The Quantified Self

A Sociology of Self-Tracking

Deborah Lupton

polity

are generally represented as 'small' and human-made, wrought from the personalised decisions and individual objectives of the people who gather them. Yet, if these data are generated by digital devices, they are often aggregated into big data sets and become part of the digital data economy. This raises issues about data politics, security and privacy in terms of the ways in which people's personal details are accessed by other parties. These issues are the focus of the next chapter.

5

'Data's Capacity for Betrayal'

Personal Data Politics

In the previous chapter I discussed the ways in which self-trackers seek to make sense of, materialise and use their personal information. Beyond these reflexive data practices, some self-trackers confront the next level of data use: where and how their personal data are stored, how they are harvested by other actors, what these actors do with their data and how they can gain better access to them. This chapter addresses these political dimensions of personal data.

Exploited self-tracking

Several years ago, when digital technologies were beginning to be used for self-tracking, Dodge and Kitchin (Dodge and Kitchin, 2007; Kitchin and Dodge, 2011) raised some important questions about the data that are produced through lifelogging practices. Here are some of their questions: Who (other than the creator) should have access to the data archives that are preserved in a lifelog? Should other people, whose data may be included in an individual's data archives, have access to some or all of the data contained in those archives (for example, images of them or details about them?) To what extent could the material be sequestered for legal cases? To what extent would deletion of data or suspension of data

gathering from a lifelog be considered a sign of guilt if the lifelog were to be used in a legal case? Could other actors insert false information into a person's lifelog, thus creating false memories? What happens to lifelog data after the death of the creator? What are the inheritance rights? How much more valid than human memories are these data to be considered? How long will lifelogs remain an act of choice and free will – will their collection become mandatory and be imposed by (some) authorities? Should portions of a lifelog be available for erasing or modifying? What details should be preserved? Is there a need to forget misfortunes and errors? What happens if one's lifelog data are stolen and used by others? Who has control over a child's lifelog?

An important implication of automated digital recording of a greater amount of personal information is that such technologies lack the power to discriminate. They simply continue to record details, leaving no sign or mark of what is important, which details should be preserved and which could be relinquished (Kitchin and Dodge, 2011). Dodge and Kitchin (2007: 439) contend that lifelogs have the potential to fulfil a 'marketer's dream' – if that marketer is able to get access to the wealth of personal details in a lifelog, including the self-tracker's purchasing and consumption habits. The two authors envisage incidents in which third parties might use this information for social sorting, invasive profiling and disciplining. They raise the possibility of insurance companies and other commercial entities requiring access to lifelog data for the benefit of calculating risks and premiums or for the purpose of according preferential treatment to some customers – while others, who fit certain profiles, would be penalised. Dodge and Kitchin also identify the possibility that society would become more conservative once people are aware that their personal information can be accessed by others and used against them, thus making public forbidden, indiscreet or criminal behaviours.

Dodge and Kitchin were writing before the widespread use of cloud computing, the growth in the collection and use of personal data by internet companies such as Facebook, Amazon and Google and the spreading of self-tracking practices beyond the realm of the private and the consensual. The

uses of the personal data that people have generated through self-tracking – that is, the uses that Dodge and Kitchin envisaged several years ago – have largely eventuated as the two authors predicted.

I have referred throughout this book to the notion of 'lively data', as the feature involved in this notion participates in the digital data economy. As I have argued, one dimension of the vitality of digital data relates to the multiple ways in which different actors and agencies may use them for their own purposes. This multiplicity has major implications for how the information that people collect about themselves, as part of self-tracking endeavours, is used (or indeed misused) by others. The exploitation of people's personal information by second and third parties is a significant political issue, not simply because of the data privacy and security issues involved but also because of the ways in which people's personal information has become valuable for these parties. The collection of personal data is now not only a mode of consensual, individually driven imperatives for self-improvement, but also an element of (sometimes illegal) commercial profiteering, population monitoring and governance.

Indeed one might view the knowledges that are created through self-tracking practices as a new element of biopower or vitality expertise. The movement of self-tracking cultures into commercial, managerial and government domains combines the rationalities of biocapital with those of the digital data economy. Biocapital involves the derivation of value from biological entities such as human bodies (Rose, 2008), while the digital data economy positions digital data objects as valuable. Just as other forms of human life – such as human gametes, blood, tissues and cells – have become commodified and invested with monetary value, so too have the digital data assemblages that are configured on human bodies via self-tracking. Indeed the value attributed to personal digital data assemblages combines two kinds of value: one related to the digital data economy and one emerging from the capitalisation of the human body. Many self-tracking practices involve the rendering of bodily attributes and dispositions into digital data. They produce value in terms of the intimate biodigital knowledges that they generate from

individuals, and therefore self-tracking practices may be described as generating digital biocapital.

The creation of digital content – that is, prosumption on online platforms and apps – can be viewed as a form of work. Indeed some scholars have represented prosumption in general as free digital labour, in which people who generate these data do so for the commercial benefit of other actors and agencies. Their labour is exploited because, while they may benefit personally from their acts of prosumption (for example, by enjoying free access to platforms and apps and opportunities to interact with others, use the information provided there, or monitor their bodies and behaviours closely), others are profiting financially from this freely given content (Fuchs and Dyer-Witheford, 2013; Rey, 2012; Till, 2014). People are not offered financial compensation, nor do they receive it for providing their experiences. The value that prosumers derive is noncommercial, while the exchange value of the data they create is accumulated by the for-profit companies that provide the platforms for people to share their experiences or to trawl the web, harvest the data and render it into a form that is valuable for commercial entities.

