3-14-2016

Selected Issues Concerning the Ethical Use of Big Data Health Analytics

Lieke Jetten
Deloitte Risk Services B.V.

Stephen Sharon
Deloitte & Touche LLP

Follow this and additional works at: https://scholarlycommons.law.wlu.edu/wlulr-online

Part of the Health Law and Policy Commons

Selected Issues Concerning the Ethical Use of Big Data Health Analytics[*]

Lieke Jetten[**] & Stephen Sharon[***]

Table of Contents

I. Introduction

A. Big Data Generally

The terms big data and big data analytics originally derive from the terms artificial intelligence, business intelligence, and business analytics; terms used in the 1950s, 1990s, and 2000s, respectively.[1] Although some argue that big data is an ambiguous term used for many concepts,[2] most definitions share a common theme, succinctly summarized as, "large pools of data that can be captured, communicated, aggregated, stored, and analyzed."[3] Other definitions describe big data using four, or even five, "Vs": (1) Volume—some data sets are measured in exabytes per day; (2) Velocity—data are collected in near real-time; (3) Variety—data may be images, video, text, meta-data, audio, location information, etc.; (4) Veracity—data are unpredictable and difficult to de-identify; and (5) Variability—data come in peaks

---

1. *See* Hsinchun Chen, Roger H. L. Chiang & Veda C. Storey, *Business Intelligence and Analytics: From Big Data to Big Impact*, 36 MIS Q. 1165, 1166 (2012) (discussing the use of business intelligence and business analytics terms).

2. *See* Michael Schroeck et al., *Analytics: The Real-World Use of Big Data*, IBM GLOBAL BUS. SERVS. (Oct. 2012), https://www.ibm.com/smarterplanet/global/files/se__sv_se__intelligence__Analyt ics_-_The_real-world_use_of_big_data.pdf (last visited Mar. 9, 2016) ("'Big data'—which admittedly means many things to many people—is no longer confined to the realm of technology.") (on file with the Washington and Lee Law Review).

3. James Manyika et al., *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, MCKINSEY GLOBAL INSTITUTE iv (June 2011), http://www.mckinsey.com/insights/business_technology/big_data_the_next_front ier_for_innovation (last visited Mar. 9, 2016) (on file with the Washington and Lee Law Review).

and valleys.[4] Smart algorithms can identify and predict behavior of target groups and give insight into real-time events. This permits decision-makers to rely on evidence using big data instead of having to rely on intuition.[5] Besides the benefits of big data, however, a simultaneously growing concern about use of (personal) data arises. According to a European Union commissioned report, Internet users are increasingly being monitored without giving personal data, through the use of digital cookies, electronic identifiers left on their browsers, or through their Internet Protocol (IP) addresses.[6] Monitoring is not limited to the Internet, however. Mobile phones, cameras, payments, store loyalty cards, biometrics, interactive services, and social media networks can all be used to capture data and monitor users. Despite these serious concerns, common reason suggests that the benefits of big data, including those to manufacturing, healthcare treatment and research, communications, and transportation currently outweigh the drawbacks.

B. Special Considerations of Health Big Data

Privacy advocates have spent the better part of a decade teaching people that their data is precious, and that once it is online, it is online forever. As this message takes hold and users have finally started to limit the data they share online, big data initiatives are asking users to freely give up their data with no direct or immediate benefit.[7] Without transparency of big data

---

4. Andrew McAfee & Erik Brynjolfsson, *Big Data: The Management Revolution*, HARV. BUS. REV. (Oct. 2012), https://hbr.org/2012/10/big-data-the-management-revolution/ar (last visited Feb. 4, 2016) (on file with the Washington and Lee Law Review).

5. *See id.* ("We can target more-effective interventions [with Big Data], and can do so in areas that so far have been dominated by gut and intuition rather than by data and rigor.").

