2018

Health Data and Privacy in the Digital Era

Lawrence O. Gostin
Georgetown University Law Center, gostin@law.georgetown.edu

Sam F. Halabi
University of Missouri School of Law, halabis@missouri.edu

Kumanan Wilson
University of Ottawa - Ottawa Hospital, kwilson@ohri.ca

VIEWPOINT

Health Data and Privacy in the Digital Era

Lawrence O. Gostin, JD
O'Neill Institute for National and Global Health Law, Georgetown University, Washington, DC.

Sam F. Halabi, JD, MPhil
University of Missouri School of Law, Columbia; and Centre for Health Law, Policy, and Ethics, University of Ottawa, Ottawa, Ontario, Canada.

Kumanan Wilson, MD, MSc
Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada.

Viewpoint pages 229 and 231

Corresponding Author: Lawrence O. Gostin, JD, O'Neill Institute for National and Global Health Law, Georgetown University Law Center, 600 New Jersey Ave NW, Washington, DC 20001 (gostin@law.georgetown.edu).

In 2010, the social networking site Facebook launched a platform allowing private companies to request users' permission to access personal data. Few users were aware of the platform, which was integrated into Facebook's terms of service. In 2014, Cambridge Analytica, a UK-based political consulting firm, developed a data-harvesting app. That app prompted Facebook users to provide psychological profiles, including responses such as "I get upset easily" and "I have frequent mood-swings" as part of a "research project."[1]

The Facebook platform allowed users to share their friends' data as well, enabling Cambridge Analytica to access tens of millions of personal profiles, identifying voters' political preferences. The controversy revealed risks to identifiable health data posed by social media and web services companies' practices. After the Cambridge Analytica controversy, Facebook suspended a project that aimed to link data about users' medical conditions with information about their social networks.

Individuals often reveal detailed, sensitive health information online. Through wearable devices, social media posts, traceable web searches, and online patient communities, users generate large volumes of health data. Although some individuals participate in online patient forums and wellness information sharing apps under their own names, others participate via pseudonyms, assuming their privacy is preserved. Many users believe their data will be shared only with those they designate.

> Individuals' health data are now solicited, aggregated, analyzed, shared, and sold in ways poorly understood and largely unregulated.

Special Legal Status of Health Information

Personal health information (defined as identifiable data relating to past, present, or future physical or mental health) has a unique quality and is deserving of special legal protection.[2] Unauthorized disclosure of sensitive health information can be embarrassing and even result in discrimination.

The Health Insurance Portability and Accountability Act (HIPAA), enacted in 1996, safeguards the collection, storage, and disclosure of identifiable health data, but only for "covered entities," defined as health plans, clearinghouses, and health care entities and practitioners that electronically transmit health information.[3] Firms such as Facebook, Google, and Twitter are not "covered entities." HIPAA also does not protect deidentified data. Yet data anonymized by one source could be deanonymized when combined with data from other sources.[4] Various federal laws and regulations also safeguard health information privacy, including the Privacy Act, Common Rule, Substance Abuse Confidentiality Regulations, and the Genetic Information Nondiscrimination Act. These statutes, however, do not generally apply to online data. State laws similarly protect health information privacy but vary from state to state.[5]

Gaps and Inconsistencies in Legal Protection

Individuals' health data are now solicited, aggregated, analyzed, shared, and sold in ways poorly understood and largely unregulated. Federal and state laws, such as HIPAA, safeguarding data in clinical settings, health insurance, and research do not govern most internet health data.[6] Thus, private firms can effectively ascertain and then use health data for various purposes targeting consumers and patients based on profiles assembled from tracked user behavior, data purchased from other sources, and predictive analytics. Currently, major gaps and inconsistencies exist in health information privacy safeguards.

Who Is the "Formal Custodian" of Web-Based Health Data?

Because HIPAA extends only to covered entities, it does not apply to data collection by social media, wellness apps, and similar services. The Federal Trade Commission is the federal government's primary privacy and data security enforcement agency, but its Health Breach Notification Rule applies to firms that manage health data "primarily for the individual."[7] Most apps and websites fall outside this definition.

Artificial Distinctions

HIPAA creates artificial distinctions between data generated in clinical or health insurance settings and in online settings. The HIPAA rule affords robust protection for the former, but scant protection for the latter. Even for HIPAA-covered entities, the law does not reach deidentified health data, even as ways to reidentify individuals from other sources proliferate.[8] Web privacy is governed by "terms of service," which are opaque and poorly understood by users. In effect, sharing data is a *fait accompli*: the "price of entry" for the website. Terms of service often explicitly state that user information may be shared with advertisers and marketers. Yet there are few legal means to hold these data brokers as accountable as HIPAA-covered entities. Thus, if data brokers irresponsibly manage

sensitive health information, consumers have no effective means for recourse.

