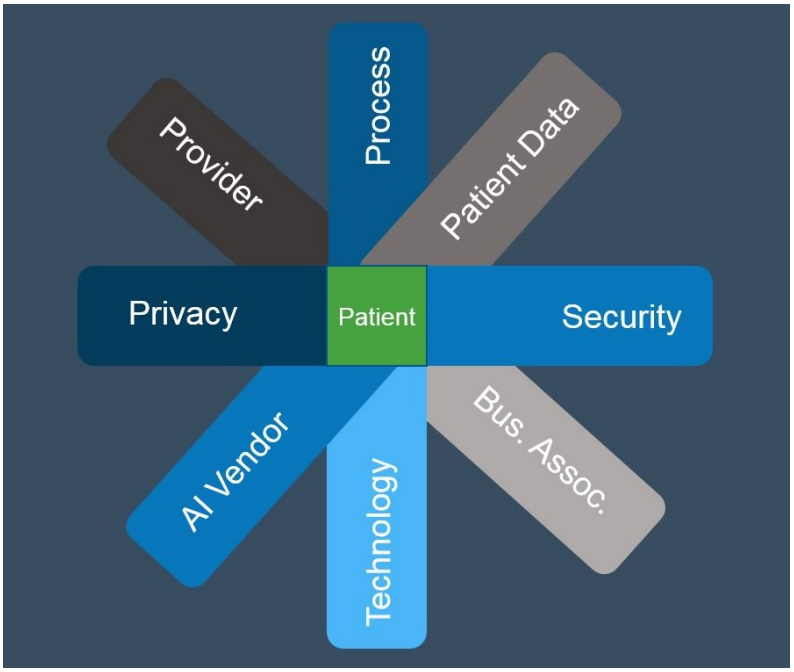
Data Empathy and AI (Part 2): Legal Theory Disruption and Ethics



AI Is Making Decisions

It is argued that AI programs may, by their nature, operate independently from the humans who programmed them and can make their own independent decisions, assessing patterns in large data sets and making predictions on those patterns, which adds another wrinkle in the equation of assessing responsibility for liability. The more robust and complex the AI tool, hypothetically, the more remote the distance becomes between the human who programmed the tool and the free-thinking tool itself - but is this not a fallacy – can the software actually enable itself independent of its creator. Unleashing the power of certain AI tools, programming such systems that learn from ingested inputs and can change and learn and modify itself on an ongoing basis, having its outcomes based on its artificial learning, makes the entire process unpredictable. Predictability, ironically, is a lynchpin of efficient jurisprudence and efficiency in the courts. Further, legal doctrine is largely based on human actors and the conduct flowing from those actors. An ever increasing autonomy inherent in AI, therefore, is muddling this doctrine.

When machines are programmed to ingest inputs and make independent decisions in the way in which it "thinks" through to the answer, and then creates the next version of that thought map by continuing to learn from the last decision it has made, in essence, relying now on its own inputs to make machine created data driven decisions, perplexing results and unpredictable outcomes may occur. How does one know that the model is behaving as one intended and that the interpretation is supportable, and does that person know why the model made a particular decision? This is a nightmare for the science of the law where predictability is a guiding light in risk mitigation and assessing liability. At the end of the day, legal concepts, such as causation, in relation to AI may prove to be so difficult to prove that it clogs the court systems, extends litigation and prevents consistent rulings across jurisdictions.

AI Status as a Person

AI is not a legal person. AI is also not the agent of a principal human (since an agent must also be a person) and therefore the fiduciary rich body of agency law cannot currently confer this status upon AI software. The question is - should AI tools be given the same status as corporations have been given under certain law, namely, the status as a "person?" When an AI system malfunctions or an error occurs, can the error be traceable back through a series of actions or causations to its root cause, or, will the trail be lost in the AI itself as the AI tool makes its own value decisions. If it is thinking like a person, should it not be treated like a person, not unlike corporate law divining the status of a person on a corporation? The question remains whether the person who created the AI program or the AI program, perhaps ten generations of thinking beyond the original programmer, can be held civilly or criminally liable for the action the AI program has taken? Further, from an evidentiary perspective, will the outcomes, decisions and predictions made by the AI tool be excluded or permitted to be used in a trial? Current state is clear: AI is not a legal person.

