2020

The Healthcare Privacy-Artificial Intelligence Impasse

Charlotte A. Tschider
Loyola University Chicago School of Law, ctschider@luc.edu

THE HEALTHCARE PRIVACY-ARTIFICIAL INTELLIGENCE IMPASSE

By Charlotte A. Tschider

INTRODUCTION

With the advent of the Internet, wireless technologies, advanced computing, and, ultimately, the integration of mobile devices into patient care, medical device technologies have revolutionized the healthcare sector.[1] What once was a highly personal, one-to-one relationship between physician and patient has now been expanded, including medical device manufacturers, third party healthcare system providers, even physician-as-a-service for interpreting the data complex systems churn out. The introduction of technology to the healthcare field has, at an ever-increasing rate, transformed human health management.

I. THE NEW CONSUMER HEALTH MARKETPLACE

Amongst rising healthcare costs, the easy access of the Internet, and a cultural desire to self-manage many health concerns, individuals—with the assistance of consumer electronics—are performing more traditional medical activities on their own.[2] For example, for cancer detection, there are now several apps that enable an individual to use their smart phone to detect a high probability of cancer simply by taking a picture of the affected area.[3] For example, UMSkinCheckApp is a skin cancer detection app created at the University of Michigan and available free to the public.[4] These apps usually are run on an artificial intelligence ("AI") infrastructure, commonly being fitted to power advanced computing applications, like those that can detect

---

[1] Nathan Cortez, *The Mobile Health Revolution?*, 47 UC DAVIS L. REV. 1173, 1177-78 (2014).

[2] Nicolas P. Terry, *Will the Internet of Things Disrupt Healthcare?*, 19 VAND. J. ENT. & TECH. L. 327, 328-30 (2016).

[3] Nicole Galan, *How Effective Are Skin Cancer Apps for Early Detection?*, MEDICALNEWSTODAY (2019), https://www.medicalnewstoday.com/articles/285751#popular-skin-cancer-apps.

[4] ROGEL CANCER CENTER, MICHIGAN MEDICINE, *UM SkinCheck App* (2020), https://www.rogelcancercenter.org/skin-cancer/melanoma/prevention/app.

cancer. In fact, AI in skin cancer diagnosis may actually be better than your physician at diagnosing skin cancer.[5]

Due to its tremendous capability for identifying probabilities for relational data sets, AI has been positioned to solve the most complex of health challenges. AI can revolutionize at-home care for older adults, reducing nursing home costs and improving quality of life.[6] AI can improve diagnostic processes for rural, generalist, or low-resourced doctors, whose patients may not have access to specialists.[7] AI can automate certain healthcare tasks or low-risk work, freeing busy physicians to focus more fully on complex cases.[8] AI can enable greater quality of life by untethering patients from chronic, time-consuming, and low-compliance tasks, such as managing diabetes.[9]

II. ARTIFICIAL INTELLIGENCE SOLVES HEALTH PROBLEMS AND CREATES NEW TENSIONS

The root of all of these AI promises, however, is data. AI lives and dies based on the data supplied to it. This means that AI, to deliver effectively on these promises, must be continuously provided high-quality, broad data sets.[10] It likely also means that probabilistic relationships between these data and AI functionality must be tested on an ongoing basis to harness the self-learning capability of unlocked machine learning, neural network, and deep learning applications.[11]

One point of significant challenge is the tension between AI data needs, which are both significant and often unknown at the time of collection, and individual data protection concerns. Although this short article will focus on privacy considerations, data protection reflects both individual privacy and

---

[5] Gigen Mammoser, *AI May Be Better at Detecting Skin Cancer Than Your Derm,* HEALTHLINE (JUNE 11, 2018), https://www.healthline.com/health-news/ai-may-be-better-at-detecting-skin-cancer-than-your-derm#1.

[6] Shourjya Sanyal, *How Is AI Revolutionizing Elderly Care*, FORBES (Oct. 31, 2018), https://www.forbes.com/sites/shourjyasanyal/2018/10/31/how-is-ai-revolutionizing-elderly-care/#61b5a239e07d.

[7] W. Nicholson Price II, *Medical AI and Contextual Bias,* HARV. J.L. & TECH 1, 18 (forthcoming, 2019).

[8] Thomas Davenport & Ravi Kalakota, *The Potential for Artificial Intelligence in Healthcare*, 6 FUTURE HEALTHCARE JOURNAL 94, 95-97 (2019).

[9] Jessica Kent, *Artificial Intelligence, Big Data to Improve Diabetes Management*, HEALTH IT ANALYTICS (Aug. 13, 2019), https://healthitanalytics.com/news/artificial-intelligence-big-data-to-improve-diabetes-management.

[10] Charlotte A. Tschider, *Deus ex Machina: Regulating Cybersecurity and Artificial Intelligence for Patients of the Future*, 5 SAVANNAH L. REV. 183 (2018).

[11] "Locked" algorithms are those that have been fixed at a point in time, usually after the algorithm has been tested, making them unable to organically self-learn.

the protection of personal information, as in cybersecurity where there are significant safety concerns for AI systems.[12] The reality is that AI technology safety and reliability goals complicate or frustrate data protection core elements. The nature of data storage and system infrastructure for AI application in healthcare creates a host of other challenges related to data protection.

