The ethics of uncertainty for data subjects

Philip J. Nickel, Eindhoven University of Technology, p.j.nickel@tue.nl

[Manuscript version. Final version to appear in J. Krutzinna & L. Floridi, eds., *The Ethics of Medical Data Donation* (forthcoming).]

Abstract:
Modern health data practices come with many practical uncertainties. In this paper I argue that data subjects' trust in the institutions and organizations that control their data, and their ability to know their own moral obligations in relation to their data, are undermined by significant uncertainties regarding the what, how, and who of mass data collection and analysis. I conclude by considering how proposals for managing situations of high uncertainty might be applied to this problem. These emphasize increasing organizational flexibility, knowledge, and capacity, and reducing hazard.

Keywords:
Ethics of data donation; practical uncertainty; opacity of algorithms; profiling; trust; value-based health care; systemic oversight; privacy-by-design; data professionalism

I. Uncertainty and data ethics

Modern mass data collection and analysis promise great innovation in the health domain, as well as significant uncertainty. The CEO of one of the world's largest technology companies has said that "fear of data-mining" leads to 100,000 preventable deaths per year (Hern 2014). One plausible explanation for such fear is that health data subjects feel uncertain about the implications of data innovation.

In this chapter the uncertainty surrounding emerging technologies is analyzed as a practical problem for data subjects. The style of ethical analysis employed here is somewhat new. Brey writes that "the main problem for the ethics of emerging technology is the problem of uncertainty"; however, in contrast to the present approach, he proposes "anticipatory technology ethics that tries to forecast various possible future developments" (Brey 2017, 175, 178). The analysis in this chapter is complementary to such an "anticipatory ethics" but does not aim to forecast future developments. Instead, it looks at the causes

and practical consequences of uncertainty for data subjects in the present tense.[1]

Some rough definitions of key concepts are needed for this analysis. *Practical uncertainties* are defined as things we *do not know* and have an *interest* in knowing (Goldman 1999; Fallis 2006). Their ethical significance therefore has two dimensions: (a) the features of data practices that create unknowns; and (b) the interests of data subjects that are impeded by these unknowns.

Data subjects refers to those people whose data is collected and processed, whether they provide this data voluntarily or not. For example, a person who gives access to genetic test results or uses 'wearable' or in-home medical data-collecting devices, or consumer smartphones with built-in health services is a data subject. In cases where people provide highly explicit and voluntary consent to the transfer of data, we can speak of *donation*. However, in what follows it will be argued that within the context of current data practices it is often unclear whether a data transfer can really count as a donation, because whether it is truly a donation is itself a morally significant uncertainty.

The argument here concerns health data but is also relevant for many other domains where personal data is shared and collected on a mass scale, such as social media, financial planning, workplaces, and urban spaces. The boundaries between health data and other kinds of personal data are blurring to some extent: "The traditional boundaries of primary and tertiary care environments are breaking down and health information is increasingly collected through mobile devices, in personal domains … and from sensors attached on or in the human body…. At the same time, the detail and diversity of information collected in the context of healthcare and biomedical research is increasing at an unprecedented rate" (Malin et al. 2013, 2). An extension of this point is that a great deal of not-seemingly-health-related data can be used for medical and health purposes (Prainsack 2017).

In accordance with the definition of practical uncertainty above, the argument to be pursued here can be expressed in the following way:

1. Fundamental features of our data practices, including the open-endedness of data to new insights and applications, the opacity of data analysis (here referring to the inaccessibility and/or incomprehensibility

[1] In this sense, my approach is what Brey would label a "generic approach" to the ethics of emerging technology that considers "inherent features of the technology" rather than an "anticipatory approach" that uses "foresight methods" (op. cit., 178-179). However, it analyzes uncertainties *about* the future.

of how algorithms analyze data), and the persistence of data, imply uncertainty regarding the what, how, and who of data practices.
2. Two important epistemic interests of data subjects are threatened by this uncertainty: (i) having trust in the institutions that manage data, and (ii) knowing one's ethical obligations with respect to data sharing.
3. Therefore, other things equal, we should take feasible policy measures to mitigate uncertainty.

In line with this argument, Section II discusses some endemic aspects of our data practices that create uncertainties, and Section III addresses our interests in having knowledge in the domain of health data. Section IV concerns possible strategies for mitigating uncertainty. Such strategies, if effective, could make it less ethically problematic to obtain the many benefits associated with mass data collection and analysis, and could help people overcome the "fear of data mining" mentioned at the beginning of this section.

II. What features of data practices create unknowns?

Three features of data and data practices — *open-endedness*, *opacity*, and *persistence* — together give rise to significant uncertainties for data subjects. These uncertainties are distinctive because they cannot easily be avoided by engaging in best practices for risk reduction (for example, through better data security). To some extent they are part and parcel of any scenario for mass collection and processing of data. They are not futuristic. They are implied by many practitioners' statements about current practices, both routine and avant-garde, as well as in current interpretation of these practices. Those familiar with data ethics might find the features of data practices discussed in what follows unsurprising. What is new here is how they are conceptualized and deployed in relation to uncertainty. The focus is on uncertainties that arise, not when something goes wrong with the management of data, but rather when it is being used as its controllers intend: uncertainties due to the very nature of digital data as a form of information and our practices of using it.[2]

[2] Collingridge (1980) is famous for arguing that we can only control the risks of technological innovation early in its development, but we can only know what risks to try to prevent after it is well underway. The uncertainties I focus on here cannot easily be prevented for another reason, which is that they are almost inseparable from the underlying data practices, strongly linked with the transformative potential of those practices, and therefore not likely to be eliminated.

