a single person, let alone a whole community (Nafus 2016). The technical systems that people access as consumers show data in ways that are quite different from what data scientists can see, who have fuller access to those systems. Indeed, this unequal access has been a chief complaint about contemporary big data cultures (Andrejevic 2014). Not only can large institutions black box the algorithms they use to parse data, they can also take advantage of the cross-population insights that cohere only when data points are collected across many people (Pasquale 2015).

While this is clearly a problematic aspect of institutional power, there is also a significant technical challenge that has largely been overlooked by social-scientific commentary. Even if companies or publicly funded research projects did become more transparent about what they do with data, either by choice or through legal mandate, what tools are realistically available for data subjects – the people to whom data refers – to create alternative views, or appropriate the data in ways that suit their agendas better? What tools would facilitate meaningful participation in big data? The systems that professional data scientists use offer the broadest possible flexibility to make a wide range of calculations, but their use presupposes advanced mathematical and computer science skills. Yet my research had been showing that 'ordinary' people do in fact have something to say about which

patterns matter and which patterns do not. Big data is not just about esoteric calculations that require machine learning algorithms. A simple recurrence in time can still show something conceptually simple and useful beyond expert circles, yet still be hard to achieve in a spreadsheet.

Threaded through the methodological work described below, I have also been participating in technical work to address this problem. The software project I co-lead, Data Sense, attempts to facilitate 'lay' data exploration by people who do not possess coding or statistical skills. The project is a technical response to what I had been seeing as research participants encountered data with me. This software is but one of many developments in the technological infrastructure that makes it conceivable for non-experts to use sensor data for their own purposes.

A secondary goal for this chapter, then, is to contribute to the discussion about the entanglements between materials, sociality and methods. The introduction to this volume argued that, in a data-saturated world, it should not be surprising that data might be found in ethnographic accounts as artefacts that participants already have, and as artefacts that ethnographers set out to collect. When I began using this approach in 2007, I did not encounter research participants who already had data I was looking to collect. Now when I do fieldwork participants regularly offer their Fitbit data, or ask the research team to look more deeply at data they have, like air quality data. The tools and the distribution of labour in research have changed, too. When I started this work, it was in collaboration with a computer scientist, who departed our lab leaving a sizeable hole in the lab's research practice. I did not know at the time that the software I would later help develop, in conjunction with other technological developments on the market, would also fill some of that hole. Although I cannot do everything a computer scientist can do, changes in the available technology mean that as a non-engineer I can now get some meaningful data out of a sensor system, parse it in some limited ways and know what to ask for when the situation requires someone writing a script to do the job. The interdependencies between materials and methods, and the distribution of labour, become clearer as they evolve in the long term. That trajectory is what I wish to show.

This transformation in research practice underscores the argument made by the social life of methods approach (Law et al. 2011; Ruppert

et al. 2013; Savage 2013), which suggests that methods cannot be disentangled from the ongoing social lives they are designed to comprehend. What starts as one person's method becomes another person's everyday practice. The changes in my research practice also echo the inventive method (Lury and Wakeford 2012), in that they show that 'it is not possible to apply a method as if it were indifferent or external to the problem it seeks to address ... if methods are to be inventive, they should not leave that problem untouched' (Lury and Wakeford 2012: 3). Indeed, I hope to suggest that taking the social life of methods seriously also means intervening in the tools that sustain cultures of big data. It is not enough to comment on the emergence of new technical systems as if we ethnographers were untouched by them. Nor is it enough to merely appropriate the techniques and technologies of other research practices to conduct ethnography. While not every anthropologist needs to become invested in actively shaping big data systems, I hope to demonstrate that an ongoing anthropological and sociological voice in their making is worthwhile both for our discipline and for our research participants, who are as entangled in a data-saturated world as we are. The materials matter to what others can do with big data, and what we can do with it, too.