The exploitation of prosumers' personal information frequently occurs when people use apps and other software for self-tracking. Many commercial companies are interested in the type of details about health, physical activities and consumption patterns that are revealed by the information collected by self-trackers on their bodies and lives (Till, 2014). For example, when people engage in user experience platforms such as PatientsLikeMe, they are encouraged to share the information they have collected about their bodies, medical conditions and treatments with other patients with the same condition. These data are valuable not only to other patients, for the insights they offer them, but also to the platform developers, who on-sell these data, and to other third parties, who use them for research into medical conditions, for clinical trials of new pharmaceuticals, or for purely commercial purposes – as do medical device manufacturers and pharmaceutical companies, for instance (Lupton, 2014).

The burgeoning business of data harvesting and data brokering involves a process whereby companies are scraping the

web for whatever they can find about people; in other words it involves the sale of the data that have been generated through the use of apps and other software. Data-harvesting and brokering companies use the information they can find online or have bought from developers in order to construct 'profiles' that provide detailed descriptions of the behaviours and health states of the people profiled. Drawing on this information, some companies create lists of people who have been sexually assaulted, diagnosed with a mental health condition or a sexually transmitted disease, designated as impulse buyers or credit risks, or accused of wrongdoing. These lists are sold to marketers, financial institutions and potential employers (Pasquale, 2014).

The advent of big data, together with the opportunity to mine them for personal information, has created new possibilities for social and economic discrimination against the disadvantaged and against minority social groups. Here one could mention the potential for predictive privacy harms, which covers cases where individuals are adversely affected by assumptions and predictions made about them on the basis of preexisting digital data sets (Crawford and Schultz, 2014; Robinson, Yu, and Rieke, 2014). Predictive algorithms that draw on personal digital data are used now in many social and economic domains. This new practice can affect people's access to healthcare, credit, insurance, social security, educational institutions and employment options and render them vulnerable to unfair targeting by policing and security agencies (Crawford and Schultz, 2014; Nuffield Council on Bioethics, 2015; Rosenblat, Kneese, and boyd, 2014). What is more, it can be difficult to challenge such assessments or to seek to have certain personal details removed from digital data sets, even if the data on which they are based are proven to be inaccurate.

Some employers have begun to use the algorithms of specially designed automated software for the purpose of selecting employees; they are also engaging in online searches through search engines or professional networking platforms such as LinkedIn in order to seek out information on job applicants (Rosenblat, Kneese et al., 2014). Now that diverse databases holding personal details on various aspects of people's lives can be joined together for analysis,

information on features such as a job applicant's health status or sexual orientation may become identifiable (Andrejevic, 2014). One recent study found that Google directs fewer higher-paid job advertisements to female than to male users in search of employment sites, in a clear case of algorithmic discrimination based on gender (Datta, Tschantz, and Datta, 2015).

Insurance and credit companies are scraping big data sets to develop customer profiles, with the result that disadvantaged groups suffer further disadvantage by being targeted for differential offers or excluded altogether because they are not viewed as profitable or as poor credit risks (Libert, 2014; Robinson et al., 2014). Data brokers in the United States use available personal data to calculate certain predictive 'health scores' on patients with the help of digital data; such scores include the Affordable Care Act (ACA) individual health risk score, which is used for assessing the risk factor for an individual who requires healthcare (Sarasohn-Kahn, 2014). Some American hospitals are purchasing from data-broking companies data on their patients' credit card transactions and information about them in public records and and in customer loyalty programs, in an attempt to use predictive algorithms for creating models that identify 'high-risk' patients. These patients will then be contacted by the hospital as part of an intervention program that seeks to prevent ill health and reduce healthcare admissions (Pettypiece and Robertson, 2014).

The legal implications of the use of personal data archives for evidence are just beginning to emerge. In 2014 the first known case where an individual's self-tracking data (collected by her Fitbit physical activity tracker) were used as legal evidence in a personal injury lawsuit received media attention. A Canadian fitness instructor sought to use her physical activity data, collected by her Fitbit, to demonstrate reduction in her activity after an injury. Her lawyers used the data analytics platform Vivametrica to compare this woman's physical activity data with those of the general population. Commentators on this case speculated that similar self-tracked personal data could be used in the future not only to support people's lawsuits, but also as evidence to prosecute them in litigation (Olson, 2014a).

Pushed and imposed self-tracking

The growing adoption, by actors and agencies, of self-tracking practices and rationales beyond the realm of the consensual and the personal raises questions about the extent to which people are now being pushed or even coerced into taking up self-tracking. Advocates who encourage people to take up self-tracking are particularly visible in the domains of patient self-care, health promotion, preventive medicine and health insurance. In the persuasive computing and digital health literature, the personal data that are generated from self-tracking are represented as pedagogical and motivational – a means of encouraging self-reflection or emotional responses such as fear, guilt or shame that will lead to the desired behavioural changes. While many people may choose to engage in these types of enterprises willingly, as part of their personal goals and motivations, there is abundant evidence in these programs that they are strongly associated with the objective of persuading people who are otherwise reluctant to participate in them. Hence the motivation for self-tracking is viewed as requiring impetus from the external agency that is attempting to change people's behaviour.