Before exploring these three uncertainty-creating characteristics of modern data practices, a brief remark is needed about what uncertainty means here. In practical ethics we are often concerned with *risky* uncertainties: possible, unwanted future states of affairs (e.g., possible data thefts). Risk and uncertainty are often defined so that they refer to distinct phenomena: 'risk proper' is probabilistic uncertainty about future unwanted events where both the probabilities and the possible outcomes of these events are known and quantifiable, whereas 'uncertainty proper' is a lack of knowledge about which outcomes are possible and/or their probabilities (Knight 1921; cf. Altham 1983). Some authors draw a further distinction between uncertainty and ignorance, where ignorance involves inability to predict outcomes or plausible scenarios (Wynne 1992, cited in Dereli et al. 2014). The kind of uncertainty to be discussed in what follows lies in between the categories of uncertainty proper and ignorance: we can identify some plausible horizons of possibility, but not exhaustively or quantifiably.[3]

Open-endedness

Open-endedness — the potential for creating and applying data-based knowledge in new ways — creates significant uncertainties for data subjects at the time when their data is collected and afterwards. Data is multiply interpretable, especially when combined and used for new purposes. Different algorithms and analytical lenses, such as different strategies of classifying, combining, and finding patterns in data, allow for new predictions and correlative generalizations. This is essential to the promises of data collection and analysis: "the value of big data lies in the unexpectedness of the insights that it can reveal" (Barocas & Nissenbaum 2014, 60). Since we cannot form expectations about these important insights, open-endedness creates significant uncertainties.

There are at least two dimensions of open-endedness. The first is the *fecundity* of inferences that can be drawn when a dataset is larger or better organized, or where more powerful analytical tools are used. The second is *recontextualization* of data across contexts. We can think of the first dimension

[3] Consequences known to be harmful for some individuals are likely to be directly caused by the further development of big data practices, such as the ability to reidentify unidentified ("anonymous") data subjects using new and more powerful analytical techniques. This could in some cases lead to loss of insurance or other harms for re-identified individuals. For example, in Lippert et al.'s (2017) controversial study, recognizable images of the faces of individuals were said to be reconstructable using data from gene sequences. There is dispute about whether the results really prove what the authors say (Erlich 2017). Irrespective of this dispute, my point here is that the looming possibility of such techniques creates horizons of uncertainty that exist long before any future harms that result. These uncertainties are ethically significant in their own right.

as the *depth or power* of the inferences we can make from a set of data, and the second dimension as the *practical applicability* of these inferences in a diverse range of contexts in real time.[4]

A real-life example of open-endedness from the health domain is the vision of the 'value-based health care' movement. This movement proposes to align payment for health services closely with actual health outcomes, creating a transformation of health care. Its founders have maintained from the beginning that data collection and analysis are necessary instruments for this transition because they make it possible to develop and apply a nuanced health-improvement metric for reimbursing health costs across the board. One early proponent, focusing on the inefficiencies of the American health care system, devotes several paragraphs to the importance of data as a means toward value-based care:

> Measurement and dissemination of health outcomes should become mandatory for every provider and every medical condition … We need to measure true health outcomes rather than relying solely on process measures, such as compliance with practice guidelines, which are incomplete and slow to change. … Among our highest near-term priorities is to finalize and then continuously update health information technology (HIT) standards that include precise data definitions (for diagnoses and treatments, for example), an architecture for aggregating data for each patient over time and across providers, and protocols for seamless communication among systems (Porter 2009).

This underlying idea has persisted both in value-based health care and in other similar movements such as the Institute of Medicine's 'learning health care system': with new sources of data and analytical tools, we can explore new ways of modeling and addressing the causes of inefficiency and suboptimal health outcomes (Committee 2013; Mulley et al. 2017). Both fecundity of insight *and* recontextualization of real-time decision-making are needed for the envisioned transformations.

Other examples emphasize recontextualization to a greater degree: cases in which an integrated situational awareness is stitched together from data originating in multiple contexts, creating a single 'dashboard' or 'visualization'

[4] Some commentators have raised epistemological concerns that big data is overhyped as a scientific field and may not withstand scientific scrutiny of its knowledge claims (Mittelstadt & Floridi 2016, Lipworth et al. 2017). My argument does not depend on the validity of the relevant knowledge claims as a whole, but rather their plausibility.

for decision-making. Suppose two large sets of data on treatments, costs, and patient outcomes, one collected by hospitals, and a second collected by general practitioners, are being combined for the first time. If health care is managed through substantially separate structures, then mutual access to this information holds the prospect of bringing about better integration and continuity of health care for both hospitals and GPs (e.g., Sheaff et al. 2015, 57). Similar contextual awareness is anticipated elsewhere in health care: for example, in the integration of "informal health and fitness data collected by the user together with official health records collected by health professionals" (Gay & Leijdekkers 2015). These cases stress the recontextualization of information, but also promise insight when complementary data is combined.

Profiling people in multiple and unpredictable ways is an ethically relevant aspect of open-endedness in data analysis. Profiling has been defined as "the process of 'discovering' patterns in data … that can be used to identify or represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent an individual subject or to identify a subject as a member of a group (which can be an existing community or a discovered category)" (Hildebrandt 2009, 275). This and other definitions explicitly relate to both dimensions of open-endedness: fecundity ("discovering") and recontextualization ("application", "identification"). Profiling is particularly relevant to data subjects in a health context because it has the potential to classify them for diagnosis, treatment, and reimbursement in unpredictable ways. For example, it might be used as a reason to choose a particular diagnostic, or to deny treatment altogether.

Opacity

A second source of endemic uncertainty in our data practices is the use of opaque algorithms and 'deep learning' to analyze data (Kennedy et al., 2015; Rieder & Simon 2017). Consider a widely discussed recent example in which a deep learning algorithm was trained to identify profile photos from a prominent social media site as being gay/ lesbian or straight (Wang & Kosinski 2018). The algorithm was able to determine this with considerable accuracy, better than that of human raters. However, because the training was automatic and data-driven, it is not known what features the machine correlated with sexual orientation identity.

This example shows that one possible reason why data analytics is opaque is that deep learning techniques do not disclose the underlying pattern of their

learning (Mittelstadt et al. 2016). Some algorithms are highly complex, and some are modified frequently (Rieder & Simon 2017, 6). However, not all algorithms are complex or difficult to understand; there are also other reasons why data analytics is opaque. One is secrecy: algorithms are often not shared due to intellectual property issues, competitiveness, inertia, or concerns that they will not withstand scrutiny (Burrell 2016; Christophersen et al. 2015; Gillingham 2016; Stodden 2010).

