Algorithmic Bias and a Risk Management Framework

WHAT ARE THE RISKS OF ARTIFICIAL INTELLIGENCE/MACHINE LEARNING TOOLS?

We have all heard some of the troubling stories involving emerging tools powered by artificial intelligence (AI)/machine learning in which algorithms yield unintended, biased, or erroneous results. Here are a few examples:

- A selection app that prefers one gender over another or is not accessible to all
- A selection app that prefers certain backgrounds, education, or experience, with no showing of job relatedness or business necessity
- Facial recognition software that struggles with different skin tones
- An employment screening tool that does account for accents
- A clinical decision support tool for evaluating kidney disease that gives doctors inconsistent advice based on the patient's race
- Triage software that prioritizes one race over others

Organizations that use these AI tools do so at increasing legal, operational, and public relations risk.

AI-powered tools that are allowed to operate while being unchecked pose real risk because the output of the AI tool might not be as reliable as it seems on first blush. This could be because the data that the tool was trained on might not be appropriate for the population that is relevant to its intended use.
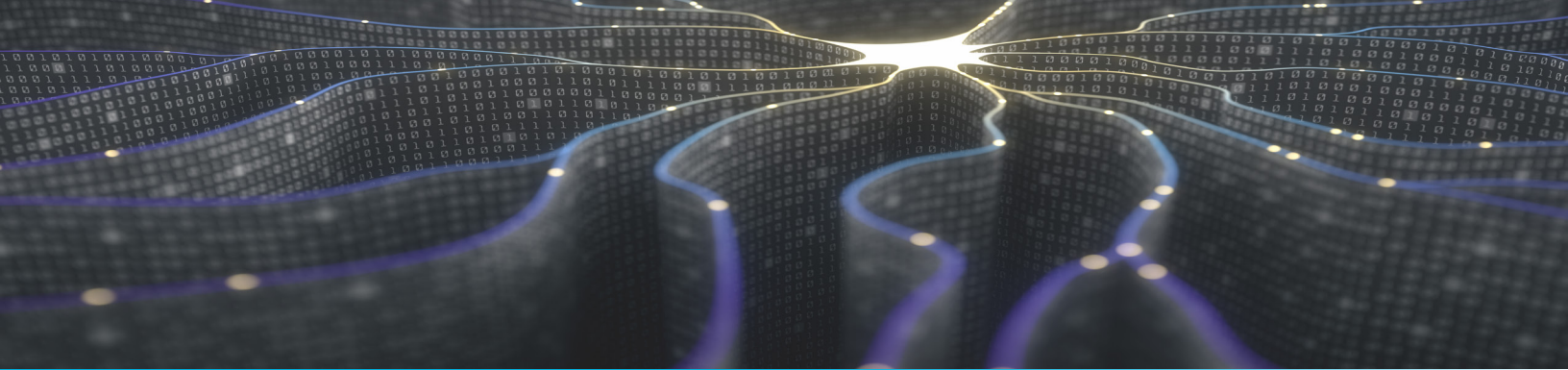
For health care organizations seeking to create or use these AI tools to benefit from the efficiencies, they first should consider creating a risk management framework to help manage these risks at the front end. Human governance is an important element for operating trustworthy AI tools. Unfortunately, the impact of algorithms on organizations and targeted populations is poorly understood and rarely measured at this time. There also is the absence of much regulation by government specific to AI. General privacy and consumer protection laws are relevant but not specific to these AI tools. Although that will be changing in the near term, organizations should be creating appropriate guardrails to show that you are a responsible organization that creates and uses only trustworthy AI tools.

WHAT DOES THE LEGAL LANDSCAPE FOR AI LOOK LIKE RIGHT NOW?