Such a perspective on encouraging self-tracking draws on traditional paternalistic approaches to health promotion and health education, in which lay people are positioned as ignorant or lacking motivation and self-control (Crawshaw, 2012; Lupton, 1995b; Petersen and Lupton, 1996). The recent interpretation of this paternalism as 'nudging' (Thaler and Sustein, 2009) adopts an explanatory framework that attempts to preserve a veneer of choice and voluntary behaviour change by making paternalism seem to appeal to strategies that subtly encourage such change. Nudges are designed so that they are not readily obvious to their target groups, or they appear to be easy to respond to without great deliberation or motivation; thus they are viewed as consensual rather than imposed. They may be deceptive or manipulative in the way they achieve their ends. This a type of 'soft' or 'libertarian' paternalism that adheres to the neoliberal model of governing populations, in which coercion is largely replaced by psychological models of behaviour that encourage people to

take up self-care practices for their own health, happiness and productivity. At its heart is the belief that, left to themselves, people would not readily take up behaviours deemed to be wise, productive and conducive to the ideal of the responsible entrepreneurial citizen; hence they must be 'encouraged' to do so by other actors and agencies.

Some writers in the field of persuasive computing and in that of nudging design are beginning to discuss the possibility of developing wearable technologies or smart objects that not only monitor people's bodies or interactions but actively intervene to discipline them. An example of such an object would be a desk lamp that turns on only when a smart phone has been placed inside it, to discourage overuse of the phone. It has been suggested that future designs may include a smart sofa that can kick people off it if they have been lounging for too long, or a smart watch that informs users that they should walk to work rather than catching the train and then urges them to walk faster if they fail to demonstrate enough enthusiasm (Peters, 2015). The Apple Watch already notices how often wearers stand and move around, and sends them notifications if they are deemed by its algorithms and sensors to have been sedentary for too long.

More obvious forms of pushing self-tracking on people are appearing in the domain of insurance. Drawing on the possibilities of self-tracking technologies, insurance companies are beginning to adopt the usage-based insurance model, which is predicated on the fact that people provide individualised information to insurers for the calculation of risks and subsequent premiums. This approach to insurance moves from actuarial calculations of risk that are based on aggregated historical data to risk assessments that focus on the individual's characteristics, as derived from a long list of variables (NAIC, 2014). As I observed in Chapter 1, some car insurance companies use telematic driving-monitoring technologies to calculate their clients' risk profiles and premiums. Health and life insurance companies are also beginning to encourage their clients to upload their self-tracking health and fitness data. For example, the insurance company AIA (Acts Interpretation Act) Australia offers a Vitality life insurance program in which, as its website puts it, 'your healthy choices are financially rewarded'. Its

clients are encouraged to engage in an array of preventive health, monitoring, testing and screening programs to earn points that will then reduce their premiums. These are divided into 'know your health' and 'improve your health' activities. The 'know your health' activities include completing online tools to calculate aspects of overall health status and mental wellbeing, completing a non-smoker's declaration and seeking health, nutrition, fitness and dental assessments from providers. The 'improve your health' activities involve attending gym or fitness sessions, engaging in 'stop smoking' or weight loss programs, ordering fresh food online, and wearing digital activity wearable devices and uploading the data to the company. Each time they perform these activities, clients earn points that are then used to reduce their premiums.

Other agencies, such as retailers that offer customer loyalty programs, are encouraging their clients to allow them access not only to purchasing behaviours displayed in supermarkets and pharmacies but also to self-tracked health and fitness data, which allows them to combine various forms of data so as to make inferences about their customers' health-related habits and preferences. The Balance Rewards for Healthy Choices program is offered by Walgreens, America's largest pharmacy retailing chain. As part of a customer loyalty program, people are offered the opportunity to 'earn points for your healthy choices', to save money on products, and to 'take advantage of great, exclusive offers for members' (Walgreens, 2014). They can do so by first recording details of their physical activity, chronic disease management or progress towards a health-related goal such as losing weight or ceasing smoking and then syncing the data collected via digital fitness trackers or uploading them onto Walgreens' platform or customised app.

The Australian Coles supermarket chain has a customer loyalty program that incorporates collecting not only information on their members' spending habits in the supermarkets and liquor stores owned by the company but also health and fitness data on them from digital self-tracking devices. The company offers life insurance, and is also associated with a major private health-insurance company that offers benefits to insured clients who regularly upload health and fitness

data onto their platform. It is not difficult to envisage a scenario in which data concerning food, cigarette and alcohol purchases and health and medical information are brought together, used to make predictions about consumers, and result in a differential targeting and pricing of insurance packages.