In the methodological account that follows, I rely on a loose, somewhat shifting figure of the reflexive data subject. Like Couldry and Powell (2014), Kitchin (2014), Wilmott (2016) and others in this volume, I take the view that big data methods should make space for reflexivity and participation by data subjects. While my approach bears some resemblance to 'trace ethnography' (Geiger and Ribes 2011) in studies of computer-supported collaborative work (CSCW), where researchers immerse themselves in online documentation to understand the social relations that generated them (see also Dumit on 'data archaeology' in Chapter 11 below), like Dumit I place much more emphasis on speaking to live human beings about data. Here, participants' reflections can extend into a research practice in its own right. In some projects, participants have simply commented on data as presented to them, while in others they have posed their own research questions, used their own methods and formulated notions of appropriate data collection, parsing and visualisation. In some others still, participants wished to exercise some control over the research agenda while leaving it to researchers to specify how their agendas translated into numerical form. Where reflection ends and

research begins, or where an ethnographer's methods end and another's begin is not something that can be assumed *a priori*. Nevertheless, in a data-saturated world, it is safe to assume that reflexive data subjects have their own agendas with data in some way.

In what follows, I focus on three particularly illustrative projects, one on computer use, one on stress among family caregivers and one on environmental health. I show the methodological considerations that emerged in attempting to give participants data to reflect upon, and how the methods shifted in response to changing the technological circumstances and the positionalities of the people with whom I was working.

Ethnomining

In 2007, our team received a request from a strategic planner hoping to collect some market intelligence about how people were using laptops. He wanted a research project where software would detect computer use, as opposed to ethnographers asking people about how they used their computers, or observing them. He wanted those measurements to fit into the demographic categories that form the currency of market research. Do women use laptops more frequently than men? Do young people's laptops use more power than older people's, because perhaps they use video more heavily? At Intel, anthropology had traditionally been used for identifying underlying social changes that created risks or opportunities for the business, and my colleagues and I were not keen to abandon this remit in order to answer a basic market research question. However, this was 2007, and device telemetry (i.e., using software to create data about how a device is used) had not yet become standard practice. This made it a technical research problem alongside a social one, and so we took on the project despite some reluctance.

We formed a small team consisting of a computer scientist and two anthropologists, including myself. Aware that this method would never adequately capture what computers meant to people, we designed the project to mix ethnographic interviewing with what would later be called big data. Our team developed and placed software on 169 people's laptops that recorded everything that happened on those laptops on a second-by-second basis over the course of a month and a half (keystrokes, power state location, etc. – everything

that was detectable, excluding the contents of documents or online materials). The computer scientist we worked with, Tye Rattenbury, shared our scepticism about reducing matters to demographic categories, but he also showed us that there were alternative ways of parsing data. His epistemological flexibility meant we did not encounter the transversals described by Madsen et al. (Chapter 8 above). Tye did not just meet us half-way, but got curious about anthropological approaches to research, which in turn stoked our own curiosity about how he was thinking about data. This situation created the mutual trust necessary to experiment with new methods together.

Tye found that there simply *were* no patterns that meaningfully mapped on to demographic categories. No way of breaking down the data produced anything statistically significant in those terms. In an attempt to see if there were other patterns to see, he printed out a full wall's worth of data visualisations. Each chart showed when a participant's computer was used lightly or intensely for the study period. The only pattern that seemed to hold was that these 169 people tended to not use their computers during normal sleeping hours – not exactly a huge finding. When it became clear that the original request was utterly unworkable, the project began to rely more heavily on the ethnography. For a sub-set of research participants we printed out large, poster-sized plots of the data, and asked them what they thought the data showed. That work demonstrated that, across the board, people were using computers in times they thought of as interstitial – in times between the times structured by kinship, waged labour and so forth. We called this approach 'ethnomining', because at the time working with big data was not 'analytics' but 'data mining'. The favoured terms are inevitably fleeting.