A complicating issue is that data subjects rarely have the concepts needed to understand the actual algorithms and deep learning techniques themselves. However, that is not in itself an epistemological barrier to knowledge. On many views of knowledge in a social world, it is socially constituted. Laypeople can have knowledge that is partly constituted by the knowledge and understanding of others, including experts (Faulkner 2011, Goldberg 2010). A serious problem arises principally when experts do not have this knowledge themselves, *or* when they do not carry out their functions in a way that confers socially constituted knowledge upon data subjects. For example, a plausible condition on socially constituted knowledge is that there is some person or collective of persons that has understanding and is willing and able to provide an articulate explanation when asked or challenged.[5] When trained scientists working with opaque algorithms do not understand or show willingness to articulate how conclusions are being derived through data analysis, this condition fails. This creates significant uncertainty about how data analysis is applied to health data now, and especially about how it could be applied in the future.

Persistence

Data is long-lived and duplicable; here I call the combination of these two features *persistence*. Unlike most collections of physical biological materials used for scientific and therapeutic purposes, once a collection of health data is gathered it is feasible to preserve it indefinitely and give access to it prolifically. For physical biological materials, this necessitates storage and, in cases of cell cultures, *in vitro* reproductive techniques. For health data on a large scale, this necessitates computer storage and various means of sharing or giving access to large amounts of data.[6] Because it is quite feasible to store, copy, and access

[5] Here my analysis differs from Burrell's (2016) in that I do not regard widespread "technical illiteracy" about data analysis as a basic form of opacity. Ordinary people can unproblematically obtain "second-hand knowledge" from experts in many domains, even when they are technically illiterate.

[6] Collections of biomedical samples or 'biobanks' are always associated with data, and especially where population-level biobanks are concerned this data component is just as important as the 'wet' biological component (as in the definition of the Council of Europe 2006).

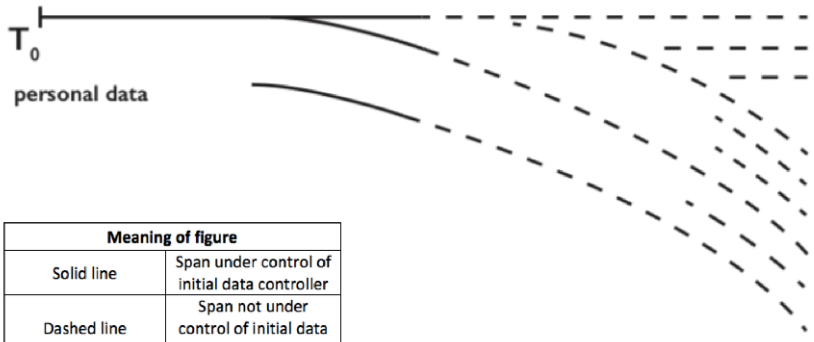
data at a “medium” scale (i.e., well below the limits of Moore’s law), this leads to a potential for reproductive profligacy of health data that extends indefinitely into the future.

Persistence is a relevant source of ignorance for the data subject because many different institutions and organizations with different interests and motivations store, share, and analyze data. Vayena & Blasimme describe a “data ecosystem” with an “increasing number of stakeholders” including “the data analytics industry ... [and] social media giants … [that] enter the domain of health bringing corporate cultures that are not necessarily aligned with existing regulations in health research” (2018, 121). Commercial organizations and governmental and academic institutions often cooperate in data-intensive projects, and the boundaries of data (access) are often not limited by institutional, regional, or national boundaries. Data about a person from one context or jurisdiction can be copied multiple times and shared with many different kinds of entities in different contexts or jurisdictions, with different motives (profit, surveillance, efficiency). Moreover, the results of data analysis, such as the results of profiling, become data entities of their own, which also share these features of longevity and shareability and can be distributed and reused for new purposes. In combination, these factors imply that multiple entities and types of entities (e.g., commercial entities, research entities) are likely to control one's personal health data in the long term, and that data profiles concerning data subjects are likely be generated which endure and are shared across contexts and jurisdictions.[7]

A figure (Figure 1) helps to visualize this as a source of uncertainty. The lines, starting at T_0, represent the lifespan of the data. The solid lines are those parts of the lifespan under the intentional control of the original recipient or collector of data. The dashed lines represent the parts of the lifespan that are not under the intentional control of the original recipient or controller of data. These dashed lines are particularly uncertainty-inducing because they are no longer governed by the same assumptions that the data subject might have reasonably made at T_0 about the motives and interests of the entities possessing the data. The lines (both solid and dashed) can branch, of course, because parts of the data can be given away or duplicated. In addition, new branches, consisting of analyzed data or profiling data based on the original

[7] Deidentification of original data shared by data subjects does not prevent those subjects from being targeted in a way that resembles profiling. For example, data from patients at a particular medical practice can be deidentified and used to make generalizations about the practice, which are then used to target those very patients.

data but not identical to it, can start independently. These are represented as solid or dashed lines starting at times after T_0.



Meaning of figure	
Solid line	Span under control of initial data controller
Dashed line	Span not under control of initial data controller
Branching line	Data is duplicated
Beginning of new line after T_0	Related but independent data is created

Figure 1. The persistence of personal data

Endemic uncertainties combined

The open-endedness, opacity, and persistence of data and data practices together create a host of unknowns for data subjects. Such unknowns include: how collected data about the subject will be combined in the future, how the combined data will be used to measure, classify, and profile the subject, and what implications new metrics and regimes of access to information will have for the subject. We can think of these as *unknown knowns*: they are forms of knowledge and knowledge-based power that will come into being in the future but are currently unpredictable and unknown.[8]

[8] Slavoj Žižek introduced this term in relation to former U.S. Secretary of Defense Donald Rumsfeld's infamous speech about 'known unknowns' and 'unknown unknowns', in order to refer to things one does know, but does not realize or admit that they know (2004). It was later used as the title of a film about Rumsfeld by director Errol Morris. My use of the term departs from these earlier uses.

III. Two epistemic interests of data subjects

In section I, practical uncertainty is defined as matters that we *do not know* and have an *interest* in knowing. We have so far been focusing on those aspects of our data practices which create unknowns. Now let us turn our attention to the second part of the definition, which refers to the *interests* we have in understanding and knowing.[9] In what follows I focus on two high-priority interests that are particularly impacted by the unknowns discussed in the previous section: one's ability as a potential data subject to *trust* others with one's data, and one's ability to determine one's own *moral obligations* in relation to oneself and others where data issues are concerned. For each of these interests, I begin by presenting a case in which the relevant interest is intuitively present.