Corporate wellness programs in the American workplace represent an instance where the boundaries between voluntary self-tracking and pushed, or even imposed self-tracking can be blurred. As discussed in Chapter 1, in the United States many employers take responsibility for securing a proportion of their employees' health-insurance coverage; they do this as part of a benefit package, in the absence of nationalised public healthcare systems such as those offered in other western countries. For this reason employers have a financial interest in promoting wellness programs among their staff members in addition to attempting to reduce absenteeism and subsequent productivity loss due to illness. The Affordable Care Act allows American employers to provide financial incentives for their staff members' participation in workplace health-promotion initiatives and demonstrations of progress towards attaining personal health goals – namely incentives in the form of payments of up to 30 per cent of these members' health-insurance premiums (Zamosky, 2014). Wearable technology manufacturers such as Fitbit are brokering deals with employers and insurance companies to sell fitness and activity trackers and data analytics software as part of these wellness programs (Olson, 2014b; Zamosky, 2014).

There is a fine line between consensual, pushed and imposed self-tracking. While some elements of self-interest may still operate and a discourse of 'choice' may be employed, people may have little option of opting out. In the case of workplace wellness programs involving the self-tracking of physical activity or body weight, for instance, wearing the devices and allowing employers to view employees' personal data may be presented as optional. However, failure to participate in the program may lead the enforcement of higher health-insurance premiums by an employer, as is happening in some American workplaces (Olson, 2014b). At its most coercive, imposed self-tracking is used in programs involving the monitoring of location and drug use for probation and

parole surveillance, drug-addiction programs, and family law and child-custody monitoring.

Personal data security and privacy

There are many significant issues concerning the security and privacy of the personal information that self-trackers upload to apps and other software. Developers often fail to inform users that their data are available to third parties (Ackerman, 2013; Sarasohn-Kahn, 2014). In the United States, where many internationally popular apps are developed, there are no legal requirements that app developers provide privacy policy statements in their information for users. A recent study of privacy policies on mobile health- and fitness-related apps found that many lacked any kind of privacy policy, few took steps to encrypt the data collected, and many sent such data to a third party not disclosed by the developer on its website (Ackerman, 2013).

The US Federal Trade Commission found that 12 free health and fitness apps focusing on relevant behaviours or on conditions such as smoking cessation, physical activity and pregnancy shared user data with a total of 76 third parties. These data in some cases included geolocation, gender, names and email addresses, exercise and diet habits and medical symptom searches (Kaye, 2014). A study of over eighty thousand health-related web pages found that 90 per cent of them leaked user information to outside parties, including commercial data brokers (Libert, 2014). Sensitive medical conditions can become identifiable through the examination of other data sets, such as purchasing habits (Rosenblat, Wikelius, boyd, Pena Gangadharan, and Yu, 2014). Several researchers have demonstrated how easy it is to de-anonymise digital data about individuals using a small amount of additional information, often on the basis of patterns of behaviour or joined-up data sets that can then re-identify people (Singer, 2015).

Personal medical details are also very valuable to cybercriminals. It has been estimated that the digital data black market is now more profitable than the illicit drug industry (Ablon, Libicki, and Golay, 2015). Data security is becoming

increasingly more difficult to protect as 'smart' online objects connect with each other and share data, and as personal data are uploaded to cloud computing archives in increasingly large amounts (Barcena, Wueest, and Lau, 2014; Kitchin, 2014). Hackers can gain access to personal data at two key points: when these are being transmitted from one location to another, such as from a personal device to a cloud computing database; and when they are kept in databases (Barcena et al., 2014). If strong data encryption and authentication protocols are not employed, hackers are able to gain access to personal data more readily.

Cybercriminals are frequently targeting the American healthcare system for illegal access to details such as names of patients, diagnosis codes and health-insurance policy numbers. They then use these details to gain access to pharmaceuticals, to make fraudulent health-insurance claims or to sell the data themselves in the black market (Humer and Finkle, 2014). Hackers have already accessed the types of information that workplaces often request their employees to provide as part of wellness programs or health-insurance plans – for instance information on sexual activity, stress levels and mental health, drug consumption, preexisting medical conditions and blood-test information (Pettypiece, 2014). Private health information details have been subject to numerous privacy breaches. Since 2009 over one thousand incidents have been reported to the US Department of Health and Human Services, all related to the hacking of digitised health information that should have been protected by the Health Insurance Portability and Accountability Act (Pettypiece, 2014).

Many internet and mobile technology users face difficulties in understanding or accessing the terms and conditions of the software and hardware that they use (Nissenbaum, 2011; Rosenzweig, 2012; Tene and Polonetsky, 2013). Some self-trackers may be unconcerned that their personal information is being used for profit or managerial purposes by others, or may view this as a trade-off designed to secure their ability to use various devices or software. Sometimes users agree to the use of their personal data by third parties as an unavoidable part of accepting the terms and conditions of devices, apps and platforms, or customer loyalty schemes (although

to what extent users actually read through the fine print on these documents is not known). In other cases the users' data may be accessed for the purposes of others without the users' knowledge or consent. However, in the wake of the publicity stirred around Edward Snowden's revelations about governments' surveillance of their citizens and extensive news coverage of the ways in which big data are being harvested for commercial purposes or illegally accessed by hackers, people are becoming more aware of how often they are digitally monitored by others. There is a growing sense that individuals are being placed under dataveillance without their knowledge or express consent (Crawford and Schultz, 2014; Hartzog and Selinger, 2013; Polonetsky and Tene, 2013; Wellcome Trust, 2013).