With this approach we were able to show both qualitatively and quantitatively what we initially could not see by playing data scientists. Contrary to the mythical 'power user' we had been asked to find, all of the computer users interacted with their computers in short sprints most of the time. My colleague Ken Anderson knew this to be a long-standing phenomenon (he had coined the term 'Internet snacking' some years before), but we did not expect that it suffused computer use so completely. Once the ethnography showed us what was worth looking for, Tye made the relevant calculation that Intel used to redesign a chipset to accommodate interstitial use – not the sort of



Figure 10.1 Interviewing with big posters of data about computer use

change that our strategic planner expected, but a result none the less (Anderson et al. 2009).

The research payoffs for our odd, early experiment surprised us. We were surprised by the ease with which people were able to make sense of, and reflect on, a large complex dataset when it pertained to themselves. We worried that the complex plots would be illegible, but in the end what mattered was their ability to see the dates and times, and flurries of activity. The temporal patterns facilitated discussions about daily habits, what it means to keep a schedule, what constitutes productivity and relaxation etc. Some participants would use the data to show their lack of a regular schedule, and to insist on how busy they were. Others would use a tightly orchestrated regular schedule to insist on the same thing. This got us into a discussion about the relation between discourses of busy-ness and practice that we could not have had otherwise. We were also surprised that people we interviewed requested to keep the data posters we had made for them.

They enjoyed having their lives reframed in a way that was open to their own interpretations. The data was 'real', in the sense that it faithfully recorded bursts of keystrokes, but it was an open, partial record that said nothing about what those keystrokes were about. These were partial indications whose full meaning only the data subject could supply.

At a practical level, there were many aspects of our workflow that proved difficult. There was the temptation to underestimate the amount of technology support that research participants need to ensure that data is collected at all. In this first project, we had written our own software that was not easy to install, and checking that data was indeed being collected was non-trivial. We ended up having to send the software on USB sticks through the mail, and asking participants to return them full of data, but we could not rely on the exact specifications of each person's software and hardware set-up. Each computer behaved just that little bit differently. In fact, this proved so unwieldy that, when we did an extension of this study in Portugal and Bulgaria, we structured the research so that there was an initial interview where we got to know the person or family, and did the installation of the technology ourselves.

Another difficulty was the sheer amount of work involved in designing, testing and compiling the plots. Just as the first round of interview questions never really 'works' in traditional fieldwork, so too we needed to test whether the printed-out plots were evocative at all, using our own data. We could not adjust the plots on a person-by-person basis because of the time pressure involved. If too much time had elapsed between the data collection period and the time when we sat down with participants to talk about that period of time, their memories would fade. Even just running the scripts to generate the plots, and print them, was itself significant labour.

In a predecessor for this study (Aipperspach et al. 2006), my collaborators learned that visualisations that paint a 'complete' picture make for poor interview materials. In that work, researchers had created a visualisation of patterns of people's movement inside a home, and interviewed the data subjects about what it might mean. The responses there had been quite flat. Participants were more likely to say 'yup, I must have been on the couch a lot' than to tell stories about what being on the couch meant to them. In the ethno-mining study, there was a 'flaw' in the data that proved advantageous. The location

data in our study was not human-readable, and often reported a change in location when there was none. This vagueness proved to be exactly the invitation we needed to get people beyond 'yup, I was there' to thinking through where they could have been, or if it was a computer glitch. The thinking out loud in turn brought out discourses of busy-ness and routine that proved significant to the work. As we proceeded in these interviews, we had to work out a delicate balance between giving participants enough of an explanation of how the graphs worked so that they stood a chance of interpreting them at all, and giving not so much that we ended up with flat responses, or telling them what to think. Finding this balance has remained an ongoing issue.