6. *See Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the European Union*, EUROPEAN COMM'N (June 2011), http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf (last visited Mar. 9, 2016) ("Online activities are particularly closely monitored.") (on file with the Washington and Lee Law Review).

7. *See id.* (discussing the ways in which user data is collected and the benefits—or lack thereof—that users receive).

practices, users may continue reducing the data they share, especially health data, because they do not understand the collective value of their data. It is in this environment that the review boards of today and tomorrow must operate.

Traditionally, internationally accepted frameworks have guided decision-makers and health professionals to decide what they should and should not do when dealing with health data.[8] This was necessary because health data has customarily warranted special protections.[9] In the era of rapid technology advancement, however, previously accepted frameworks are no longer sufficient for three reasons. First, innovation is outpacing the frameworks themselves, many of which reflect the world of data collection decades ago. Second, health data is increasingly being collected outside of traditional healthcare settings. Third,

---

8. These international frameworks include: Nuremberg Code (*available at* http://www.hhs.gov/ohrp/archive/nurcode.html); The Belmont Report (*available at* http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html); WMA Declaration of Helsinki (*available at* http://www.wma.net/en/30publications/10policies/b3/);
CIOMS/WHO International Ethical Guidelines from Biomedical Research Involving Human Subjects (*available at* http://www.recerca.uab.es/ceeah/docs/CIOMS.pdf); WHO Guidelines for Good Clinical Practice on Pharmaceutical Products (*available at* http://apps.who.int/medicinedocs/pdf/whozip13e/whozip13e.pdf); Council of Europe's Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine (*available at* http://www.coe.int/nl/web/conventions/full-list/-/conventions/treaty/164); UNESCO Universal Declaration of the Human in Genome and Human Rights (*available at* http://portal.unesco.org/en/ev.php-URL_ID=13177&URL_DO=DO_TOPIC&URL_SECTION=201.html); WHO Operational Guidelines for Ethics Committees That Review Biomedical Research (*available at* http://www.who.int/tdr/publications/documents/ethics.pdf); EU Parliament and Council of the European Union Directive 2001/20/EC on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Relating to the Implementation of Good Clinical Practice in the Conduct of Clinical Trials on Medicinal Products for Human Use (*available at* http://ec.europa.eu/health/files/eudralex/vol-1/dir_2001_20/dir_2001_20_en.pdf); OECD Privacy Principles (*available at* http://oecdprivacy.org/); FTC Fair Information Practice (*available at* https://www.ftc.gov/reports/privacy-online-fair-information-practices-electronic-marketplace-federal-trade-commission).

9. *See* 45 C.F.R. § 164 (2015) (outlining regulations to protect health data).

data are then shared with third parties not only for research, but also for commercial gain.

The new environment in which health data are collected supplement traditionally crucial questions such as: "How to secure data and limit access?," with new questions such as: "Should a software developer be held criminally responsible if a flaw in their software permits a malicious individual to kill someone by tampering with a remotely accessible insulin pump?" In this scenario, the insulin pump was connected to the Internet to permit remote control but also so its data could be shared and analyzed as part of a big data initiative. The benefits to the diabetic patient (and others) are real, but are they worth putting people's lives at risk? Big data initiatives in other industries typically do not implicate life or death decisions with the frequency of health big data.

C. Big Data Questions

Big data brings forth a multitude of proposed benefits but also raises equally powerful disincentives: (1) Big data may lead to the creation of better tools, services, and public goods, but will it do so at the expense of privacy incursions and invasive marketing?; (2) Big data may help us understand political movements and online communities, but will it also be used to track protesters and suppress freedom of speech?; and (3) Big data may positively transform how we study human communication and culture, but will it also narrow the range of research options and alter what "research" means?[10] Recent developments have raised new thought provoking questions, and this Article aims to discuss two of these unique questions that apply to the healthcare industry specifically:

Is it ethical to control or influence user behavior through health big data?

Should potentially life-saving information be shared with a patient who has consented to sharing data for an unrelated purpose?