Bias and Discrimination

The manner in which AI is programmed could create bias, based on the creator of the code, namely, the programmer itself who may, even unknowingly, influence the software's tendencies to develop other biases (i.e., racial, age, gender, societal, etc.). This bias could target certain populations without regard to others; certain criteria in one hand while ignoring others. The data itself may contain bias - those inputs that can lead the tool to go in one direction, which it may not have gone in had a different set on exclusions and inclusions been taught to the AI tool. Has the resulting care been provided, or not, on the basis of race, gender or other protected status. Inputs may be tainted based on disparate treatment that is provided to certain individuals or groups. It is therefore important to understand the nebulous world of bias, to train programmers and scientists to defend against bias in building algorithms and to continue to hone AI tools to minimize its effects. Further, in the maintenance of these programs, a dedicated staff, intimate with the nuances of the code, must be employed in order to operate the AI tool efficiently, to mitigate against unintended consequences, to be the eyes and ears of the AI tool and to construct a better, defensible position should harm result from the AI tool.

Ethical Considerations

Ethical considerations pervade the provision of healthcare using AI tools. Can a provider get relevant data without being too invasive? Does the provider have enough data to input or know what metrics truly matter to build a functioning model? Are mature analytics in place? What is the status of the quality of the data? Is there a process to identify relevant data? Is the right data being utilized in order to achieve the results desired? Is the data complete? If AI is a black box, and it cannot demonstrate the paths to outcomes and conclusions, should it be utilized in patient care and at what level? Do patients have a right to know this path and right to understand the inherent risks and also results gleaned from his or her data used to make those decisions? Is there a mechanism to mitigate against harmful, incorrect, intentional and unintentional bias in the AI systems? Is the implementation of AI technology eroding patient primacy and autonomy? Once it has been determined legally that certain data *can* be used, the second question of *should* it be used must also be addressed. Thoughtful ethics, in short, must be a primary driver in building and implementing AI.

The Complexity of Explanation

Some AI tools are referred to on occasion as "black boxes" which, due to their nebulous and complex algorithms whose decisions are not easy to reverse engineer or easy to understand. However, someone, a human, created it. AI is, at its heart, source code and data. The more likely one is able to explain the mechanism and the nature of the AI tool the more likely one may be able to prove and defend its fairness in practice. This would require a mastery of the technology, a well thought out plan that has been executed, thoroughly tested and documented prior to production, and proactive monitoring on an ongoing basis to remain accountable to the results it creates. As seen below, explanability and interpretability has a direct impact upon certain legal theory including causation and foreseeability.

The Legal Framework

The laws and regulations that will govern AI will continue to be battle tested, evolve, take steps forward and eventually stabilize. As with any technology to legal theory equation, the law always lags behind and the gap grows larger as technology outpaces the reactive nature of the creation of law. Whether this regime relates to initiatives being pressed by the FDA nationally, or in conjunction with international parties, or another agency with dominion over these types of solutions, it will take time to settle through the risk benefit analysis to the greater good, ultimately helping to establish safety standards, best practices and potential safe harbors for certain uses. First, however, AI will claw its way through the world of first impressions, resulting in lawsuits that press at the corners of legal theory.

AI as a Regulated Device

The FDA is responsible for regulating the quality, performance and marketing of Devices and responsible for ensuring premarket approval of new Devices. The FDA is currently looking to define a common set of considerations related to Software as a Medical Device ("SaMD") and to articulate the final definition of SaMD. We have seen expansion over the years of what is considered a "Device" including Medical Mobile Devices and now, ultimately, SaMD. AI is SaMD when it is intended to treat, diagnose, cure, mitigate, or prevent disease or other conditions and are medical devices under the FD&C Act. The intended use of AI/ML-based SaMD is a critical determinative element. There will be a risk-based approach to categorize AI based on intended use and a spectrum of risk to patients and the level of regulatory control the FDA determines is needed to legally market the Device (similar to existing FDA theory – the higher the risk to patients the higher the scrutiny and regulatory involvement prior to hitting the market). Ultimately, regulators and agencies like the FDA, nationally and internationally, are crowd sourcing towards a consistent global regime that will take into account local nuances, but clearly, SaMD is in its sites.

Privacy, Security and Cyber

With a tool such as AI, true, authentic privacy runs the risk of becoming, outside of theory, non-existent. AI programs are used in every industry for tracking and predictions based on an individual's personal and societal preferences as well as where that individual travels, lives, drives, works, eats and breathes. The amount of data that is accumulated in this matrix of various AI tools pointed at the consumer is staggering.

In the healthcare context in the US, HIPAA is obviously paramount. The HIPAA Privacy Rule, sets standards for individuals' privacy rights and how their health information is used. It assures that individuals' health information is properly protected while allowing the flow of health information needed to provide and promote high quality health care, attempting to strike a balance that permits important uses of information, while protecting the privacy of patients. While patients have fundamental rights to privacy, the balance permits the provider to access higher volumes of data to paint the full, complete and accurate picture of the patient for further care or analysis. Further, the HIPAA Security Rule focuses primarily upon reasonable controls to ensure the confidentiality (means that e-PHI is not available or disclosed to unauthorized persons), the integrity (means that e-PHI is trustworthy, not altered or destroyed in an unauthorized manner) and availability (means that e-PHI is accessible and usable on demand by an authorized person) of PHI through administrative, technical and physical safeguards. In short, and this points towards data management program strategy, the more data stored in different places, the more points of entry exist. In opposite but equally perplexing, the higher the concentration of data in one centralized location, the higher the risk related to a single breach.