Caregiving

While the ethnomining process was used more or less in a similar way in subsequent projects (see, for example, Nafus and Beckwith 2016), by 2015 the material circumstances had shifted in some important ways. One was that consumer-grade sensing technologies like Fitbits had proliferated, which meant that the possibility for using sensor data in research expanded beyond click data or document version data to include what could be sensed about a body, or an environment. Culturally, new expectations had developed about what that data indicated. Participants in the Quantified Self community, for example, had begun to use data from these devices to turn medical research methods 'inside out' (Greenfield 2016), and discover something medically and personally meaningful about their own bodies, that sometimes drew on medical expertise, and at other times challenged it. Indeed Data Sense started in part as an effort to help them do this. More broadly, data analysis tools and infrastructures had become more stable and easier to use. New tools like Tableau meant that with good familiarity with data structures it was possible, without coding, to do many data visualisation tasks inappropriate for spreadsheet programs. Application Programming Interfaces (APIs) – an important way of getting data out of a sensor device – had also become fairly common and more reliable. Many companies had come to understand the importance of offering 'raw' data export, often with vociferous pressure from the many Quantified Self participants who needed that capability to do research on themselves.

These changes set the stage for the Atlas of Caregiving project, a Robert Wood Johnson-sponsored pilot project led by the Family Caregiver Alliance to examine stress among family caregiving (Mehta and Nafus 2016). In this work, we asked caregivers to use a set of sensing gadgets that detected stress-related phenomena for a 24-hour period. We also asked them to keep a log of their activities, so that we could interview them about it afterwards and together identify what specific aspects of stress were causing the most trouble. As my collaborator Rajiv Mehta put it, the primary purpose was to 'see what we could see'. The sensors were being evaluated for their research value at the same time as we were trying to understand the substantive issue of stress among caregivers.

This was also an action research project, meaning that the goal was not merely to respond to the scholarship on family caregiving but to test whether collecting and examining data were directly useful to the caregivers involved, as is the case in the Quantified Self community. The people who participate in Quantified Self are different reflexive data subjects from family caregivers. Quantified Self participants tend to have elaborate ideas about what kinds of data collection, parsing methods and experimental design are going to yield the most insight, while family caregivers are usually preoccupied with other things. Still, their homes were not free of data, but contained elaborate spreadsheets, filing systems, medical data and the occasional Fitbit that kept things afloat. Both Mehta and I had been deeply involved in the Quantified Self community, and so we had a loose sense that, if caregivers were shown a new way of collecting data, they might stand to benefit in practical ways from reflecting on it.

In this study we spent a good deal of time evaluating newly available technologies for their suitability in research. The criteria we used for evaluation were the following (for further details, see Mehta and Nafus 2016).

¹*Signal*: What is the data claiming to indicate, and is that claim credible? In our case, we were looking for indicators of stress, but there are no direct indicators of stress to be had. Instead, there are proxies, like heart rate, electrodermal activity (EDA) and sleep disturbances. These only partially indicate some types of stress, some of the time, and also indicate a whole lot else like exercise or a humid room. We also examined whether the device is prone to returning missing data, or implausible data such as spikes in heart rate reaching 170.

Sampling rate: How often the device takes a reading (known as sampling rate) can affect what a researcher can do with the data afterwards. A device that takes a reading every second can tell a researcher about some types of physiological responses that one taking a reading every minute cannot. However, a file that takes a reading once every second is also 60 times larger than a file that takes a reading once a minute, and that impacts in the requirements for storing the data, the ease of moving it between collaborators and ease of working with it. Fine-grained data might or might not be necessary, depending on the researchers' conceptualisation of how long the phenomenon in question plays out, and the participants' ability to recall the details of what was happening (which in our case was not at the second-by-second level).

Parsers: Sometimes data can be parsed in different ways to extract different kinds of indicators. Activity trackers regularly parse movement data into steps or sleep, for example. Parsing algorithms like these are rarely available off the shelf and not easily deployed. In our case, we contacted stress-sensing researchers who allowed us to use a parser that would convert EDA data into more sophisticated signals that they believed are more likely to be tied to stress.

Battery and data storage: As with the initial 2007 study, devices do not always behave in the way that one would hope. Batteries might need frequent recharging, or the device might fill up with data and need to move it off the device and into long-term storage more quickly than a research participant can get to a connection. Some devices successfully sync better than others, which impacts the level of intervention necessary from participants or researchers.