---

10.   *See* Danah Boyd & Kate Crawford, *Critical Questions for Big Data*, 15 INFO., COMM. & SOC'Y 662, 663 (2012) (asking these questions).

II. Ethics of Controlling User Behavior

A. What Is the Risk?

Profiling and targeting of users through prescriptive analytics can start driving consumer behavior, rather than the other way around.[11] This means that if a consumer acts upon a personalized offer, more refined personalized offers will follow, which in turn slowly blocks the consumer from making other choices because they are less aware of alternatives. Removing options from consumer consideration interferes with the principle of free choice. This practice may be interesting commercially, but ethically, it is a slippery slope.

One example of how an awareness of data monitoring led to a change in user behavior occurred in the years following Edward Snowden's leaks of National Security Agency (NSA) monitoring capabilities and activities.[12] With knowledge that the NSA monitored user activity online, search terms for sensitive health keywords, such as "therapy," decreased.[13] People's personal space and right to control their own actions were violated. Further troubling is that the consequences of this violation extended beyond merely avoiding certain searches to people avoiding searching for help with their life-threatening mental illnesses, affecting them and the public.[14]

---

11. *See* Frank Buytendijk & Jay Heiser, *Confronting the Privacy and Ethical Risks of Big Data*, FIN. TIMES (Sept. 24, 2013), http://www.ft.com/cms/s/0/105e30a4-2549-11e3-b349-00144feab7de.html#axzz3zLchx46s (last visited Feb. 5, 2016) ("[T]he profiling and associated prescriptive analytics start driving customer behaviour, rather than the other way around.") (on file with the Washington and Lee Law Review).

12. *See id.* ("One result of NSA whistleblower Edward Snowden has been the global surge in discussion about privacy and big data.").

13. *See generally* Alex Marthews & Catherine Tucker, Government Surveillance and Internet Search Behavior (Apr. 29, 2015) (unpublished manuscript), https://papers.ssrn.com/sol3/Data_Integrity_Notice.cfm?abid=2412564 (last visited Mar. 9, 2016) (discussing the drop in certain search terms after the Snowden disclosures) (on file with author).

14. *See generally id.* (noting consequences from the reduction of certain searches).

From pedometers to light bulbs, the private sphere in which people are tracked and monitored is rapidly increasing.[15] As the number and variety of Internet-enabled devices increases, wherein even furniture can interact with people, a slew of new data is created, just waiting to be mined or exploited.

Paranoia and conspiracy theories aside, there is a possibility that people will become risk-averse and modify their behavior, as continuous monitoring leads to feelings of "creepiness."[16] Even if behaviors do not change, that may indicate a significant change in sociological norms, namely, desensitization to being monitored.

Controlling user behavior comes in another less obvious form. A benefit of the big data trend for organizations and research institutions is that they may profit from the performance of medical and imaging procedures, either as it enables them to supply medical monitoring devices—for example, to monitor a patient's vital signs after surgery—or to gain insight from the acquired data, which can have commercial value. They therefore have a financial incentive to perform procedures which may contribute to research but which are not medically necessary for the patient.[17] Researchers seeking to increase their data set, and institutions seeking to maximize profits, may consciously or subconsciously influence patient behavior. By definition,

---

15.   *See Internet of Things: Privacy & Security in a Connected World*, FED. TRADE COMM. (Jan. 2015), https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf (last visited Mar. 9, 2016) ("The IoT [Internet of Things] explosion is already around us, in the form of wearable computers, smart health trackers, connected smoke detectors and light bulbs, essentially any other Internet-connected device that is not a mobile phone, tablet, or traditional computer.") (on file with the Washington and Lee Law Review).

16.   *See* Mike Gualtieri, *Digital Creepiness: How Not To Spook Your Customers*, FORRESTER (July 7, 2014), https://www.forrester.com/Digital+Creepiness+How+Not+To+Spook+Your+Cust omers/fulltext/-/E-RES117347 (last visited Mar. 9, 2016) (defining creepiness as "[a] feeling by a customer that a digital experience offered by a company knows more about them than they should and is using that information in a way that makes the customer feel uncomfortable") (on file with the Washington and Lee Law Review).