The mass media are replete with such statements as 'Google/Facebook/Amazon knows you better than you know yourself'. The argument is that the internet empires' capacity to collect routine transactional data on users and to apply their algorithms so as to interpret and predict their habits and preferences provides insights on features that users themselves may not have known they possessed. The implications for self-tracking practitioners have also been identified. For example, in an article for the technology website PandoDaily entitled 'You are your data: The scary future of the quantified self', the author speculates on the ways in which personal data may be used for surveillance by others – including credit card companies, insurers and employers:

> As we document and share more of where we go, what we do, who we spend time with, what we eat, what we buy, how hard we exert ourselves, and so on, we create more data that companies can and will use to evaluate our worthiness – or lack thereof – for their products, services, and opportunities. For those of us who don't measure up compared to the rest of the population, the outcome won't be pretty. (Carney, 2013)

The knowledge that the big data empires and security organisations appear to have about people often unsettle people (Wellcome Trust, 2013). Some find this apparent superior knowledge about themselves 'creepy' (Tene and Polonetsky, 2013). Many express powerlessness in the face of the

authority that internet empires have to collect, own and harvest their personal data (Andrejevic, 2014; Andrejevic and Burdon, 2015).

A study carried out by the Pew Research Center in late 2014 (Pew Research Center, 2014) found that the Americans they surveyed were displaying caution about how their personal online interactions and data were being monitored by security agencies and commercial entities. Their respondents were concerned about their personal data security. Nearly all of them were aware of the implications of Snowden's revelations about how the government was monitoring their private online communications and expressed the belief that people had lost control over how their digitised personal information was collected and used by companies. The people surveyed demonstrated a universal lack of confidence in the security of online communication channels and were highly aware of the difficulty of preserving anonymity on the internet. The respondents viewed their social security numbers as the most sensitive piece of personal information that they wished to protect, and this was followed by their health and medical information as the next most sensitive category.

A Wellcome Trust study that conducted qualitative research with British people similarly found that many participants viewed health- and medical-related information differently from other kinds of data. Participants saw the collection and sharing of their own data – their medical records – across healthcare sites in a positive light, as beneficial to their own healthcare. However, they were less sanguine about these private data being shared outside the NHS (National Health Service) system, and especially with employers and private companies that may seek to profit from the data (Wellcome Trust, 2013). In their British study, Dennison, Morrison, Conway, and Yardley (2013) found that several participants expressed concern about the security of the personal data they uploaded onto self-tracking apps and about the ways in which third parties might use this information. They were particularly sensitive about the possibility that details about their mental or physical health might be used by commercial entities that intended to target them with advertisements or might be broadcast on social media sites without their permission.

Nafus (2014: 217) uses the evocative phrase 'data's capacity for betrayal' when discussing the unintended consequences of engaging in sensor-based self-monitoring. Her participants were using home energy-monitoring systems. Some of them were concerned about the possibility that criminals may hack into the data and recognise when a home's inhabitants are out and may steal people's possessions, or that energy companies may use people's detailed energy use data for their own purposes. This sense of betrayal was also evident in another study – of Australian families that used home energy monitors (Snow, Buys, Roe, and Brereton, 2013). One participant in this project recounted an incident in which her husband had been examining their home's energy use from his digital device at work. Her own energy use had been noted and remarked upon by an onlooker who knew the couple and went so far as to telephone her to comment on her energy use. She was confronted by this loss of privacy. A teenage girl described how her parents could monitor when she was using the air conditioning at home by reviewing the energy monitoring system data; this discomforted her and made her feel under their surveillance. Such experiences reveal how self-monitoring can easily slide into surveillance by others, who could be members of one's own family.

In response to these issues, privacy and human rights organisations have begun to call for legislation and bills of rights that promote greater transparency in the ways in which big data are used by second and third parties. Critics have contended that a new 'digital divide' is emerging, in which powerful institutions and organisations such as the internet empires have control over digital data while others are excluded from access (Andrejevic, 2013, 2014; boyd and Crawford, 2012).

In February 2015 the Nuffield Council on Bioethics published a report on the ethics of the collection and use of data in medical research and healthcare that refers to the personal data gathered voluntarily by people as part of self-tracking practices (such data are referred to in the report as 'patient-generated data'). The report's authors are strongly in favour of better control over the security and privacy of such information – so much so that they discuss drawing up a legal framework for dealing with these issues and imposing

criminal penalties on the misuse of these types of data. They emphasise the importance of (1) developing ethical principles for the use of medical and healthcare data – principles that should be grounded in ideas of respect for persons, privacy and human rights; (2) incorporating the full range of values and interests of all actors involved; and (3) maintaining effective accountability in relation to data initiatives. Similarly, the Insight Ireland Centre for Data Analytics produced a white paper that set out a 'Magna Carta for big data' (Predict, 2015). The white paper's authors contend that the rights of all stakeholders – commercial bodies, the government and the public – need to be acknowledged by policy development. This entails protecting the privacy of the public appropriately while ensuring that government, research and commercial use of big data can still take place.