A. Interests in trust

> **SHARA**
> Shara is considering going to a hospital because she believes she may have been exposed to HIV in a sexual encounter (although she believes the risk is very low). She believes she could obtain a prescription for post-exposure HIV prophylaxis. However, she is not sure of the implications of trusting the hospital or the pharmacy with these 'data points'.

Can Shara trust the hospital and other relevant institutions with her data? Considering the uncertainties associated with institutional data practices, it might be rational for Shara to make a strategic assessment of whether the risk of HIV infection makes it worth visiting the hospital under these circumstances. She does not know who will come to possess her data in the future, how they will analyze it, and for what purposes. For all she knows, she could be profiled as being high-risk and denied service or offered different care in the future. From the perspective of individual rational choice, if not from the perspective of public health (Ford et al. 2015), such uncertainty could tip the balance in favor of not seeking treatment. This is an urgent epistemic and practical problem for Shara.

[9] Two main senses of 'interest' are operating here in a way that is mutually reinforcing: something can be *in my interest* to know, or it can be interest*ing*, or both. I use the term 'interests,' rather than 'rights' or 'needs,' because the latter terms presuppose that epistemic concerns are so strong as to be ethically overriding (i.e., to serve as 'trumps' over other values). I think it will be clear that the interests in question *are* sometimes sufficiently weighty to override other values or interests, but this need not always be the case.

As this example illustrates, one weighty epistemic interest of data subjects is to have sufficient reason to trust entities such as governments, corporations, research institutions, and hospitals with data. Data practitioners and scholars have remarked upon this interest in trust, particularly in the health (care) domain, where trust is a bedrock value (Larson 2013; Lipworth et al. 2017). Our attitudes about the social, political, and technological world depend on trust. Trust frames how we think of our prospects for cooperation, and the responsibilities of others.

Trust involves a complex of predictive and normative expectations based on the interests, motives, and past performance of the trusted entity (Voerman & Nickel 2017). With a few notable exceptions (Hardin 2006), many scholars, including some philosophers, take for granted that one can trust institutions (Hawley 2017). We can hold this kind of trust towards specific organizations (such as Harvard University or the NHS) and various functional human roles within them (e.g., the role of data scientist or clinical researcher). Trust in institutions is based on our ideas of the norms and functional aims that govern and define organizations and the roles within them, in addition to individual characteristics such as goodwill or moral character that ground person-to-person trust (Baier 1986, Holton 1994). Trust in institutions is distinctive in that it does not normally involve the expectation that the trusted entities will be specifically responsive to the trust one places in them. In this respect, it differs from trust between intimates (Faulkner 2011).

People's interest in trust is not merely to have trust, but to have it in the right circumstances and for the right reasons. Normally, this aspect of trust is backed by having a reliable grasp of the interests, functions, and norms that motivate and explain the trusted entity's behavior. This idea of a reliable grasp can be cashed out in more 'internalist' or 'externalist' ways. Internalism means that one's warrant for trust consists primarily of items to which one has "direct and unproblematic access" (Bonjour & Sosa 2003); externalism means that it can also substantially consist of items in one's social or physical environment to which one does not have access. Manson & O'Neill (2007) put forward an ideal of 'intelligent trust' that emphasizes the virtues or talents of an individual truster in making good choices about whom to trust. Others advance a notion of 'healthy trust' or 'sound trust' that emphasizes the importance of the environment as well as the individual in creating the conditions for epistemically grounded and non-exploitative trust (Boenink 2003, Voerman & Nickel 2017). Loosely speaking, the first account emphasizes the internal

aspects of warranted trust, and the second account emphasizes the external aspects. (However, the notion of "intelligent trust" could also be given an externalist interpretation, as Sosa does for the idea of intellectual virtues more generally (Bonjour & Sosa 2003).) Either way, intelligent or healthy trust depends on a stable, reliable ascription of norms and functional aims to the institutions we rely on.

The endemic uncertainties of our data practices, explicated in the previous section, threaten this epistemic basis for trust. They make it very difficult to have a stable, reliable grasp of norms and functions of the entities we rely on, or even to determine which entities are actually involved. Uncertainties about what kinds of organizations and institutions will come to possess one's data in the future, and about how data might be used for profiling, make it difficult to trust because such uncertainties threaten the warrant for trust. A data subject may reasonably wonder whether the new metrics of value-based health care might in the future be used to profile her (perhaps using an opaque algorithm) as being a poor prospect for health outcomes, or whether her data will be transferred to new entities whose motives and interests oppose her own. When such scenarios cannot be defined, the epistemic grounding for trust in health care institutions is missing. The 'what', 'how', and 'who' of trust cannot be specified.

Brown & Calnan (2012) have analyzed situations like Shara's in which there is a high degree of uncertainty and institutional complexity in clinical contexts, in terms of trust. They argue that trust becomes an explicit problem in such contexts because its rational basis is threatened (ibid., 4). However, trust remains salient as a possible way of "bridging" uncertainty (ibid., 53ff.). I follow their analysis when looking at data practices. Trust remains a possible strategy for navigating situations that arise in the midst of those practices, even when the uncertainties surrounding our data practices threaten and undermine its familiar epistemic foundation. However, such a strategy is like the Biblical house built on sand, whose foundation is unstable. It can be occupied, but existential threats to it cannot be rationally put out of one's mind.

B. Interests in knowing our own obligations

> **CARLA**
> Carla has recently moved to a new area. She has a serious health problem. When she arrives at the hospital to get medical treatment for

her problem, she chooses to conceal a past pregnancy and a past depression, preventing both events from becoming data points.

Is Carla failing to act in good faith? Is she unfairly advantaging herself over others in order to gain access to health care, by leaving out something that is relevant to clinical decision-making? Or is she simply protecting her privacy from exploitation by commercial and research organizations she does not endorse? Knowing one's obligations determinately means being able to answer these questions. Being morally responsible as a citizen and a member of the moral community seems to require such knowledge. Intuitively, knowing one's own obligations determinately is an important human interest.