17.   *See* Paul R. Helft & Christopher K. Daugherty, *Are We Taking Without Giving in Return? The Ethics of Research-Related Biopsies and the Benefits of Clinical Trial Participation*, 24 J. CLINICAL ONCOLOGY 4793, 4794 (2006) (questioning whether procedures as a part of research are always necessary).

unnecessary medical procedures expose patients to the unnecessary risks of those procedures.

Using big data, not the needs of the patients themselves, as the driver to initiate medical treatment shocks the collective conscience of the medical field and shatters its foundations in the Hippocratic Oath. But what if the equation changes so that the patients provide informed consent—a subject to lengthy discussion on its own[18]—to the potentially superfluous procedures and even receive payment in exchange for their participation?[19] This gray area is the emerging domain of the ethics board.

B. The Role of an Ethics Board

The Privacy and Ethics Review Board (PERB) has a responsibility to spot the issues raised above and should aim to identify and mitigate risks to consumers. Before the PERB can meet to achieve this lofty objective, its members must be chosen. Members can and should come from different backgrounds with different motivational interests, but this means their selection and balance is particularly important. A big data code of ethics, stating the values an organization commits to upholding, unites the employees, contractors, and PERB members and fosters a more consistent and fair PERB. While involving privacy and ethics specialists is an obvious start, involving the primary groups likely to be affected by the choices being made, such as the consumers from whom data is collected and industry professionals, is also advised. Consumers should be empowered with the ability to review data collected about them, have an

18. *See generally* Frederik Zuiderveen Borgesius, Behavioural Sciences and the Regulation of Privacy on the Internet (Oct. 2014) (unpublished manuscript), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2513771 (last visited Mar. 9, 2016) (discussing informed consent as a way to protect patient privacy) (on file with author); *Unified Ethical Frame for Big Data Analysis*, INFO. ACCOUNTABILITY FOUND. (March 2015), http://informationaccountability.org/wp-content/uploads/IAF-Unified-Ethical-Frame.pdf (last visited Mar. 9, 2016) (same) (on file with the Washington and Lee Law Review).

19. Hugo Greenhalgh, *Sell Your Own Personal Data*, FIN. TIMES (Oct. 16, 2015), http://www.ft.com/cms/s/0/8e51ecd4-7327-11e5-bdb1-e6e4767162cc.html (last visited Feb. 5, 2016) (on file with the Washington and Lee Law Review).

ability to withdraw it at any time, and should receive a clear benefit from sharing their data. In the scenario of unnecessary medical procedures, the threshold for informed consent should be raised, and counsel should also be consulted.

III. Communicating Unexpected Findings to Data Subjects

The second ethical question this Article explores is whether to communicate unexpected findings to data subjects. For example, when analysis on fetal stem cells in prenatal testing reveals that the mother has cancer, should she be informed of her condition? Should the answer depend on the religion or ethnicity of the patient? Consider that in Judaism, one is obligated to share medical information that may save a life,[20] but in China, the cultural norm is not to inform an ill patient of bad news because they are thought of as a child that needs to be protected by avoiding upsetting news.[21]

Even if a principal investigator wants to inform a data subject about the presence of disease, this might be difficult to do if the data has been de-identified in an effort to protect the subject. This creates a counterintuitive incentive to either not de-identify data, or to separately store data that would permit re-identification. Some may argue that the risk of loss of privacy for everyone in a study is worth the potential benefit of a subgroup of members. Unfortunately, this position may discourage participation and decrease the value of a study. These trade-offs need to be considered and acted upon accordingly.