Liability

The lynchpin of liability, accountability under the law, will need to be carefully crafted and vetted. In fault based actions related to AI, the question is whether we have clear standards related to how fault will be proven? If there is an action taken or omitted in the generation of decisions in AI, and the output harms an individual, is it clear who the acting party is who should be held liable for the action – is it someone in the chain of treatment; the third party software company; the subcontractors who wrote work for hire code for the third party; the provider's data scientist; the provider's leadership; or the AI software, free thinking and evolving its intelligence over time. There are several theories of liability that may be disrupted by the advent of AI that are defined below.

Negligence/Malpractice and Standards of Care

Negligence, in general, relates to conduct that falls below an established legal standard for protecting another against unreasonable risk of harm (i.e., determined by the reasonably prudent man). In healthcare, physicians' actions are judged, not by a general reasonable standard, but against a reasonable physician standard who has the same expertise, skills and knowledge under similar circumstances (i.e., did the physician fail to comply with customary medical practices). Standards of care may advance with medical technology, as custom changes over time leaving physicians wondering what the actual standard of care could be for which they may face liability. Currently, is it customary for physicians to utilize AI to make clinical decisions? If not, then could the physician be seen as not comporting to current custom if harm results from the AI tool? Similarly, in ten years will it be customary, for instance, to use AI to reinforce a physician's interpretation? If the failure has to comport with current custom, and that custom changes over time, then the standard is likely to shift so that, when most physicians are using a certain AI tool in clinical practice, the use of AI may be seen as customary in the field and physicians will need to ensure compliance with that evolved standard. Further, we may reach a day when failing to use certain AI tools in a clinical setting, when the AI tool is more reliable than a human interpretation, is a breach of a new custom, creating liability by falling short of using an AI tool that has proven its staying power in medicine. Thus, AI marks an advance in medical technology that does not yet arguably have a set judicially accepted standard of care.

Causation and Foreseeability

By utilizing the theory of causation, fault and liability, in general, may be attributed and assessed according to an actor's actions or omissions. Determining fault leads to the assessment of compensation for injury whether under theories of tort, contractual breaches (and in limited ways, related to strict liability). The theory of causation, which factually demonstrates fault, therefore, is critical in the enablement of compensation for injury. Such causation needs to demonstrate a connection between the duty owed to the plaintiff, on one hand, and the damages on the other. Under tests like proximate cause, one will only be liable for risks, associated with conduct, that are reasonably foreseeable. Likewise, one should not be liable for results that were unforeseeable and for which one could not have known to act in a manner to limit the risk of harm. If the results of the conduct cannot be foreseen, then liability will not be warranted.

If AI, in a black box context, is reading patterns or producing results that are unforeseeable, or reaching conclusions that a human being could not have made, then it, arguably, cannot be held liable for those results. Neither can its programmer who

similarly doesn't understand the foreseeable outcomes. Logic would seem to dictate that someone making medical decisions he or she cannot justify through an explainable medical standard using a tool that will provide unpredictable and unforeseeable results is, counterintuitively, actually being negligent, but if the risk of harm is unforeseeable, and the black box nature of AI is unforeseeable, then there may be no recourse for resultant harm and the actor may be shielded in this cause of action. It is difficult to imagine this logic helping to build a legal framework of predictability to determine when a party should be held liable and when another should not be. Yet this is exactly with which we are now confronted – without foreseeability, as AI becomes more autonomous and distanced from its programmers, arguably, nobody will be liable for resultant harm under negligence theories since causation and foreseeability standards will not support the predictable assignment of fault.

Data Quality

Data quality should be viewed as one element in assessing one's duty of care. It is unreasonable to use bad data to make data driven decisions, particularly with sensitive populations or critical decisions such as those related to healthcare treatment. Therefore, with a duty of reasonable care in mind to select and utilize high reliable data to feed the AI programs, a provider will be in a much more defensible position should an unfortunate injury result from the AI tool. Further, what is important is the combination; the triangulation of provider employees entering data, the quality and richness of the data itself, and the ingestion of that data into AI applications – it is a cluster of symbiosis – each piece accommodating and enhancing the other; each becomes, in combination, the whole picture related to whether or not a standard of care was satisfied by the physician and provider treating patients using AI.