International Data Environment

Data collection, analysis, and transfer rarely occur within national borders. Instead, data are accumulated and used throughout the world. The US government negotiates privacy frameworks, which principally ask voluntarily participating firms to process data consistently with their terms of service.[9] The 2018 European Union General Data Protection Regulation offers more robust "data accountability," including transaction-by-transaction consent, limits on the scope of consent, and disclosure of "categories of recipients" to whom data may be transferred.[10] Yet, firms can circumvent General Data Protection Regulation rules through adroit construction of terms, location of user portals outside the European Union, and separation of entities providing services and collecting data. In addition, these rules will not prevent the illegal theft of data.

Fixing Federal Law and International Frameworks

Individuals are entitled to post personal health information on social media and other sites. Companies should be able to use these data, with the users' informed consent. Still, law reform should ensure that terms of service are transparent and comprehensible so that consumers can make an informed choice. Thus, users should readily be able to understand when companies, researchers, or clinicians seek access to their personal health information.

Data protection laws, therefore, should extend beyond health system settings, encompassing rapidly advancing data collection technologies. Technology enables data to be amassed, stored, matched, and analyzed for beneficial purposes. Massive data storage also can be vulnerable to cyberattacks and inadvertent release of sensitive data. Data may be used in novel ways through machine learning and artificial intelligence, exacerbating data privacy and security risks.

Reform need not be difficult. A Department of Health and Human Services task force developed guidance for ethical use of patient-generated health information apps that could be extended to all app developers and social media. Department of Health and Human Services recommendations include consent boxes based on models drawn from Food and Drug Administration Nutrition Facts Label and the Schumer Box for credit card disclosures. Apps or websites that gather and share personal health information, as defined by HIPAA, would be required to transparently specify the data to be used or disclosed; the entities disclosing and receiving those data; the expiration date of authorization; and the right to revoke authorization. The law should afford users an effective means to exercise their rights, without loss of service. Users who opt out of data collection terms could pay a reasonable fee to use the service.

HIPAA and other federal and state privacy laws are too focused on formal data custodians and data collected in the narrow contexts of treatment and medical research. The law should do more to affect companies that now collect and transfer personal health data as readily as HIPAA-covered entities. Doing so would allow individuals to share their information with greater awareness of downstream uses. At the same time, it would permit companies to use that information for everything from advertisements to wellness apps. The increased transparency would also foster public trust in emerging information technologies.

ARTICLE INFORMATION

Published Online: June 20, 2018. doi:10.1001/jama.2018.8374

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

REFERENCES

1. Funk M. Cambridge Analytica and the secret agenda of a Facebook quiz. *NY Times*. https://www.nytimes.com/2016/11/20/opinion/cambridge-analytica-facebook-quiz.html. Published November 19, 2016. Accessed May 20, 2018.

2. The Health Insurance Portability and Accountability Act of 1996 (HIPAA), Pub L No. 104-191, August 21, 1996; Security and Privacy 45 CFR Part 160.103 (2002).

3. Standards for Privacy of Individually Identifiable Health Information. Final Rule, 45 CFR parts 160 and 164 (2002).

4. Leaf C. The biggest share in the sharing economy. *Fortune*. http://fortune.com/2015/08/07/digital-health-data/. Published August 7, 2015. Accessed May 15, 2018.

5. Lee LM, Gostin LO. Ethical collection, storage, and use of public health data: a proposal for a national privacy protection. *JAMA*. 2009;302(1):82-84. doi:10.1001/jama.2009.958

6. Ostherr K, Borodina S, Bracken RC, Lotterman C, Storer E, Williams B. Trust and privacy in the context of user-generated health data. *Big Data Soc*. 2017;4(1):1-11. doi:10.1177/2053951717704673

7. Health Breach Notification Rule. Final Rule. 16 CFR Part 318 (2009).

8. Tanner A. How data brokers make money off your medical records. *Scientific American*. https://www.scientificamerican.com/article/how-data-brokers-make-money-off-your-medical-records/. Accessed May 18, 2018.

9. US Department of Commerce. Privacy shield framework. https://www.privacyshield.gov/Program-Overview. Published July 12, 2016. Accessed May 8, 2018.

10. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and repealing Directive 95/46/EC. https://www.eugdpr.org/. Accessed May 16, 2018.