Burden on participants: A device that was particularly unsightly, physically hindering or inappropriately gendered might be difficult for participants to use for any sustained length of time. Technology babysitting, like syncing data from device to online storage, might require unavoidable work from the participant. As our study also involved participants keeping a log of their activities, which was a significant burden, we also limited electronic data collection to a 24-hour period specifically so that no one would have to recharge a battery or deal with syncing issues on top of keeping a log, and that on top of caregiving.

Compatibility with participants' systems: Our device-vetting process revealed that many devices and mobile phone apps are specific to a

particular operating system, or otherwise depend on what participants have access to. Setting up an in-home sensor might require high-quality broadband access that participants don't always have. In the caregiving study, we did not load software on to people's devices largely to avoid this problem, but that choice meant we could record only for a short amount of time, and record one participant at a time. In subsequent projects where the research questions demanded a longer sensing period, we had to make sure to use devices that could work with both Apple and Android systems, and do more technical support to ensure they were successfully loaded on to the participants' phones.

Data export and data format: Not all devices provide the ability to export data. Some popular devices offer export of daily summaries, but require customers to pay more or have a developer account to get the full dataset available (software like Data Sense and Zenobase provides a way to work around this problem). Many potentially viable devices were rejected for this reason. Also, not all data types are interoperable. We did not fuse together numerical data, handwritten logs and images taken once every minute into an interactive visualisation because this would have required building bespoke software. Similarly, timestamps can be inconsistently formulated across different devices, making it tricky to compare across different data. Some data analysis tools will help with the timestamp alignment, and someone more technically minded can overcome the time-formatting issues in a spreadsheet program. In the end, we showed participants each data type (noise level, heart rate etc.) separately in part to avoid these issues.

Data giveback: Data giveback can be as simple as sending participants spreadsheets, but this is not likely to be particularly useful. Many consumer-oriented sensors provide interfaces that research participants can use to peer into their data, but sensors designed for research use might not provide this. It can also be a source of frustration if the interfaces are poorly designed or inappropriate to the context (for example, if the device 'nudges' the user into exercising more when the data is being collected to indicate stress). It might be sufficient to provide participants visualisations produced by the project.

Because we were sensing a 24-hour period, and people quickly forget the details of a single day, our turnaround process was much tighter than in the computer-use study. We held an initial interview where we also set up the gadgets, revisited the participants the next

day to collect the gadgets, and then Rajiv immediately produced plots in Tableau which we used the following day for interviewing. Because we had selected a research-grade sensor, we did have to write a script to be able to plot the data, but this would not have been necessary with a consumer-oriented device. Still, it also required learning to use new software (Tableau) and developing an overall strategy for the types of plots that would be necessary. Here, we focused on simple area graphs for each of the numerical datasets, one that contained data for the full duration of study and another that zoomed in to times that appeared notably stressful.

As with previous studies, what the data evoked proved much more ethnographically productive than the stories it told directly. Because our set-up involved multiple modes of sensing, the discrepancies between the different accounts often proved interesting. In the interviews, images from body-worn cameras and handwritten logs helped participants recall the events of the previous day. The logs helped us understand how participants framed those events, and indicated their subjectively experienced stress. One particularly telling interview shows how these modalities came together. Nadine's son, a Type 1 diabetic, had a continuous glucose monitor, which sent her alarms on her phone when his glucose was out of range. Sometimes it indicates a real problem, and sometimes the alarms were spurious.

On our second interview Nadine explained how the alarm woke her up multiple times in the night, and early in the morning her son had an infection she had to take care of. Later that morning the glucose monitor was going off again, and she was desperately trying to get hold of her son and her husband, both of whom could resolve the problem. So far, there are three stressors in the account: a medical situation, the difficulty reaching the people who can resolve it and the need to respond to the situation on very little sleep. This incident was playing out while she was in a meeting with one of her volunteer organisations. She said, 'So in the school meeting it was really the worst, and I felt so bad because I'm constantly texting and there's this counsellor sitting with a direct view of me ...' So now we have a fourth stressor, which is the social stigma of appearing to not be properly paying attention at a meeting.