III. PRIVACY GOALS AT CROSS-PURPOSES WITH AI

In information privacy, a number of key tenets have emerged from at least 20 years of active development in the international statutory space. Not only does the Health Insurance Portability and Accountability Act (HIPAA) in the United States mandate certain required steps with relation to Protected Health Information (PHI), a type of personal information statutorily defined, many omnibus international data protection laws similarly require certain steps be taken (and certain actions restricted) based on the privacy rights of a natural person.[13]

First, the most significant principle of privacy law is "data minimization."[14] Data minimization is the concept that data should not be collected, used, or stored more than is absolutely necessary. This concept is integrated into other traditional privacy concepts like use restrictions. Data protection laws require a privacy notice be displayed prior to data collection, and subsequent use of data must be consistent with what has been previously and specifically disclosed in the notice.[15]

AI creates major challenges for data minimization in that 1) data maximization, rather than minimization, is a central success factor for AI, 2) often AI data scientists do not always know which data elements will be useful to make a specific decision, and 3) data previously collected can be tremendously useful for future AI uses. These three problems cannot be solved under existing HIPAA requirements. First, covered entities (organizations regulated under HIPAA) must only collect data necessary to treatment, payment, or healthcare operations.[16]

For any data collection or use outside these purposes by a covered entity or a third party of the covered entity, express authorization must be completed, requiring not only explicit consent from the patient but also an

---

[12] Katherine Wellington, *Cyberattacks on Medical Devices and Hospital Networks: Legal Gaps and Regulatory Solutions*, 30 SANTA CLARA HIGH TECH. L.J. 139, 188–90 (2014).

[13] COUNCIL ON FOREIGN RELATIONS, *Reforming the U.S. Approach to Data Protection and Privacy* (Jan. 30, 2018), https://www.cfr.org/report/reforming-us-approach-data-protection.

[14] 45 CFR 164.502(b), 164.514(d).

[15] 45 CFR 164.520.

[16] 45 CFR 164.506.

expiration date or event.[17] This presumes that data use will end. Further, data uses are restricted to those specifically communicated in the privacy notice or authorization, restricting downstream or different uses. In AI, often the nature of data use is not determined before collection because the algorithm itself determines which data are useful, not the data scientist.

One way around this issue is to remove identifying features of PHI so that data may be used more flexibly. When data are de-identified under HIPAA's de-identification safe harbor, data may be used as broadly as one wishes; even data sales are permitted.[18] However, AI may also be used to re-identify patients from seemingly de-identified data sets. As previously described, AI functions only when it is supplied a steady stream and volume of data.

Usually these data are collected in "big data" sets, or large aggregated databases stored at a cloud provider, a third party that can manage data of this volume. To fill out the database, frequently organizations will purchase other de-identified data sets, such as from insurers, or adopt other data sets that may be publicly available. When these data sets are combined and an AI utility runs, an individual patient may be reidentified to a significant probability, effectively undoing de-identification. HIPAA does not address the challenges of reidentification, only of de-identification.

Finally, data are tremendously useful. Without data, AI systems fail. This means that healthcare data will become absolutely central to the success and safety of AI tools, including everything from diagnostic apps to complex medical devices, like the artificial pancreas. The necessity of data creates unique challenges for downstream use: 1) data use may be limited in time, 2) data may be impossible to recollect, due to lack of contact or reidentifiability, and 3) data are valuable. The underlying function of historical privacy laws, like HIPAA, is that activities begin and end, rather than continue and change. For this reason, HIPAA demands that data use outside primary data collection purposes like treatment, payment, and healthcare operations, should be limited, explicitly agreed upon, and terminated upon some date or action.

This perspective is informed by contemplated data uses at the time the Privacy Rule was finally implemented: in 2002, typically data outside primary purposes was used for a discrete purpose, such as sending data for a

---

[17] 45 CFR 164.508.

[18] U.S. DEP'T OF HEALTH & HUMAN SERVICES, *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule* (Sept. 30, 2020), https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html.

second opinion, sending to an outside lab, referring to a third party, or for other limited purposes. Authorizations were not intended to last forever. If a time period terminates or data may be desired to be used for another purpose, it can be impossible or tremendously impractical to "reconsent" to different data uses or uses beyond the previous authorization's date or terminating event. For example, if data have been de-identified, they may have a probability of identification but may not be able to actually be directly identified where contact information is available. To reconsent, an organization would need a phone number or e-mail, which are not permitted as part of a HIPAA safe harbor de-identified data set under most circumstances. Further, even if such data is available, the data may be out of the control of the primary party after it has been de-identified.

Perhaps most important is the relative value of such data: without healthcare data from a variety of sources, AI applications will not be able to be created, and likely the goals of AI will be frustrated. Healthcare is one of the areas with the most AI opportunity, and fulfilling these goals means ready access and likely sharing of such data. Although individual privacy rights must be observed, this may mean a shift in how we think about individual privacy.

CONCLUSION

Reworking privacy commitments in an AI world is an important endeavor. It may mean that we reconceptualize what these rights must be against a broader data need. It will likely include investment in better approaches to reduced identifiability that protect patients while promoting data use and sharing that will not reidentify. It may also mean permitting, at least for AI, broader declarations in privacy notices that put patients on notice for the use of AI while also permitting broader use. Without an approach that balances both invention and patient protection, we cannot realize the incredible potential of AI in healthcare.