Dilemmas 2015 Papers from the 18th annual International Conference Dilemmas for Human Services: Organizing, Designing and Managing

New ethical dilemmas arising from the growth of personal health

Margunn Aanestad[1,2]

[1]Department of Informatics, University of Oslo, Norway
[2]Department of Clinical Medicine, UIT Arctic University of Norway margunn@ifi.uio.no

Abstract

The aim of this paper is to encourage critical investigations of information governance related to personal health data. Presently, most societies have regulations that balance the protection of privacy of personal health data with the need to disclose personal data for public good. This balance seems to be challenged by recent technological and policy developments, and this paper seeks to illuminate some of these developments and the associated dilemmas that emerge. Firstly, the challenges of handling novel data types from new sources (such as smartphones, devices and sensors) is still poorly understood and regulated. Secondly, not only are medical and societal gains predicted from giving access to personal health data, but also expectations of economic value creation circulate. Thirdly, the emerging data-intensive scientific practices depend on data, both as raw material and for building the scientific toolbox (models, methodologies and repositories). Being observant on these developments can facilitate interventions into the ongoing "ethics-in-the-making" of these crucial processes.

Key words: personal health data, ethical dilemmas, open data, sharing, disclosure.

http://dx.doi.org/10.15626/dirc.2015.03

ISBN: 978-91-87925-74-0

Introduction: Balancing personal privacy and public good

The health information collected from individuals in diagnostic or treatment situations may be valuable beyond the immediate situation and for other parties than the directly involved participants. Sometimes clinical information is re-used for research purposes, and sometimes it is reported to national health authorities that conduct disease surveillance. Most societies have regulations that address the dilemma between protecting the person's privacy and disclosure of data for public good purposes. Sometimes the interests of the collective trump the interests of the individual. For instance, it is common that health authorities have legal rights to obtain data for disease surveillance and public health purposes. The data collected for these purposes are often anonymized and aggregated data, but can also be detailed, personal data, e.g. relating to specific diagnoses and diseases. This data collection typically happens through health providers, who are obliged to report incidences of diseases to designated health agencies or registries. Health registries in this category are frequently operating without requiring any explicit consent statement from the individual patient.

However, if researchers seek access to these data, this will be handled differently. The regulation of health-related research is often based on the Nuremberg Code from 1947, the Declaration of Geneva (1948) and the Helsinki declaration (1964). These regulations say, among other things that researchers who want to obtain data from registries, patient-related documentation, or directly from the patient, need to obtain patient consent (in most cases). The consent has to be informed and voluntary, as well as linked to the specific purpose of the research. The patient should be able to decline to participate, or pull out at a later point in time, without consequences to his or her treatment. Thus, in the context of research, the wellbeing of the individual trumps the interests of the society and science. When it comes to the regulations regarding the rights of employers and insurers to request personal health information, there is greater variation among countries' approaches. There is, however, usually in place an ethical and legal framework that stipulates a certain information governance regime. The argument of the present paper is that recent technological and policy developments seem to challenge these information governance regimes. We will briefly consider some empirical examples. First, new data types from novel sources generate new usage possibilities that emerge outside the current regulated space (both legally and socially). Secondly, the expectations for the economic value generated from 'Big Data' seem to challenge the pre-existing balance between public and private goods. Thirdly, the emergence of data-intensive science requires new modes of accumulation and usage of personal data in order to develop knowledge and methods. The paper will briefly present these examples and the dilemmas they give rise to. As these issues are significant and highly complex, the ambition of the paper is to give an initial indication of important trends and issues, rather than a comprehensive discussion.

New data sources and usage potentialities

This section argues that with the emergence of novel data sources and usage possibilities we may find ourselves in legally and socially unregulated spaces. Mobile phones and personal wearable devices capture extensive data sets, including location, altitude, sound, image, Bluetooth proximity, EEG, ECG, pulse, temperature and numerous other information elements. These devices may be used to monitor lifestyle and health-related activities (fitness, stress level, sleep pattern, nutrition, activity etc.) or to monitor disease through measuring biomedical and clinical parameters. In many cases these data are transmitted from the device to a receiver, such as an app on the person's smartphone, from where the data may again be uploaded to a remote storage solution. This data uploading can happen in the context of a service provision, where the user subscribe to e.g. a disease monitoring service, or it can happen without the user being aware of it, such as when smartwatch apps routinely deposit the captured data remotely. It is not always transparent which data is being captured, how it is transmitted, with whom it is shared or how it can be used (Hilts et al., 2016).