In order to know our obligations in relation to data practices we must know who will have access to our data, what the data means for us, and how it will be used. If a patient's data will be used to profile her for unspecified purposes that extend far beyond the provision of medical care or for unrelated commercial purposes, then it seems intuitively that an act of concealment by the patient does not violate any moral duty of honesty or fairness to others but is rather a matter of protecting her own privacy. On the other hand, if the data is to be protected rigorously and used only in research that could benefit others similar to her, and her choice to conceal data actually hinders this goal, then arguably she can be seen as acting dishonestly and unfairly by not disclosing important facts from her medical history. This creates uncertainty about the duties and responsibilities conferred on different parties by a data transfer. The status of a patient's data transfer could be seen as a kind of *donation*, as the *price* of a service, or as a *shared burden* — a sort of *tax* — imposed for the sake of fairness and solidarity. Which of these ways of thinking about data transfers and their associated "deontic consequences" is the correct one is unclear and indeterminate in many cases.[10] Intuitively, this kind of moral uncertainty frustrates important interests of data subjects.

The linkage between uncertainty about our data practices and uncertainty about what obligations result from transfers of health data is implicit in Cohen's (2017) argument that people have a duty to share health data as a matter of solidarity. The argument starts with the crucial assumption that the possessor of shared health data will be either a government agency or a hospital system "committed to improving healthcare … for the people it serves," not a for-profit commercial entity (ibid., 210). When this assumption is reliably satisfied, we can think of health data sharing as having the status of a

[10] Compare Tutton's (2004) discussion of how to frame the "sharing" of biological materials in biobanking.

reciprocal shared burden or a tax, where everybody has an obligation to contribute, and gratitude and specific goods and services are not expected in return. Conversely, though, if there is significant uncertainty about whether data will be used for purposes unrelated to health, for commercial purposes, or by new organizations and institutions, then there will also be uncertainty about the conclusion that people have a solidaristic obligation to share health data.

An important corollary of the linkage between data practices and uncertainty is that *uncertainties about data transfers challenge the very idea of data donation*. Making a donation (i.e., gift-giving) is an act associated with other morally-laden acts and attitudes such as gratitude, and is not easy to combine with other moral regimes such as that of communitarianism and solidarity, or that of a commercial exchange (Herman 2012). In real practice, giving away data is often thought of as the *price* of using web-based services. Data is a kind of bartering chip that one uses to pay for these services. The idea of "Web 2.0" has been coined for a business model in which users are *prosumers*, who both *produce* content and data for Internet sites and applications and also *consume* — often "for free" — the valuable services that websites and apps deliver (Toffler 1980; Ritzer & Jurgenson 2010). Prosumption is a business model for many health data companies (Prainsack 2017). So long as we remain confused about whether a given data transfer is our contribution toward carrying a shared burden, as Cohen argues, or a bartering transaction, as the business model of prosumption implies, then it will not possible to consider that very transfer of data to be a pure donation at the same time.

In the remainder of this section, I address a philosophical objection to the idea that uncertainty really threatens our epistemic interests. (Those who are not worried about such an objection may choose to skip to the next section.) So far I have relied on the intuition that knowing our moral obligations determinately makes one better off. However, according to some philosophers, even if we do not know the outcomes of our actions determinately, or even the various possible ways of valuing possible outcomes, we can still calculate our moral meta-obligations (Lockhart 2000; Zimmerman 2008, 38; Barry & Tomlin 2016; Lazar 2018). The underlying strategy for determining our meta-obligations is to consider every plausible valuation of different possible actions (the possible obligations about which we are uncertain), and then use a meta-principle to calculate one's unique meta-obligation given these possible valuations. For example, suppose our imaginary patient Carla does not know whether her data transfer would count as a

donation of data, a price that she pays in exchange for medical service, or a tax associated with the shared burden of the medical system. By taking each of these deontic statuses as members of a set of possible valuations *V*, she can apply a suitable meta-principle to calculate her unique actual obligation. An example of such a meta-principle might be, "If any of the valuations in *V* implies an obligation not to do *x*, then there is a meta-obligation not to do *x*."

This objection maintains that one's unique obligation can be specified just as well for situations of significant uncertainty as it can be for situations in which the outcomes and valuations are certain. If our knowledge of our meta-obligations under conditions of uncertainty is just as satisfactory, ethically and practically, as our knowledge of our determinate obligations, then our epistemic interest in knowing our obligations can be satisfied perfectly well even under conditions of uncertainty. This would undermine the claim that our epistemic interest in knowing our obligations is threatened by uncertainty about data practices.

This is a deep objection deserving a thorough philosophical treatment. Here I offer three preliminary replies. The first is that there is nothing that prevents us from holding that situations in which it is rational to act on a meta-principle under moral uncertainty are situations in which we are worse off, other things equal, compared to situations in which it is rational to act on a determinate principle. The second is that, empirically, people have a strong aversion to uncertainty (sometimes called "ambiguity" in the relevant empirical literature), at least in contexts where quantifiable options are directly compared to ambiguous, uncertain ones (Fox & Tversky 1995). The third is that getting the *outcome* wrong will normally be more likely if an agent does not know her own obligations determinately, than if she does. This is true even if she acts blamelessly because she acted according to an appropriate meta-principle. Being more likely to get the outcome wrong makes her worse off even if it does not reflect directly on whether she is to blame. In sum, we can accept meta-principles for situations of moral uncertainty without giving up the empirically-supported intuition that moral uncertainty threatens our interests in an important respect.

IV. Strategies for mitigating uncertainty

Risk scholars have proposed structured guidelines for mitigating situations of high uncertainty, focusing on two main strategies: *increasing systemic resilience* and *reducing hazard* (Renn 2008). Since these strategies are well-

established, it is useful to consider how they could be applied to the uncertainties surrounding data practices. Systemic resilience refers to *flexibility* and *organizational capacity* in monitoring and responding to ongoing hazards. Hazard reduction, by contrast, is a matter of limiting what is at stake in uncertainty. Below I attempt to identify instances of each strategy from the literature on data governance and consider whether they are likely to mitigate harms to data subjects' epistemic interests. Doing so highlights the need for further research about the feasibility of such strategies, as well as the feasibility of supplementary strategies that more directly address the problem of practical uncertainty about health data.