One compromise to the scenario presented above would be to provide participants with a choice to opt-in to receiving incidental findings. This permits participants to retain control over their

---

20. *See* Eliyahu Fink, *Privacy and Government Snooping in Jewish Law*, ORTHODOX UNION (June 18, 2013), https://www.ou.org/life/news-op-ed/privacy-and-government-snooping-in-jewish-law (last visited Feb. 5, 2016) (explaining the obligations to save another's life under Jewish law) (on file with the Washington and Lee Law Review).

21. *See* Jessica H. Muller & Brian Desmond, *Cross-Cultural Medicine a Decade Later: Ethical Dilemmas in a Cross-Cultural Context*, 157 WEST J. MED. 323, 325 (1992) ("Because a patient is suffering already from the illness, it is unnecessary to make them suffer even more by discussing the reality of the disease.").

data without impacting the security of every participant's data and permits religious and ethnic preferences to be taken into account. It also avoids having to rely on a PERB to make this decision on behalf of participants. Another option to minimize risk is to have a third party store the re-identification information.

IV. How Do Other Industries Handle Similar Ethical Issues?

Ethical conundrums are not unique to the healthcare industry. One source of valuable guidance is to explore what other industries are doing in the face of similar challenges. For years, people have been asking, "Should an autonomous car seeking to avoid a collision swerve and hit a person or maintain course and risk serious injury to the driver?" What if swerving would hit multiple people, and the driver is alone? In the case of autonomous cars, these choices are programmed at the factory, in a vacuum, devoid of unique circumstantial inputs.[22]

Pilot programs are bridging the gap between the automotive and health industries through car seats with health sensors that can be used to detect when the driver has a heart attack or is falling asleep. The benefits of a car that can drive its occupant to the hospital are indisputable, but they also raise the question of whether these benefits outweigh the ethical and privacy risks concerning misused data.

There will never be universal agreement with the programmatic choices built into our cars, but developing and utilizing "ethics-by-design" principles—in the same vein as "privacy-by-design"[23] principles—will at least make the decisions

---

22. *See* Will Knight, *How To Help Self-Driving Cars Make Ethical Decisions*, MIT TECH. REV. (July 29, 2015), http://www.technologyreview.com/news/539731/how-to-help-self-driving-cars-make-ethical-decisions (last visited Feb. 5, 2015) (discussing the ethics of self-driving cars) (on file with the Washington and Lee Law Review).

23. "Privacy by design" refers to the principle of building in privacy features or decisions into early stages of design, to avoid time-consuming fixes afterwards when violations occur. *Introduction to PbD*, INFO. & PRIVACY COMM'N ONT., https://www.ipc.on.ca/english/privacy/introduction-to-pbd/ (last visited Feb. 5, 2016) (on file with the Washington and Lee Law Review).

more understandable, if not more fair. Supporters of this concept further argue that ethical considerations should not be the exclusive domain of ethicists, but rather should be incorporated at every phase and in every aspect of design.[24] Others advocate for the inclusion of philosophers to assist with the ethical investigations.[25] Health big data practitioners must be equally forward-thinking or risk potential government regulation if their PERBs are ineffective.

V. Conclusion and What's Next?

A. What Are the Issues We Have Not Thought About Yet?

With the onslaught of Internet-enabled devices in the past few years, new questions about how to protect consumer privacy and how to ethically use data have arisen frequently. This rapid evolution requires constant innovation in how PERBs operate. Checklists and questionnaires are no longer sufficient to identify the majority of risks and concerns. PERBs must be prepared to innovate as quickly as the industries they seek to examine.

B. How Will Privacy and Ethics Review Boards Need To Evolve?

Responsible data stewardship and effective management of ethical and legal risks for big data initiatives can only be achieved with a well-designed PERB that operates nimbly and transparently. It must morally assess new data initiatives by providing a predictive and all-encompassing framework of how various involved stakeholder groups may respond and the possible moral and ethical challenges that need to be overcome.