This asymmetry, which opens up for exploitation, is in itself problematic (Zuboff, 2015). In addition, when we know that such data are linked to health insurance programs, or used in court cases[1] it becomes pertinent to investigate the mechanism of transmission and specifically the possibility for interception and manipulation. In addition to the registration of personal data, these devices allow for other types of aggregate or population-level tracking, such as using the Bluetooth radio signal to track how people move around within a physical space (e.g. consumers in a shopping mall[2] or citizens and tourists in an urban space[3]). Despite the fact that the data collected do not contain identifiable information, this is still a type of usage that we currently are not accustomed to and which may create controversies.

Other examples of novel data sources that escape our current awareness are e.g. data from apps that automatically report technical data such as usage logs, patterns of user actions, and errors, to the app developer, in most cases with the

[1] Unknown Author. (2015). "Police charge woman for making up a rape after she was exposed by her own FitBit," News.Com.Au, June 24, 2015. http://www.news.com.au/lifestyle/police-charge-woman-for-making-up-a-rape-after-she-was-exposed-by-her-own-fitbit/story-fneszs56-1227412671705 and Christina Bonnington. (2014). "Data From Our Wearables Is Now Courtroom Fodder," Wired, December 12, 2014 http://www.wired.com/2014/12/wearables-in-court/

[2] McCarthy, Bill (2015). Using Location-Based Analytics to Understand the Customer Journey. ShopperTalk. http://www.shoppertrak.com/using-location-based-analytics-to-understand-the-customer-journey

[3] Paul Lewis. (2008). "Bluetooth is watching: secret study gives Bath a flavour of Big Brother," The Guardian, July 21, 2008 http://www.theguardian.com/uk/2008/jul/21/civilliberties.privacy

intention to improve the application and/or service. While the content data as well as metadata of e.g. a medical app will be subject to regulation, this other kind of data from the app is not regulated (Andersen, 2013). Similarly, the equipment in smart homes, e.g. digital locks, sensors, and control systems, collect data about the inhabitants and visitors of the house. These types of "infra-data" (a notion intended to differentiate them from the more well-known category of meta-data, or "data about data") seem currently to be below the horizon of regulators.

The economic potential of personal health data

Collecting and using personal health data is often justified by referring to the collective good. Traditionally this referred to public health issues, e.g. the need to conduct disease surveillance in order to detect causes of diseases and contain epidemics. However, recently the collection of personal health data also gets justified by referring to the potential socio-economic value of these data. The visions associated with "big open data" often predict entrepreneurial innovation with economic significance if data is "released" (i.e. openly published). Recent controversies in UK around the care.data scheme illustrate this extension of the justification for data collection.

In 2014, NHS England attempted to initiate a scheme of extended collection of data. Data had been collected from hospital records for a long time, and data from the Hospital Episode Statistics are published on the UK government's "Open Data Hotel" (data.gov.uk). The ambition of the care.data scheme was to also collect data from General Practitioners' (GPs) patient records, as this would enable analysis of patient trajectories and make available more complete information, e.g. on medication usage. In the data set to be extracted on a monthly basis from the GPs patient record system, the patient's name would not be included, but NHS number, birth date, postal code, gender, and ethnicity would be reported. In other words, the data would be classified as "re-identifiable". The GPs' obligation to maintain patients' privacy under the Data Protection Act had been suspended under a new law (the Health and Social Care Act 2012), which obliged them to report to the new Health and Social Care Information Centre. This law also bypassed the need to obtain patients' consent for 'secondary use' of information collected. As the start of the care.data scheme drew closer, a growing debate picked up, triggered by concerns with extraction of sensitive and re-identifiable information. It was also fuelled by the reluctance with which the NHS England allowed an opt-out option, as well as an ill-managed information campaign during January 2014. However, a core point in the controversies was the lack of information about how the third parties' access to data would be handled. The care.data solution was not meant to support the provision of care (this was the role of the Summary Care Record), but to facilitate research, planning and monitoring, both for the NHS and for third parties. Potential users of data would be organisations within the NHS (such as commissioning bodies) but also outside of the

NHS. Subject to approval, this could potentially be health charities, universities, hospital trusts, think-tanks, pharmaceutical companies and other private companies. In the policy documents multiple purposes of usage were indicated: health surveillance, quality assurance, audit, health service research, and health service planning. However, the details around the regulatory mechanisms (approval of eligibility) were not in place at the time of debate. Controversy peaked in late January and early February, until the NHS backed down on February 17th 2014 and decided to postpone the scheme. At the time of writing, it has not yet been restarted.