Compliance

When AI is unleashed, particularly related to patient centric care, it must be thoroughly tested and highly reliable. Documenting this testing and its outcomes in the test environments, as well as tracking its progress and evolution within the production environment will be critical for a provider, not only in defensibly demonstrating compliance with its own policies and current regulation, but also ethically related to the provision of healthcare. Lastly, how the tool was rolled out, monitored and maintained over time may be a critical element in determining compliance with a standard of care.

Respondeat Superior

Healthcare organizations may face vicarious liability as an employer for the tortious acts of an employee. In order to be liable, the employer must have demonstrated a failure to exercise due care in hiring, training, or supervising employees, or for failing to maintain adequate facilities and/or equipment. A provider that does not adequately train a physician to use AI or to explain the AI process risks, or implements a faulty or unreliable technology, or an otherwise malfunctioning AI Algorithm used by a physician in the practice of medicine, may all be arguments in favor of holding the provider liable for its physician's actions.

Products Liability

For healthcare organizations or physicians if a patient is injured by a product that is "not reasonably safe" due to a defective design, manufacture, or warning, the organization and/or the physician may face products liability challenges. What has remained a long standing element of this legal theory is that the liability relates to a tangible product (i.e., tangible personal property). In this context, in general, software is not tangible personal property and therefore is not subject to products liability actions. However, once software is embodied as a key ingredient of a tangible product, the question is open whether that interpretation can change and whether or not creative plaintiffs' counsel can start eroding this theory and claiming the software itself is the product subject to a products liability action. Clearly, the law will be tested in this direction as manufacturers continue to assert that software cannot be part of a products liability action.

There is also the issue of the "Learned Intermediary Doctrine" which holds that the physician has the primary responsibility of warning patients of the hazards of particular products, which includes AI tools. Otherwise, this duty, in general, resides with the manufacturer of the product. However, this intermediary step serves to shield the manufacturers from liability as the manufacturer will not be held to holding a duty to the patient as it has passed this duty on to the physician. Although there are several exceptions to this theory, one can see how manufacturers have a built in design incentive to actively engage physicians to ensure, in the event of harm and an adequate failure to warn of the risk, that they can push the liability to the physicians through this doctrine.

Further, regardless of this doctrine, how would either the manufacturer or the physician adequately warn about risks associated with AI results if neither fully understand the processing and decision making of the AI tool itself? If there is no way to adequately quantify the risk, then there is no way to actually articulate the risks to the patient, and therefore, in the

ensuing harm to the patient, it should be quite easy to state that the patient never received any adequate warning on the potential risks of harm related to the AI tool used in the practice of medicine.

Contract Law

AI, as software that ingests and uses data, will continue to be governed by the current body of software contract law when multiple parties are involved, for instance, in development, programming, licensing and hosting software (for instance, SaaS arrangements) and related to data restrictions and security. In this respect, commercial contracting should look to any nuances in the software being used in AI and the underlying data issues that are inherent in its use to develop the appropriate clauses between parties. Most of these clauses already exist within a commercial context and center on the software and the data itself.

Intellectual Property Law

For AI, there may be three bodies of law to review: patent law – whether the AI technology inventions are capable of patent protection; copyright law – whether the AI code or algorithm can seek protection under the US Copyright Act depending on its submission with other attributes and knowing, in general, that algorithms on their face, cannot seek copyright protection; and trade secret law – whether the know-how, process and methods can be categorized with independent value with a reasonable confidentiality protection in order to shelter under a trade secret umbrella. Further, the input data and the output data could be protected by the provider under applicable copyright law.

Data Governance and Management

Much of these theories and considerations all converge on how thoughtful an organization has been in globally thinking through, constructing, resourcing and holding accountable its data management program, which will include AI as a subset. What are the operational considerations to using AI, blending legal theory and ethics into operations? Is the organization cognizant of likely pitfalls in standards of care and foreseeability? Is the organization well informed related to potential risks to patients and causes of actions based on AI tools? How is AI rolled out, who is trained, what is monitored and how is the system maintained? How is the organization empowering its data and otherwise managing the data and using AI tools for the primary benefit of the patients? Is AI aligned with enterprise strategy and has the value of the asset been articulated? Has the organization developed AI uniformity and measurable standards? Lastly, has the organization evolved inspiring leadership in AI driven data practices and the promotion of accountability to the notions of patient primacy and Data Empathy in the provision of healthcare. If not, a promising set of solutions like AI tools may ultimately prove to be a nightmare that keeps on giving.

Kevin Michael Mooney, Esq.
Senior Director, Enterprise Data Governance
The Cleveland Clinic Foundation
216-312-4887
mooneyk@ccf.org