A. Systemic resilience through flexible systemic oversight

First, I consider a strategy to increase systemic resilience. The General Data Protection Regulation (GDPR), taking effect in the European Union in 2018, may appear to be such a strategy. It places new governance requirements on data controllers and takes steps to harmonize governance across member countries. It requires that people be informed when they are being profiled (Regulation 2016, §60), and that people can find out the "logic involved in any personal data processing" (§63).

Despite these measures, one recent study finds that there is significant uncertainty about how the GDPR will be implemented in practice, and that there is likely to be a tradeoff between disruptive innovation and strict regulatory compliance (van den Broek & van Veenstra 2018). There are also other reasons why the GDPR does not solve the problem of uncertainty for data subjects. First, it only directly protects citizens of the EU. Second, many data subjects in the EU do not care about or understand the rights to know and to respond to data processing articulated in the GDPR, and consequently those rights do not protect them from uncertainty. Thirdly, even those who do care about their rights often give legal consent to data collection and processing because doing so is instrumental to obtaining services. Such acts of consent do not generally have the function of reducing uncertainty, as anyone who has clicked through an online consent form can ascertain.

Vayena & Blasimme (2018) have put forward the idea of flexible "systemic oversight" to avoid tradeoffs between innovation and regulatory compliance, while countering the impact of uncertainty. The idea is to create a comprehensive framework for governance that is reflexive, inclusive, and responsive. Systemic oversight is meant to allow for innovation while providing

"adaptive and flexible mechanisms" for oversight, in which there is "deliberative democracy" through "collective engagement of research participants in decisions about data governance" (ibid., 124-125). In relation to uncertainty, "oversight mechanisms should not be seen as procedures for prospective risk assessment, but rather as *adaptive* instruments that respond to change" (ibid., 124).

As applied to the problems discussed here, the idea is that regulatory processes resulting from collective, democratic processes will protect data subjects' interests and thereby make the act of sharing health data more rational. Mechanisms of deliberation and collective engagement would also increase well-grounded trust (in line with the account offered above in section III.A), so long as an alignment of interests results that favors data subjects and is available to them as a warrant for trust.

Flexible systemic oversight might be taken to mean that *relative to a given jurisdiction and use of data*, individuals could be given explicit guidance about their obligations and protected from the unexpected consequences of their choices. This could help to mitigate the effects of uncertainty about one's data-relative obligations. For example, within a given health care administrative region, the choice could be made to impose a solidaristic model of shared burden, in which everybody transfers data for the sake of common benefit. In cases where there were data leaks or unforeseen effects of profiling, a compensation scheme could be introduced to remedy the impacts, as proposed by Prainsack (2017). Such a regional choice could relieve people of the burden of uncertainty about their data-relative obligations.

Flexibility and systematicity are potentially at odds with one another, however. Flexibility implies that there is temporal and local adaptation to particular institutional situations and innovation regimes. However, this flexibility may actually prevent the formation of stable expectations that simplify trust decisions and make one's obligations as a data subject clear across different boundaries and jurisdictions. Regulation must address data that crosses the clinical/ non-clinical boundary, data that crosses institutional and national borders, and data that is commercialized or exploited within public-private partnerships. Flexibility and adaptability seem to imply variability. In that case, flexible and adaptable regulatory processes may be less effective at conveying to people that their interests are being consistently protected, and less effective at establishing a simple and clear set of obligations in respect of health data, compared with truly systematic (hence inflexible) regulation with clear-cut restrictions extending to all uses and jurisdictions. More research is

needed to clarify how flexibility might be balanced with systematicity, and what the impact will be on the expectations and obligations of data subjects.

B. Hazard reduction through privacy-by-design

Now I turn to a hazard reduction strategy. To begin with, it is important to note that it is somewhat unclear how to think about hazard reduction as applied to the data domain. In the domain of system safety, hazard reduction denotes the removal of hazardous substances and processes from a system, or their replacement with less hazardous substances and processes (Leveson et al. 2009). There is no direct analogue in the domain of privacy.

However, there are some crude parallels. We could, for example, encourage data obsolescence by default, or disallow "hypercollection" (Prainsack's (2017) term). Data obsolescence implies that after a certain time period, data is always deleted by default (it "obsolesces"), unless it has been specifically saved because of its demonstrable significance. Limiting hypercollection, by contrast, means that the default is not to collect or combine data in the first place unless there is a specific research motivation for doing so, such as a specific, powerful research question to be answered.

Although these measures could substantially protect privacy, they would carry an unacceptable cost in the health domain. They would block innovation and cost lives. Strictly limiting data collection or encouraging data obsolescence is difficult or impossible to combine with transformative initiatives such as the value-based health care movement considered above, in which massive and persistent data collection and analysis is built into the paradigm. Obsolescence would severely limit the value of the careful work that goes into creating a dataset.

It might be possible to think of the analogue of hazard reduction strategies in the domain of privacy as a broader set of measures that limit the degree to which personal data is threatened in the first place. *Privacy-by-design* is the name applied to strategies in which privacy safeguards are built in to a technology and attended to in the primary process of technology development and implementation, incorporating physical, technical, and procedural safeguards along the way.

Understood in this way, privacy-by-design may be too broad and vague to capture the simple and obvious logic of hazard reduction, but is nonetheless

promising as a strategy of mitigating uncertainty. Its effectiveness will depend upon the specifics of the situation and the way it is carried out. An important point to keep in mind is that some health data innovation appears inseparably to depend on information that can indirectly lead back to the subject (e.g., as a member of a relevant class or group) or can reidentify the subject when combined with other data (Mittelstadt et al. 2016).

C. Concluding Reflections

Health care faces major challenges such as the difficulty of efficiently caring for an aging population, and the increasing incidence of chronic disease that are expensive to treat. Data-based innovations are one of the main ways that technology can help meet these challenges. Lives can be saved and improved by the insights gained through health data collection and analysis. At the same time, however, these innovations create many uncertainties for ordinary people. In this paper, I have argued that these uncertainties are an ethical problem for data subjects.