---

24. *See* Laurel D. Riek & Don Howard, A Code of Ethics for the Human-Robot Interaction Profession (2014) (unpublished manuscript), http://papers.laurelriek.org/a-code-of-ethics-for-the-human-robot-interaction-profession-riek-howard.pdf (last visited Mar. 9, 2016) ("The general public and professional ethicists will surely have input, but it would be best if attention to HRI ethics began within the practitioner community in order to facilitate the incorporation of ethical perspectives in every phase of HRI research, development, and marketing.") (on file with author).

25. *See* Knight, *supra* note 22 (arguing in favor of philosopher input to the ethical dilemmas of self-driving cars).

The structure and composition of this board is the first step toward engaging stakeholders and fostering independence. The guidelines for the similarly purposed Institutional Review Boards (IRBs)[26] are a good start, but without modification, they may fail in this new arena. A standard IRB, consisting of at least five individuals, contains: (1) at least one scientific member; (2) at least one non-scientific member; and (3) at least one member not affiliated with the organization.[27] An ideal PERB builds on this model by including underrepresented groups needed for an independent and effective judgment.

The PERB should be a stable internal body, containing external stakeholders involved on an as-needed basis. This is necessary to establish that affected parties are sufficiently represented and to establish that appropriate specialists are consulted for previously un-thought-of issues. The PERB should therefore include: (a) at least one lawyer (internal/external); (b) at least one business representative (internal); (c) at least one representative of a key stakeholder group, in order to safeguard their engagement and provide feedback in the moral deliberation process (external); (d) one Secretary to moderate a structured conversation (internal); and (e) an open seat, called "markets," in which market or country representatives sit, in order to map out cultural or national differences in moral judgments and ways to reach consensus (internal/external).

C. Final Thoughts

Big data promises abundant opportunities for added value. The potential annual value of big data to Europe's public sector administration was calculated to be €250 billion.[28] Using big data, managers successfully rely on evidence, rather than intuition, permitting companies to predict consumer behavior and

---

26. *See Institutional Review Board Guidebook*, OFF. HUM. RESEARCH PROTECTION (1993), http://www.hhs.gov/ohrp/archive/irb/irb_guidebook.htm (last visited Feb. 5, 2016) (providing IRB guidelines for protecting research participants) (on file with the Washington and Lee Law Review).

27. *Id.*

28. Manyika et al., *supra* note 3, at preface.

health risks in ways previously thought impossible.[29] Besides the undisputable benefits of big data, however, there is a simultaneously growing concern over the use of personal data, especially personal health data.

As the examples above and elsewhere illustrate, the risks of big data extend beyond feelings of creepiness, to life-and-death situations. Whereas creepiness may be overcome by both educating and engaging data subjects into moral data decisions, giving data owners back some degree of control,[30] and some other larger challenges need deeper consideration and discourse.

The future of ethical use of our data depends on an evolution from IRBs to PERBs containing a wider variety of interested stakeholders and specialties. This Article is a starting point to facilitate internal discussions regarding the structure, purpose, and challenges of a PERB. Success will depend on the engagement of affected stakeholders and the respect and weight given to their opinions.[31] Continuous attention to ethical concerns by everyone involved, not just the PERB, will strengthen the image of health big data and help to support realizing its full value.

---

29. *See* McAfee & Brynjolfsson, *supra* note 4 (explaining that businesses can "make better predictions and smarter decisions" using data).

30. Alistar, *Big Data Is Our Generation's Civil Rights Issue, and We Don't Know It*, SOLVE FOR INTERESTING (July 31, 2012, 12:40 PM), http://solveforinteresting.com/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it/ (last visited Feb. 5, 2016) ("Data does not invade people's lives. Lack of control over how it is used does.") (on file with the Washington and Lee Law Review).

31. *See generally* Göran Svensson & Greg Wood, *The Dynamics of Business Ethics: A Function of Time and Culture—Cases and Models*, 41 MGMT. DECISION 350 (2003) (providing examples of this).