Apple's Tim Cook has taken a major stance by arguing that personal data and security are extremely important and should be protected. Apple's policy is that their product is the devices they sell, not the personal data that are generated by using the devices (Heath, 2015). For example, Apple announced in September 2014 that it was improving personal data encryption on its iPhones and iPads, following similar moves by Google and Yahoo. However, iPhone and iPad users are still encouraged to sign up to Apple's iCloud data-syncing and storage service, and the information and images that are stored there may be accessed by hackers or government security agencies. While these data on iCloud may also be encrypted by Apple, Apple uses its own password to encrypt them, and it may be forced to decrypt them at the government's request (M. Lee, 2014).

Communal self-tracking and taking control of personal data

What may be termed 'communal self-tracking' involves the consensual sharing of a tracker's personal data with other people, as a central feature of self-tracking practice. The people who take part in this process may use social media, platforms designed for comparing and sharing personal data,

and sites such as the Quantified Self website, in order to engage with, and learn from, other self-trackers. Some people attend meetups or conferences in a desire to meet face to face with other self-trackers and share their data and evaluations of the different techniques and devices for self-tracking.

The Quantified Self website often refers to participants as engaging in a community and encourages the sharing of personal data with one another. Indeed an emphasis on this process as part of the ethos of the quantified self has been evident since the earliest days of the Quantified Self movement. In his first article on the quantified self for *Wired* magazine, Gary Wolf (2009) asserted that self-tracking involves the sharing of data and collaboration on ways of using them, and therefore it is not a 'particularly individualistic' practice.

Self-trackers may share their data on the Quantified Self website or on other sites, on their own blogs or on social media sites such as Twitter, where the hashtag #quantifiedself is often employed to draw other self-trackers' attention to their posts. Some people choose to tell a very personal story, perhaps about how they used self-tracking in response to grief about the loss of a family member, or in response to their struggles with eating disorders, bowel problems or weight. As I noted in Chapter 4, this kind of sharing involves emotional disclosure to the group or online community. Others focus on how they use particular methods or devices and thus engage in a more technical exposition (Barta and Neff, 2014).

Notions of 'small data' and 'big data' are part of these discussions of how personal data may contribute to shared goals. There are various interpretations of what the term 'small data' means, which are inflected via the contexts in which the term is discussed. One definition that recurs in popular forums presents small data as information that individuals, organisations or businesses collect on themselves of their own will and for their own purposes. Small data are defined as personal and identifiable; big data as impersonal and anonymous. Small data are often represented as more contextual and easy to manage, because there are fewer data points. Information that is deliberately collected by someone for oneself, as part of self-tracking initiatives, is often represented as a form of small data.

Several commentators have begun to refer to 'the quantified us' as a way of articulating how the small data produced by self-trackers may be usefully incorporated into large data sets if one wants to 'get more meaning out of our data' (Ramirez, 2013). As one account of 'the quantified us' puts it:

> One of the ways we can transition the Quantified Self movement to have more impact, is to bridge the gap between Big and small data, and to heighten the collective relevance of the data we track about ourselves. By uncovering insights about ourselves through looking closely at others who are like us in the most meaningful ways, we can chart new paths toward becoming the people we want to be. (Jordan and Pfarr, 2014)

As this suggests, the concept of 'quantified us' still focuses firmly on the individual's agenda. The idea is to draw on others' pooled data to further one's own interests and goals: 'Quantified Self can provide added value, when you start sharing your data online and other self-trackers share their data as well. All this [*sic*] combined data provide an enormous amount of extra information for you' (de Groot, 2014). Therefore, while there is constant reference among members of the Quantified Self movement to the 'quantified-self community', this community largely refers to sharing personal data with one another or learning from others' data or from self-tracking or data visualisation methods, so that one's own data practices may be improved.

This perspective is also evident in the discourse of organisations such as the Small Data Lab, which are beginning to be established in order to provide software and assist people in harvesting their own data so that they can access 'the big insights and meaning this small data contains [*sic*] within' (Small Data Lab, 2014). In this initiative, the personal by-product data that people contribute to big data sets are reclaimed and returned to these individuals for their own use. The ideal is to create a 'rich personal data ecology', in which the various forms of data that people generate can be archived and joined together in 'personal data vaults' to provide insights for those users (Paz, 2013).

This drive towards 'sharing your numbers' and recounting experiences of self-tracking fits into the wider discourse of

sharing personal details and experiences with others, which underpins many activities on Web 2.0 social media platforms (Beer and Burrows, 2013; John, 2013). In this discourse of sharing, help and support from others, and building better information from aggregated data sets, individualism as expressed in self-tracking cultures can have a strongly participatory dimension. Individualism remains a key attribute; but it is contended that one can achieve the optimal self more quickly as part of a participatory culture. Self-entrepreneurialism is represented both as contributing to the broader knowledges developed via digitisation and as benefiting from digitisation, in a synergistic or cybernetic relationship of self to others. In this context self-reinvention and reflexivity are shared undertakings.