We then looked at the pictures together. She had written down the incident as happening at 8:15 am, which was the start time of the meeting, and had marked it as having very high subjectively felt stress

levels. Her heart rate data appeared to show a fairly large surge just after 8:15, and so at first we thought we had a small discrepancy in time (8:15, as she had remembered it, versus just after 8:15 when the heart rate spiked). When we looked at the camera data for 8:15, though, we saw her not in the meeting but still bicycling to the meeting. The surge lasted for the duration of the incident as she describes it, which means that the sensors were picking up on some biological manifestation of what she was experiencing. But the minor discrepancies in the starting time suggested one additional important layer to the stress. When we showed her the camera data from that time period, she noted 'I think it was probably really that I thought, "Oh, we have this meeting afterwards. But now I need to concentrate on this."' She told us about how she was worried about letting down someone she would encounter in the next meeting, which to us spoke to the tremendous cascade of responsibilities she was carrying, and how they were interwoven temporally.

A positivist would be inclined to use the story to underscore the poor recall capabilities of human beings. We used it instead to think about how stress might be a layered phenomenon, where the first version of the story is not a lie or a mistake but the part of the story remembered as worth telling. Subsequent versions of the story, mediated through different resources, showed that a cascade of responsibilities were working together. The 'full story' is not what made it interesting; the fact that it came in layers, with different parts remembered in different ways, is what gave it meaning. We came to suspect that this might be saying something about how stress builds up, and becomes culturally and socially inflected. These nuances became visible only by making room for, and valuing, human 'error'.

Real Time Health Monitoring

In this final section, I wanted to add some thoughts about this approach for community-based participatory research (CBPR). In the caregiving pilot, there was an aspiration that the research should be useful for at least some of the participants involved. In this last project, Real Time Health Monitoring, a coalition of environmental justice groups had been looking for research that could identify the health effects of air pollution from a point source polluter in their area. Being familiar with how members of the Quantified Self community answer

questions like this, and, by this time, having a prototype of software that could ease some of the data-wrangling challenges the problem posed, I became interested in exploring what it would take to help these groups answer their question. I collaborated with Gwen Ottinger, an STS scholar who had a longer history with these groups as they advocated for high-grade air quality sensors to be put in place, and Randy Sargent, a computer scientist at Carnegie Mellon University's CREATE Lab, which had created an important infrastructure for making air quality data accessible (esdr.cmucreatelab.org). The installation of air quality sensors was itself an important victory, but the question quickly became how to make good use of them. Gwen and I worked with the coalition to enable them to set up a small, participatory pilot study that equipped nine volunteers to with various health-related devices to report symptoms. If these tended to occur at times when there was a spike in pollution, it could indicate something significant to them.

Here we followed the same device-vetting procedures as in the previous study, provided the high levels of technology support for participants that we had learned would be necessary and continued the process of meeting with people individually to explore what the data meant to them. In those individual meetings, I did not prepare visualisations ahead of time, as Data Sense enabled me to try out various visualisations interactively in the interview itself, and annotate them. This reduced substantially the amount of data preparation necessary beforehand, but limits the study to the patterns that Data Sense handles well. (There are also now other similar tools available, such as databasic.io, Fluxstream and Zenobase, which offer analogous improvements and tradeoffs.) This was a new evolution in my research practice, as for the first time it was up to me to handle the data movement from the device to the place where it would be analysed. While there were some hiccups, including one of the data collecting apps going out of business entirely during data collection, in general I found this now feasible to do without specialist skills.

