Beware Privacy Risks In Training AI Models With Health Data

**This article was originally published in Law360 Expert Analysis.*

Nov 09, 2023

Categories:

Blogs
Health Law Matters
Technology Blog

Authors:

Barbara Bennett
Neha M. Matta

Artificial intelligence models are instrumental in U.S. health care, promising better patient outcomes, new treatment discoveries, increased efficiency and more precise decision making. Health and wellness applications also are using AI to assist consumers and encourage them to improve their nutrition, exercise and lifestyle.

In each case, these models largely are powered by machine-learning algorithms. The algorithms are trained on and analyze vast amounts of data to generate insights, discoveries and enhanced personal care.

This training data may include health information protected by the Health Insurance Portability and Accountability Act, and other personal health and wellness data. The success of AI models trained with this health data depends on careful data privacy analyses.

AI and Training Data

Data — and lots of it — is key to achieving the promise of AI. While there are many risks to be evaluated with respect to using protected health information or consumer health data as AI training data, one of the most important is privacy.

Tension exists with the use of any health data, not only because U.S. privacy laws prioritize protecting an individual's health information and may limit data use, but also because AI models can be unpredictable and are not yet well understood.

For example, deidentified or anonymized data can be made identifiable when combined with other deidentified data and processed by a machine-learning algorithm. Output data unpredictably may contain identifiable health information.

The High Cost of Failing to Consider Privacy Laws

The failure to consider privacy laws can be costly. Businesses training AI models with health data, and those using models trained by third parties with health data, will be under the microscope of the U.S. Department of Health and Human Services and the Federal Trade Commission, not to mention subject to other federal and state enforcement authorities and consumer class actions.

For example, in early 2021, the FTC took action against Everalbum Inc., a photo storage and organization application, after it was found to have violated Section 5(a) of the FTC Act by applying facial recognition to users' photographs without obtaining their opt-in and informed consent.

The FTC ordered Everalbum to delete any facial recognition models or algorithms developed with, trained on, or derived from the improperly obtained photos or videos.

More recently, in March 2022, the FTC took action against WW International Inc., previously known as Weight Watchers, for utilizing algorithms developed with personal health and wellness information gathered through its Kurbo healthy eating app, targeting children as young as eight without parental consent.

The FTC's settlement required the company to pay a $1.5 million fine for violating the Children's Online Privacy Protection Rule, and to delete all illegally collected data and any work product algorithms created using that data.

The expectation is that HHS will follow the FTC's lead and use disgorgement as a remedy for similar HIPAA violations. Several states also are following in the FTC's footsteps as they enact stricter legislation to combat what they consider to be deceptive data practices.

Compliance Can Be Complex and Difficult

Protected Health Information

To develop or train an AI model with protected health information, a person or entity must be either a covered entity or business associate under HIPAA and, with limited exceptions, the purpose is

required to be the payment, treatment or health care operations of the applicable covered entity.

Covered entities may find that health care operations may be the most applicable avenue to justify an AI model trained on protected health information, but the analysis is not a simple one. There are complex lines drawn between research, product development and health care operations.

In addition, there are narrow exceptions to the HIPAA prohibition against combining data from more than one covered entity. Moreover, individuals have rights to, among other things, have their data amended and deleted, which can be difficult in an AI model context.

A business associate's rights are even more restricted than a covered entity's. Whether a covered entity or business associate may use protected health information to train AI models must be assessed on a complex case-by-case basis.

It may be more efficient for a covered entity or business associate to train AI applications with deidentified data. And even then, written permission is required for business associates from each applicable covered entity. There is a risk that the data will not remain deidentified when combined and processed with other deidentified data.

Consumer Health Data

The use of consumer health data to train AI models must be assessed through the lens of the FTC and state-specific regulations.

As noted above, over the past few years, the FTC has demonstrated a pattern of strict enforcement that includes the destruction of all algorithm systems and AI models that have been built or trained with consumer health data that the FTC found was collected or used contrary to applicable law.

In addition, to further bolster consumer health data protections, the FTC has recently made clear that it will use its unfairness authority to make companies aware that they cannot use those data to engage in practices like targeted advertising.

FBT Gibbons

The agency has found that notice and choice often fail to protect consumer privacy because people feel they do not have meaningful choice or are tricked into consent. The FTC also has created a framework and issued a firm policy statement to proactively address complex risks associated with unfair or deceptive collection or use of biometric data.

The FTC will extend these analyses to AI technologies to protect against businesses using data practices that it considers unfair or deceptive to obtain consumer consent.

State law can impose further requirements.

For example, Washington's My Health My Data Act specifically defines Consumer Health Data as including "information that is 'derived or extrapolated from non-health information,' including through machine learning and algorithms."

The MHMD Act imposes several privacy obligations on regulated entities, including, for example, that they must (1) obtain prior opt-in consumer consent for the collection of consumer health data, (2) obtain a separate consent for the sharing of consumer health data, with limited exceptions, and (3) maintain a consumer health data privacy policy that clearly and conspicuously sets forth the data sources, the manner of data collection and the intended data use.

Further, the MHMD Act uniquely provides consumers with an absolute right to their consumer health data, without exception, including the right to opt out of data use and demand the deletion of their data, including from all archives and backups.

Ensuring Future Compliance

Users of AI models developed or trained by third parties are advised to conduct proper diligence with respect to the type of health data that has been used to train these models.

Obtaining proper liability assurances, including indemnification provisions, concerning the training data is among the many steps to consider when managing this risk.

Even in the absence of securing these assurances, proper due diligence will allow businesses to be aware of the risks that they

have assumed and to protect themselves accordingly with necessary policies and procedures.

Developers of AI are also advised to be diligent about their right to use training data with a view toward potential future consequences, such as the disgorgement of data or the destruction of the model in its entirety, should their rights in the data be challenged successfully.

While this article explores only one component of privacy risk with respect to developing or utilizing AI, businesses can maximize their AI opportunities by reviewing and managing other privacy risk, as well as the many other areas of AI risk — including but not limited to security, intellectual property, labor and employment, transactional, and regulatory compliance.

For more information, please contact any attorney with our Artificial Intelligence or Health Care Innovation teams.

FBT Gibbons