The imperative of being able to manage and control the continuous streams of information that are generated by self-tracking is integral to self-tracking cultures, as I discussed in Chapter 4. Reflecting on the challenges of which data to collect, how to make sense of and visualise the data, and how to apply this knowledge to one's life is part of the issue of 'controlling my data', which frequently comes up for discussion on the Quantified Self website and in members' meetups and conferences. Increasingly, such discussions incorporate examination of how self-trackers' personal data are used by other actors and agencies and how the users themselves can seek to gain greater control over where the data go and how they are used.

Nafus and Sherman (2014: 1785) contend that self-tracking is an alternative data practice that is a form of soft resistance to algorithmic authority and to the harvesting of individuals' personal data. They argue that self-tracking is nothing less than 'a profoundly different way of knowing what data is, why it is important, who gets to interpret it [*sic*], and to what ends'. However the issue of gaining access to one's data remains crucial to questions of data control and use. While a small minority of technically proficient self-trackers are able to devise their own digital technologies for self-tracking and thus exert full control over their personal information, the vast majority must rely on the commercialised products that are available and therefore lose control over where their data are stored and who is able to gain access. For people who

have chronic health conditions, for example, access to their data can be a crucial issue. A debate is continuing over the data that are collected by continuous blood glucose monitoring and whether the patients should have ready access to these data or only their doctors. As one person with diabetes contends on his blog, older self-care blood glucose-monitoring devices produce data that patients can view and act on immediately. Why should the information generated by the newer digitised continuous blood glucose monitors be available only to doctors, who review it some time later, when patients could benefit from seeing their data in real time (Dubois, 2014)? A similar issue arises in relation to the information that is collected on heart patients' defibrillator implants. The data that are conveyed wirelessly to patients' healthcare professionals cannot be easily accessed by the patients themselves. In jurisdictions such as the United States, the device developers are legally prohibited from allowing patients access to their data (Dockser Marcus and Weaver, 2012).

There is recent evidence that the Quantified Self movement is becoming more interested in facilitating access to personal data for purposes beyond those of individuals. In a post on the Quantified Self website entitled 'Access matters', Gary Wolf (2014) comments that self-trackers have no legal access to their own data, which they may have collected for years. Nor is there an informal ethical consensus that supports developers in opening their archives to the people who have contributed their information. Wolf and others associated with the Quantified Self movement have begun to campaign for self-trackers to achieve greater access to the personal data that are presently sequestered in the cloud computing archives of developers. They argue for an approach that leads to the aggregation of self-tracked data in ways that will benefit other people than individual self-trackers themselves.

Some Quantified Self movement-affiliated groups have begun to experiment with ways in which self-tracking can be used for community participation and development. Members of the St Louis Quantified Self meeting group, for example, have worked on developing a context-specific app that allows people to input their moods and identify how certain spatial locations within a community affect emotional responses. They are also developing a Personal Environment Tracker

that would allow St Louis citizens to monitor their own environmental impact and that of the community in which they live (Ramirez, 2014).

The Quantified Self Lab, the technical arm of the Quantified Self movement, has also announced that it is becoming involved with citizen science initiatives in collaboration with the US Environmental Protection Agency (Ramirez, 2015). It has now joined with the Robert Wood Johnson Foundation, an American philanthropic organisation focused on health issues, to work on improving people's access to their personal data. Both groups are also collaborating with other partners on the Open Humans Network (Open Humans, 2015), which is aimed at facilitating the sharing of people's details about their health and medical statuses as part of a participatory research initiative. Participants who join in this initiative are asked to upload the data that they have collected on themselves through self-tracking devices as well as any other digitised information about their bodies that they are able to offer for use in research studies. Part of the model that the Open Humans Network has adopted is that researchers agree to return to the participants themselves any new data that emerge from projects that use these participants' information, and participants decide which of their data they allow others to access.

A number of initiatives have developed that incorporate the aggregation of self-tracked data with those of others (apart from members of the Quantified Self movement), as part of projects designed to benefit both the individuals who have collected the data and the broader community. Citizen science, environmental activism, healthy cities and community development projects are examples of these types of communal self-tracking endeavours. These initiatives, sometimes referred to as 'citizen sensing' (Gabrys, 2014), are a form of crowdsourcing. They may involve the use of data that individuals collect on their local environs, such as air quality, traffic levels or crime rates, as well as on their own health indicators – or a combination of both. These data may be used in various ways. Sometimes they are simply part of collective projects undertaken at the behest of local agencies, but they may also be used in political efforts to challenge governmental policy and agitate for improved services or planning. The impetus may come from grassroots

organisations or from governmental organisations; the latter construe it as a top-down initiative or as an encouragement towards community development.

Self-tracked data here become represented as a tool for promoting personal health and wellbeing at the same time as community and environmental development and sustainability. As these initiatives suggest, part of the ethical practice of self-tracking, at least for some practitioners, may involve the notion of contributing to a wider good as well as collecting data for one's own purposes. Access to large data sets – rendering these data sets more 'open' and accessible to members of the public – becomes a mode of citizenship that is distributed between self, community and physical environment. This idea extends the entrepreneurial and responsible citizen ideal by incorporating expectations that people should not only collect their own, personal information for purposes of self-optimisation but should also contribute it to tailored, aggregated big data that will benefit many others, in a form of personal data philanthropy: self-tracking citizenship, in other words.