An important consequence for the present chapter is that in order to see a given data transfer as a *donation*, undertaken as an act of generosity, it is not possible to see it at the same time as a bartering chip that one exchanges for a service, or as a shared burden that one undertakes out of solidarity. Resolving the uncertainties around our health data may therefore mean making a choice between seeing particular data transfers in one of these ways or another. This may limit the applicability of the idea of data donation.

An important task of future research is to further develop the kinds of governance strategies discussed above so that they better address the specific epistemic problems for data donors and data subjects explored in this paper. Another is to seek complementary approaches that directly shore up trust and reduce the costs of not knowing one's own obligations.

Focusing on trust, for example, we might consider whether a greater emphasis on data *professionalism* could shore up trust in the face of uncertainty. Professionalism arises in situations where experts in some field of activity, such as doctors, engineers, and pharmacists, adopt formal standards for having the privilege of labeling themselves a certain way, and enjoy an exclusive right to evaluate the work of others who use the label. Professionalism is often linked to trust and trustworthiness (Pellegrino & Thomasma 1993; Manson & O'Neill

2007). The underlying idea is that the development of professions functions to signal trustworthiness to those with a practical need for the relevant form of expertise. Data science professionalism is relatively undeveloped compared with professionalism in other areas of science, engineering, and medicine. By developing it in the realm of health care data, and clearly signaling what standards go along with the relevant professional identity, we could create (and communicate) trustworthiness in this area.

As for our interest in knowing our own obligations in the domain of data practices, a plausible first step is for health care authorities to acknowledge openly that there is significant uncertainty about practices of data collection and analysis. This is a matter of showing respect for the real difficulties that data subjects and potential data donors face when trying to make well supported and ethically responsible decisions to accept or resist sharing data, and may be a way to begin addressing these difficulties.[11]

[11] I would like to acknowledge the useful feedback I received on earlier versions of this paper from participants at workshops in Eindhoven, in Warwick, and at the Oxford Internet Institute. This research is affiliated with the Netherlands Organisation for Scientific Research Responsible Innovation (NWO-MVI) project "Mobile Support Systems for Behaviour Change," project number 100-23-616.

References

Altham, J.E.J. 1983. Ethics of risk. *Proceedings of the Aristotelian Society* 84: 15–29.

Baier, A. 1986. Trust and antitrust. *Ethics* 96: 231–260.

Barocas, S., and H. Nissenbaum. 2014. Big data's end run around anonymity and consent. In *Privacy, Big Data and the Public Good,* eds. J. Lane, V. Stodden, S. Bender, & H. Nissenbaum, 44–75, Cambridge University Press.

Barry, C., and P. Tomlin. 2016. Moral uncertainty and permissibility: evaluating option sets. *Canadian Journal of Philosophy* 46, no. 6: 898–923.

Boenink, M. 2003. Gezond vertrouwen. Over de rol van vertrouwen in het bevolkingsonderzoek naar borstkanker. *Krisis* 1:53–74.

Bonjour, L., and E. Sosa. 2003. *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues.* Malden, MA: Blackwell Publishing.

Brey, P. 2017. Ethics of emerging technology. In *The Ethics of Technology: Methods and Approaches,* ed. S.O.Hansson, 175–191. London: Rowman & Littlefield.

Brown, P., and M. Calnan. 2012. *Trusting on the Edge: Managing Uncertainty and Vulnerability in the Midst of Serious Mental Health Problems.* Chicago: The Policy Press.

Burrell, J. 2016. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data and Society* January–June 2016: 1–12. doi: 10.1177/2053951715622512.

Christophersen, M., P. Mørck, T.O. Langhoff, and P. Bjørn. 2015. Unforeseen challenges: adopting wearable health data tracking devices to reduce health insurance costs in organizations. In *International Conference on Universal Access in Human-Computer Interaction,* eds. M. Antona, and C. Stephanidis, 2: 88–99. Berlin: Springer Int.

Cohen, I.G. 2017. Is there a duty to share health data? In *Big Data, Health Law, and Bioethics*, eds. I.G. Cohen, H.F. Lynch, E. Vayena, and U. Gasser, 209–222. Cambridge University Press.

Collingridge, D. 1980. *The social control of technology.* New York: St Martin.

Committee on the Learning Health Care System in America; Institute of Medicine; Smith M, Saunders R, Stuckhardt L, et al., editors. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington (DC): National Academies Press (US); 2013 May 10. 5, A Continuously Learning Health Care System. Available from: https://www.ncbi.nlm.nih.gov/books/NBK207218/.

Council of Europe. 2006. *Recommendation Rec (2006)4 of the Committee of Ministers to Member States on Research on Biological Materials of Human Origin.*

Dereli, T., Y. Coşkun, E. Kolker, Ö. Güner, M. Ağırbaşlı, and V. Özdemir. 2014. Big data and ethics review for health systems research in LMICs: understanding risk, uncertainty and ignorance—and catching the black swans? *The American Journal of Bioethics* 14, no. 2: 48–50.

Erlich, Y. 2017. Major flaws in "Identification of individuals by trait prediction using whole-genome." *bioRχiv.* doi: https://doi.org/10.1101/185330.

Fallis, D. 2006. Epistemic value theory and social epistemology. *Episteme* 2, 3: 177–188.

Faulkner, P. 2011. *Knowledge on Trust.* Oxford University Press.

Ford, N., et al. for the World Health Organization Postexposure Prophylaxis Guideline Development Group. 2015. World Health Organization guidelines on postexposure prophylaxis for HIV: recommendations for a public health approach. *Clinical Infectious Diseases* 60: S161–S164. DOI: 10.1093/cid/civ068.

Fox, C.R., and A. Tversky. 1995. Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics* 110, 3: 585–603.

Gay, V., and P. Leijdekkers. 2015. Bringing health and fitness data together for connected health care: mobile apps as enablers of interoperability. *Journal of Medical Internet Research* 17, 11: e260. doi: 10.2196/jmir.5094.

Gillingham P. 2016. Predictive risk modelling to prevent child maltreatment and other adverse outcomes for service users: inside the 'black box' of machine learning. *The British Journal of Social Work* 46, 4: 1044–1058, https://doi.org/10.1093/bjsw/bcv031.

Goldberg, S. 2010. *Relying on others.* Oxford University Press.

Goldman, A. 1999. *Knowledge in a Social World.* Oxford University Press.