The 'Information Governance Assessment' (HSCIC, 2013) claims: "Opening up valuable data to external agencies is an important government policy, but many members of the public would be uneasy about private companies benefitting from their health data. The risks associated with the sharing the data of course need to be considered against the benefits to be achieved. Many more organisations will be able to make better use of valuable data offering potential benefits to the public." (ibid. p.6). There is also a discussion on whether this is acceptable in principle, where the authors conclude that "Access to such data can stimulate groundbreaking research, generate employment in the nation's biotechnology industry, and enable insurance companies to accurately calculate actuarial risk so as to offer fair premiums to its customers. Such outcomes are an important aim of Open Data, an important government policy initiative." (p. 5). We see that the public good arguments related to research is intermingled with arguments of socio-economic development and privatized profit-making, all hidden under the cloak of "innovation"[4]. Based on a critical examination of UK policy documents in the biosciences, Edward Hockings claims that "we are witnessing a shift from rights-based approach to the adjudication of competing claims, in which benefits to the economy, for example, are seen as goods to be balanced with a data subject's right to privacy and confidentiality (Hockings, 2016, p. 95).

The hunger for data in data-intensive sciences

The frequent referrals to the potential for scientific breakthrough from collecting personal data, point to very real and significant possibilities that recent technological and methodological advances have made possible. The availability of computational and analytic power is involved with what some calls a 'fourth paradigm' in science: the data-intensive research supersedes the experimental, theoretical and computational paradigms that have come before (see e.g. Hey et al., 2009). Having available large sets of e.g. clinical, imaging, and molecular data allows the exploration of previously unknown patterns that may point to new insights, as in

[4] See e.g. David Cameron's presentation of the scheme in 2011: http://www.bbc.com/news/uk-16026827

genome-wide association studies (GWAS) or predicting adverse drug events (Raghupathi and Raghupathi, 2014).

This mode of research poses challenges to the existing consent regimes. Researchers are required to ask for informed, explicit and voluntary consent from patients, whose data is to be used in research, implying that the research project's objective and methods should be described. As scientific practices moves away from the hypothesis-driven research model, and towards data-driven and exploratory analysis, both the directions pursued and the methods employed are difficult to predict. The emerging debate on the use of broad consent and dynamic consent (Williams et al., 2015) is indicative of the underlying shifts in the debates (Karlsen et al., 2011; Steinsbekk et al., 2013).

Another aspect of this development concerns the contests over the valuable data themselves. The ethical tensions in this domain go beyond the most frequently discussed themes of privacy and security, as models of ownership, reuse and sharing of data are changing. A recent report from the Nuffield Council on Bioethics points to "the faltering ability of conventional information governance measures to keep pace with these developments" as a significant problem (Nuffield Council on Bioethics, 2015, p. xvi). The governance of shared data infrastructures is challenging (Bilder et al., 2015). For instance, there are significant costs associated with maintaining the required data repositories, in particular if curation service or other data services are provided. The initial funding models may not scale together with growth in demand and usage, and there is a need to consider other revenue streams and business models to recover the costs of running the services (Berman and Cerf, 2013).

One empirical illustration of such challenges is the tussles around the global data repositories for sharing data on genetic variants associated with inherited risk of breast and ovarian cancer (variants in the *BRCA* genes). In 1995, ten scientists concerned with keeping information in the open created the Breast Information Core (BIC) database. At the time, Myriad Genetics had acquired patent rights and a test monopoly for BRCA tests in the US jurisdiction (a monopoly they had until the patent rights were invalidated in 2013). Myriad initially contributed to the BIC database, but decided to stop sharing information in 2004 and subsequently acquired competitive advantage based on accumulating their test results (i.e., variants found) in their own, well-curated and thus high-quality proprietary database. The wider community reacted to this enclosure of critical information by establishing the "Sharing Clinical Reports Project" and "Free the Data" initiatives which encouraged GPs and patients respectively, to disclose the results from the test they had undertaken. In addition, the US government supported the establishment of the open ClinVar repository. Also other open, 'walled garden' or public-private cooperation have arisen, as actors seek to negotiate the different value logics of research, clinical use, and commercial exploitation of data (Vassilakopoulou et al., 2016).