Responses and resistances to dataveillance

As humans increasingly become nodes in the Internet of Things, generating and exchanging digital data with other sensor-equipped objects, self-tracking practices, whether taken up voluntarily or pushed or imposed upon people, will become unavoidable for many. The evidence outlined in this book suggests a gradually widening scope for the use of self-tracking, which is likely to expand as a growing number of agencies and organisations realise the potential of the data produced from these practices. As the monitoring of individuals' bodies, energy use, work productivity, moods, social relationships, purchasing habits, driving practices and so on becomes more routine and widespread, the extent to which the subjects of this tracking can opt out becomes limited. People may have few choices about whether or not to participate as data-generating subjects.

It is important, however, to emphasise that dataveillance (or any other mode of watching) is not an inevitable, fail-safe

operation. It is always responded to with resistant strategies (Raley, 2013) that may be more or less effective. While people can no longer escape being the subjects of dataveillance, they can to some extent make choices about the self-tracking practices in which they may engage and about the devices they decide to use. They may seek out developers and manufacturers who are responding to consumers' concerns about data privacy and security.

There have also been calls for the use of the policy of 'privacy by design' when developing digital devices. This concept emphasises that the protection of consumers' privacy should be a major element in the design of objects such as smart technologies. Such discussions refer to the notions of the 'user-centric internet' and 'controlled computing', where people's personal data will be protected by the judicious structuring of information systems engineering, above the demands of those who wish to profit from or otherwise use these data (Cavoukian and Kruger, 2014). As a designer of digital systems, Lloyd (2014) argues for the importance of making systems that are more transparent, so that users can understand how they operate, what information they are collecting and how these data are algorithmically interpreted. She advocates for digital systems that give over more agency to users, so that they feel more in control.

Dodge and Kitchin (Dodge and Kitchin, 2007; Kitchin and Dodge, 2011) have suggested that lifeloggers should not try to achieve the total recording of as many details of their lives as they can, as is proposed by the ideal of lifelogging. Instead, as a way of evading surveillance and the appropriation of their personal details by others, lifeloggers should seek to achieve only a partial record, by using devices that block the recording of some details or record others only imperfectly. Dodge and Kitchin (2007) also suggest that 'an ethics of forgetting' should be incorporated into the design of lifelogging devices and software as part of allowing people to forget some aspects of their lives and to evade the close surveillance of their lives exerted by others. People should be able to 'dupe the log' in order to 'unsettle the authenticity of the record' (Dodge and Kitchin, 2007: 439).

Dodge and Kitchin (2007) further assert that forgetting should be viewed as an emancipatory process, which allows

for the freedom of escaping the bounds of remembering, rather than as a weakness or fallibility, as lifelogging discourses tend to suggest. The best type of lifelog, they argue, is one that conforms to the fallibility of human memory so that it might degrade in terms of its accuracy over time, as human memory does, losing or changing some details while preserving others. The recording of an event, for example, would be an impression rather than a highly precise and accurate record. Algorithmic strategies could be incorporated into digital self-tracking devices in order to promote this type of duping of the log and to evade the 'merciless memory' of digital recording of details (Dodge and Kitchin, 2007: 443).

Various other strategies for dealing with a perceived loss of control over people's personal data have been proposed. One is that of obfuscation: the deliberate production of false, misleading or ambiguous data (Brunton and Nissenbaum, 2011). Examples of software that has been developed for this purpose include AdNauseam and TrackMeNot. These are browser extensions that have been expressly designed as political strategies for online users to avoid dataveillance by commercial companies. They do not use encryption or concealment, but instead the opposing strategies of creating digital 'noise'. TrackMeNot hides real web searches among a plethora of false ones, creating 'ghost queries'. AdNauseam works in conjunction with an ad-blocker tool. It automatically clicks on blocked ads that the user has never viewed, thus creating a false trail of information about the users' browsing habits and rendering user profiling and monitoring useless for the ad networks' databases.

Other means of engaging in counterveillance include the use of such tools as Eyebrowse, a Firefox plug-in that visualises the user's web browsing history as well as those of the user's friends. In so doing, this tool displays the data that internet companies are able to collect when people browse the internet. The use of this type of tool may be described as a self-tracking technology for revealing others' tracking of a person's activities (in other words, the tracking of tracking), with the objective of developing greater awareness of where people's personal information goes when it enters the digital data economy. Here self-tracking becomes a mode of learning about a user's participation as a subject in dataveillance.

Final Reflections

I have suggested in this book that self-tracking cultures have emerged in a sociocultural and political context in which various rationales, discourses, practices and technologies are converging. These include the following:

- concepts of the self that value self-knowledge and self-entrepreneurialism;
- ideas about the body that champion tight regulation, control and order;
- the privileging of knowledges that are regarded as scientific, and therefore neutral and objective, supposedly unsullied by human subjectivity or bias;
- a moral and political environment in which taking responsibility for one's life and health is privileged and promoted;
- the affordances of new digital technologies that are able to monitor an increasing array of aspects of human bodies, behaviours, preferences and habits in ever greater detail;
- the emergence of the digital data knowledge economy, in which digitised personal information bears significant commercial, managerial and research value; and
- the realisation, on the part of governmental, managerial and commercial actors and agencies, that they can