However, the environmental justice groups were acting as a community, not as individuals *per se*. That added additional layers of complication. In the caregiving study, we could analyse data on an individual-by-individual basis, but here the participants themselves were asking questions about what was shared between them. Specifically, they were looking for correlations with the pollution levels that

they shared physically. Data Sense was designed for individual use by Quantified Self (QS) community members, and so it creates grids of pairwise correlations for a small handful of data referring to a single individual. QS participants were largely wary of correlations across a population, and so Data Sense did not contain a notion of a population, even though here it was central.¹ What it did have was a notion of data sharing and exchange, and so, to hack this feature for community use, we created a series of grids by sharing data into one account. Each grid showed one pollutant and as many participants as we could fit in. These showed whether the group was correlating with the pollutant while exposing the individual variation. Some participants were more strongly affected than others. This attempt at numerical commensuration without a population-wide measurement would likely make epidemiologists unhappy, but it suited the purpose of discussing with the group whether there was a possible relationship worth further research. It invited conversation about whether pollution was affecting those with chronic health problems more severely, or people living closer to or further from the main polluter. Seeing the range of individual responses invited people to imagine how they were a part of the whole in a new way.

While this project did convince me that it was feasible to do this kind of data work without the aid of a data scientist, and, in the future, by citizen groups themselves with a bit of training, 'empowerment' also comes with its costs. I anticipated the heightened scrutiny, and pre-emptive discounting, any further calculations would receive because they were made by newcomers to these forms of mathematics, primarily by women, and by people who were seeking the evidence base they needed to advocate for change. The increased level of mathematical complexity also created more room to get it wrong for the wrong reasons. For example, in the course of working with the data, we discovered that some of our data had numerical qualities (non-normal distribution) that Data Sense was not optimised to handle. It took the intervention of someone more fluent in statistics to spot it, and figure out a workaround. While we put ourselves in a position to be able to defy genres of big data calculation and expose the limitations of epidemiology – to get it 'wrong' for the right reasons – we also risked stumbling into problems that others have indeed thought more deeply about, like non-normal distribution, that, if not caught, could be used by others to undermine the whole thing.

Unfriendly scrutiny is always just around the corner for these groups. At some point my inventive method – the use of consumer devices to make non-standard data compositions that suited a human lifeworld if not expert-driven health research – becomes not just an interesting, novel research approach but a matter of responsibility to others who would also have to live with this. These groups will have to decide whether to use calculations like these to make claims through idioms of science or idioms of storytelling, and my method sat precisely in the middle. The middle might be ethnographically rich, but, in some circumstances, dangerous waters. These are real pollutants, and real fights about bodily harms that are not helped by uncertainty. We swam in these waters nevertheless on the faith that doing so meant that far-away experts would no longer monopolise the tools of knowledge production. In that sense, perhaps we were not just 'seeing what we can see' but learning *how* to see, rooted in this particular social world that included data, software, hardware, people and landscape.

Conclusion

When I began working on Data Sense, I did not expect to end up hacking on it as a matter of methodological necessity. The trajectory, however, shows how the social life of this ethnographic method is caught up in the evolution of the materials at hand, and does not leave untouched the worlds it is designed to comprehend. My story has been one of decreasing reliance on collaborators to do the data wrangling, which is not a claim that collaboration is problematic, but a claim about the capacities that open up when anthropologists take the social life of methods seriously, not just as the outcome of research practice but as part of the research itself. This methodological evolution shows, from a particular point of view, the changing material circumstances as they have unfolded across the last ten years. Indeed, the changing materials available are not some exogenous force but something anthropologists can and should actively participate in. When we approach data systems as something that is a part of the social worlds in which we work, and not merely the latest flight of fancy of funding bodies dazzled by the magical power of computation, we tap in to longer-standing anthropological traditions of deep engagement with the social worlds we write about. If data is a

contemporary form of storytelling, then anthropologists are only just beginning to take the leap in telling those stories with the people we work with. It might be easier to do this with tools that we also have a hand in creating.

It is not difficult to imagine how it is that our research participants tell their stories with data, yet somehow we struggle to imagine doing so for our own. Stories with numbers are not stories we usually tell. Self-trackers and environmental justice advocates regularly receive criticism from those who they challenge about how they are not scientific enough, use poor experiment design, etc. These accusations are cut from the same cloth as the ones that anthropologists have endured for a very long time. They are born from a preoccupation with generality that both anthropologists and these groups find untenable. We can be clear-headed about the epistemological politics of large data systems and nevertheless see how supposedly dubious numbers, coming from such suspect sources as non-professionals and other researchers, might yet become powerful in unexpected ways.