Hardin, R. 2006. *Trust.* Cambridge, UK: Polity.

Hawley, K. 2017. Trustworthy groups and organizations. In *The Philosophy of Trust,* eds. P. Faulkner, and T. Simpson, 230–250. Oxford University Press.

Herman, B. 2012. Being Helped and Being Grateful: Imperfect Duties, the Ethics of Possession, and the Unity of Morality. *Journal of Philosophy*. 391–411.

Hern, A. 2014. "Google: 100,000 lives a year lost through fear of data mining," *The Guardian,* June 26, 2014, https://www.theguardian.com/technology/2014/jun/26/google-healthcare-data-mining-larry-page.

Hildebrandt, M. 2009. *Profiling and AmI*. In *The Future of Identity in the Information Society: Challenges and Opportunities,* eds. K. Rannenberg, D. Royer, and A. Deuker, 273–310. Heidelberg: Springer.

Holton, R. 1994. Deciding to trust, coming to believe. *Australasian Journal of Philosophy* 72: 63–76.

Kennedy, H., T. Poell, and J. van Dijck. 2015. Introduction: data and agency. *Big Data and Society* 2. doi: 10.1177/2053951715621569.

Knight, F. 1921. *Risk, uncertainty, and profit.* Boston and New York: Houghton Mifflin.

Larson, E. 2013. Building trust in the power of big data research to serve the public good. *JAMA* 309, 23: 2443–2444. doi: 10.1001/jama.2013.5914.

Lazar, S. 2018. In dubious battle: uncertainty and the ethics of killing. *Philosophical Studies* 175: 859–883.

Leveson, N., N. Dulac, K. Marais, and J. Carroll. 2009. Moving beyond normal accidents and high reliability organizations: a systems approach to safety in complex systems. *Organization Studies* 30: 227–249.

Lippert, C. et al. 2017. Identification of individuals by trait prediction using whole-genome sequencing data. *PNAS* 114, 38: 10166–10171. doi: 10.1073/pnas.1711125114.

Lipworth, W., P.H. Mason, I. Kerridge, and J.P.A. Ioannidis. 2017. Ethics and epistemology in big data research. *Bioethical Inquiry* 14: 489–500.

Lockhart, T. 2000. *Moral Uncertainty and Its Consequences*. New York: Oxford University Press.

Malin, B.A., K. El Emam, and C.M. O'Keefe. 2013. Biomedical data privacy: problems, perspectives, and recent advances. *J Am Med Inform Assoc* 20, 1: 2–6.

Manson, N., and O. O'Neill. 2007. *Rethinking Informed Consent in Bioethics.* Cambridge University Press.

Mittelstadt, B.D., P. Allo, M. Taddeo, S. Wachter, and L. Floridi. 2016. The ethics of algorithms: mapping the debate. *Big Data & Society* 3, 2. doi: 10.1177/2053951716679679.

Mittelstadt, B.D., and L. Floridi. 2016. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics* 22, 2: 303–341. doi: 10.1007/s11948-015-9652-2.

Mulley, A., A. Coulter, M. Wolpert, T. Richards, and K. Abbasi. 2017. New approaches to measurement and management for high integrity health systems. *BMJ* 356: j1401. doi: 10.1136/bmj.j1401.

Pellegrino, E.D. and D.C. Thomasma. 1993. *The Virtues in Medical Practice.* New York: Oxford University Press.

Porter, M. 2009. A strategy for health care reform — toward a value-based system. *New England Journal of Medicine* 369: 109–112.

Prainsack, B. 2017. *Personalized medicine: empowered patients in the 21st Century?* NYU Press.

Regulation (EU) 2016/679 of the European Parliament and the Council. 2016. *Official Journal of the European Union*. L119/1–88.

Renn, O. 2008. White paper on risk governance: toward an integrative approach. In *Global Risk Governance. International Risk Governance Council Bookseries, Vol. 1*, eds. O. Renn & K.D. Walker, 3–73. Dordrecht: Springer.

Rieder, G., and J. Simon. 2017. Big data: a new empiricism and its epistemic and socio-political consequences. In *Berechenbarkeit der Welt? Philosophie und Wissenschaft im Zeitalter von Big Data,* eds. W. Pietsch, J. Wernecke, M. Ott, 85–105. Wiesbaden: Springer VS.

Ritzer, G., and N. Jurgenson. 2010. Production, Consumption, Prosumption. *Journal of Consumer Culture* 10: 13–36.

Sheaff, R. et al. 2015. Integration and continuity of primary care: polyclinics and alternatives – a patient-centred analysis of how organisation constrains care co-ordination. *Health Services and Delivery Research* 3, 35. doi: 10.3310/hsdr03350.

Stodden, V. 2010. The scientific method in practice: reproducibility in the computational sciences. MIT Sloan School Working Paper 4773–10.

Toffler, A. 1980. *The Third Wave.* New York: William Morrow.

Tutton, R. 2004. Person, property and gift: exploring languages of tissue donation to biomedical research. In *Genetic databases: socio-ethical issues in the collection and use of DNA,* eds. R. Tutton & O. Corrigan, 19–38. London: Routledge.

Van den Broek, T., and A.F. van Veenstra. 2018. Governance of big data collaborations: how to balance regulatory compliance and disruptive innovation. *Technological Forecasting and Social Change* 129: 330–338.

Vayena, E. and A. Blasimme. 2018. Health research with big data: time for systemic oversight. *The Journal of Law, Medicine & Ethics* 46: 119–129.

Voerman, S.A. & P.J. Nickel. 2017. Sound trust and the ethics of telecare. *Journal of Medicine and Philosophy* 42: 33–49.

Wang, Y., and M. Kosinski. 2018. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology, 114, 2: 246–257*.

Wynne, B. 1992. Uncertainty and environmental learning. *Global Environmental Change* 2: 111–127.

Žižek, Slavoj. 2004. What Rumsfeld doesn't know that he knows about Abu Ghraib. *In These Times*. Accessed on 04-04-2018 at http://www.lacan.com/zizekrumsfeld.htm.

Zimmerman, M.J. 2008. *Living with Uncertainty: The Moral Significance of Ignorance.* Cambridge University Press.