Conclusion

We are witnessing a period where technological, methodological and policy-related developments challenge the pre-existing information governance regimes. There are a number of empirical domains which will be core sites for tensions emerging from shifting relationships between the individual and the larger society. For instance, the rethinking and transformation of informed consent is one such theme. This happens both to cater for the more exploratory and less predictable trajectory of research, and to rein in the temptation to reuse available data without consent, including the novel types of data available from other sources than the traditional. In addition, the issues related to data protection, privacy, anonymity and confidentiality of health information will remain core areas of concern. The emerging models of ownership and exchange of personal health data will be another core topic for researchers to follow. While public data stewardship is a well-known model, also proposals for novel ownership forms emerge, such as patient-owned cooperatives (Hafen et al., 2014) or distributed initiatives such as MyData (Poikola et al., 2015); initiatives which address the right to control data and the right to benefit from data in different ways. These are some among a larger set of ethical dilemmas that should be of core concern to scientists (Mittelstadt and Floridi, 2016). It is pertinent that researchers engage in these processes of ”ethics-in-the-making” and investigate what is at stake for the shared information resources, for the society as well as for the individual citizens.

References

Andersen, K. H., 2013. *Mobil App med helseopplysninger – en mulighetsstudie innenfor det norske lovverket*. Master Thesis, University of Oslo, 2013.

Berman, F. and Cerf, V., 2013. Who will pay for public access to research data? *Science Magazine 09*, August 2013.

Bilder G, Lin J, Neylon C., 2015. *Principles for Open Scholarly Infrastructure-v1, retrieved* 5.10.16, http://dx.doi.org/10.6084/m9.figshare.1314859.

Hafen, E., Kossmann, D. and Brand, A., 2014. Health data cooperatives—citizen empowerment. *Methods Inf Med* 53(2), 82-86.

Health and Social Care Information Centre, 2013. Information Governance Assessment. Customer: NHS England – care.data addendum. Customer Requirement Reference Number: NIC-178106-MLSWX.A0913. Date: 29/08/2013. Version: v1.0

Hey, T., Tansley, S., and Tolle, K., 2009. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, Redmond, Washington, USA.

Hilts, A., Parsons, C. and Knockel, J., 2016, *Every Step You Fake: A Comparative Analysis of Fitness Tracker Privacy and Security*. Open Effect Report (2016).
Available at: https: //openeffect.ca/reports/Every_Step_You_Fake.pdf.

Hockings, E., 2016. A Critical Examination of Policy-Developments in Information Governance and the Biosciences. In Mittelstadt and Floridi (eds.): *The Ethics of Biomedical Big Data*, pp. 95-115, Springer, Switserland.

Karlsen, J.R., Solbakk, J.H. and Holm, S., 2011. *Ethical endgames: Broad consent for narrow interests: open consent for closed minds*. Cambridge Quarterly of Healthcare Ethics, 20(4), pp. 572-583.

Mittelstadt, B. and Floridi, L., 2016. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical contexts. In Mittelstadt and Floridi (eds.): *The Ethics of Biomedical Big Data*, pp. 445-480, Springer, Switserland.

Nuffield Council on Bioethics, 2015. *The collection, linking and use of data in biomedical research and health care: ethical issues.*

Paul, G. and Irvine, J., 2014. "Privacy Implications of Wearable Health Devices," *SIN '14 Proceedings of the 7th International Conference on Security of Information and Networks.*

Poikola, A., Kuikkaniemei, K. and Honko, H., 2015. *MyData – A Nordic Model for human-centered personal data management and processing*. Ministry of Transport and Communication, Finland, report 3/2015. ISBN: 978-952-243-455-5

Raghupathi, W. and Raghupathi, V., 2014. Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2(3), 1-10.

Steinsbekk, K.S., Myskja, B.K. and Solberg, B., 2013. Broad consent versus dynamic consent in biobank research: is passive participation an ethical problem? *European Journal of Human Genetics*, 21(9), pp.897-902.

Vassilakopoulou, P, Skorve, E. and Aanestad, M., 2016. Premises for Clinical Genetics Data Governance. Grappling with diverse value logics. In Mittelstadt and Floridi (eds.): *The Ethics of Biomedical Big Data*, pp. 239-256. Springer.

Williams, H., Spencer, K., Sanders, C., Lund, D. Whitley, E.A., Kaye, J. and Dixon, W.G., 2015. Dynamic consent: a possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. *JMIR Medical Informatics*, 3 (1).

Zuboff, S., 2015. Big Other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30, 2015