Note

1 The Data Sense team, at time of writing, has since explored the design requirements that would be necessary to create appropriate cross-population aggregates.

References

- Aipperspach, R., Rattenbury, T., Woodruff, A. and Canny, J. 2006. 'A Quantitative Method for Revealing and Comparing Places in the Home'. In *International Conference on Ubiquitous Computing Proceedings*. Berlin and Heidelberg: Springer, 1–18.
- Anderson, K., Nafus, D., Rattenbury, T. and Aipperspach, R. 2009. 'Numbers Have Qualities Too: Experiences with Ethno-Mining'. In *Ethnographic Praxis in Industry Conference Proceedings 1*. New York: Blackwell Publishing Ltd, 123–40.
- Andrejevic, M. 2014. 'The Big Data Divide'. *International Journal of Communication* 8: 17–36.
- Corburn, J. 2005. *Sweet Science: Community Knowledge and Environmental Health Justice*. Cambridge, MA: MIT Press.
- Couldry, N. and Powell, A. 2014. 'Big Data from the Bottom Up'. *Big Data & Society* 1(2): 1–5.

- Geiger, R. S. and Ribes, D. 2011. 'Trace Ethnography: Following Coordination through Documentary Practices'. In *System Sciences (HICSS)*, 2011 44th Hawaii International Conference on System Science. Piscataway, NJ: Institute of Electric and Electronics Engineers (IEEE), 1–10.
- Greenfield, D. 2016. 'Deep Data: Notes on the N of 1'. In *Quantified: Biosensing Technology in Everyday Life*. Edited by D. Nafus. Cambridge, MA: MIT Press, 123–67.
- Kitchin, R. 2014. 'Big Data, New Epistemologies and Paradigm Shifts'. *Big Data & Society* 1(1): 1–12.
- Latour, B. 2004. 'Why Has Critique Run Out of Steam? From Matters of Fact to Matters of Concern'. *Critical Inquiry* 30(2): 225–48.
- Law, J., Ruppert, E. and Savage, M. 2011. 'The Double Social Life of Methods'. CRESC Working Paper Series, 95. <http://research.gold.ac.uk/7987/1/The%20Double%20Social%20Life%20of%20Methods%20CRESC%20Working%20Paper%2095.pdf> [accessed 4 March 2018].
- Lury, C. and Wakeford, N. (eds). 2012. *Inventive Methods: The Happening of the Social*. London: Routledge.
- Mehta, R. and Nafus, D. 2016. 'Atlas of Caregiving Pilot Study Report'. <https://atlasofcaregiving.com/core-research/pilot-study> [accessed 16 June 2017].
- Nafus, D. 2016. 'Introduction'. In D. Nafus (ed.), *Quantified: Biosensing in Everyday Life*. Cambridge, MA: MIT Press.
- Nafus, D. and Beckwith, R. 2016. 'Number in Craft Situated Numbering Practices in Do-It-Yourself Sensor Systems'. In *Critical Craft: Technology, Globalization, and Capitalism*. Edited by C. Wilkinson-Weber and A. DeNicola. London: Bloomsbury Press, 115–34.
- Pasquale, F. 2015. *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, MA: Harvard University Press.
- Ruppert, E., Law, J. and Savage, M. 2013. 'Reassembling Social Science Methods: The Challenge of Digital Devices'. *Theory, Culture & Society, Special Issue on 'The Social Life of Methods'* 30(4): 22–46.
- Savage, M. 2013. 'The "social life of methods": A Critical Introduction'. *Theory, Culture & Society* 30(4): 3–21.
- Wilmott, C. 2016. 'Small Moments in Spatial Big Data: Calculability, Authority and Interoperability in Everyday Mobile Mapping'. *Big Data & Society* 3(2): 1–16.