Local, state, and federal agencies are racing to implement regulations to address these issues with a common thread of identifying and mitigating bias. The U.S. Department of Health and Human Services (HHS) may soon require those using algorithms in health care to evaluate them for bias.[1] New York City requires certain AI-powered selection tools used at any stage of the employment life cycle, from recruitment to termination, to be audited for bias, and several municipalities are currently considering similar regulations. Most recently, the U.S. Equal Employment Opportunity Commission, the Consumer Financial Protection Bureau, the U.S. Department of Justice's Civil Rights Division, and the Federal Trade Commission (FTC) issued a "Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems," resolving to vigorously enforce their collective existing legal authorities and to monitor the development and use of automated systems. In the European Union, big tech companies will have to conduct annual audits of their AI systems beginning in 2024, and the upcoming AI Act will require audits of "high-risk" AI systems.[2]

1 https://jamanetwork.com/journals/jama/fullarticle/2800369.
2 https://www.technologyreview.com/2022/10/24/1062071/do-ai-systems-need-to-come-with-safety-warnings/.

WHAT IS NEEDED IS A FOCUSED AI EVALUATION BY A MULTIDISCIPLINARY TEAM!

Many companies rely solely on data scientists to find and fix problems in their AI tools. While data scientists are focused on making sure that the AI tool works from a technical perspective, they often lack the training and expertise to ensure that the AI tool is in compliance with the multitude of complex local, state, and federal anti-discrimination laws and regulations and other regulatory requirements that may apply to their algorithm. Nor can we expect data scientists to have the breadth of knowledge necessary to ensure that an algorithmic design is consistent with current law or whether the data fed to the AI tool has no fundamental flaws specific to its intended use that could create unintended bias.

Rather, there is a need for a multidisciplinary team of attorneys, social scientists, and data scientists to identify whether bias or discrimination may exist, to develop appropriate alternatives when they do, and to implement a best practices compliance framework going forward while ensuring the utility of the AI tool. This multidisciplinary approach reminds us of the way in which we address overpayment audits for Medicare billing providers. In that context, there is a need for a lawyer to advise on the law, a certified coder to advise on the technical aspects of billing and there is a need for a statistician to help with establishing a statistical sampling method to create an efficient way in which to refund moneys should there be an identified systemic error. That also is a multidisciplinary approach.

A proper risk management framework to manage potential bias concerns requires proactive measures throughout the machine-learning model's life cycle. This includes independent testing for the presence of bias before final deployment and periodic checks after to assure consistent compliance during use.

To measure the potential biases systematically and carefully, one might consider applying an evidence-based evaluation framework that relies upon the Artificial Intelligence Risk Management Framework developed by the National Institute for Standards and Technology (NIST) and its more specific guidance in assessing bias. In those materials, NIST highlights that the bias issue is a socio-technical challenge that requires a multidisciplinary approach to solve. We recommend that organizations embrace this view of having three distinct but integrated pillars:

- *Social scientists* who study the root causes of algorithmic discrimination as the starting point for investigation
- *Data scientists* well versed in the technical challenges of machine learning and designing systems for transparency and effectiveness; and
- *Attorneys* who can discern whether the output of an algorithm complies with applicable legal and regulatory standards while maintaining functionality and maximizing potential privilege assertions

We also recommend to our clients that organizations who intend to create or use AI tools should consider creating a compliance program for those tools that could be modeled after the seven elements of an effective compliance program – as stipulated by the DHHS Office of the Inspector General. At least the organization would have a defensive position should there be issues with the AI tool at a later time.

Let's embrace the benefits of these new AI tools to achieve the efficiencies well needed by our health care industry – but let's do so as corporately responsible organizations who care about the risks as well as the rewards.

Bradley Merrill Thompson

Chairman of the Board and Chief Data Scientist at EBG Advisors and
Member of the Firm at Epstein Becker Green

202.861.1817
bthompson@ebglaw.com

Having earned his Master of Applied Data Science in 2022 from the University of Michigan, where he received his law degree more than 35 years earlier, Brad stays up to date on the latest tools and metrics for assessing bias in machine-learning models.

Lynn Shapiro Snyder

Member of the Firm and Board of Directors

202.861.1806
lsnyder@ebglaw.com

Lynn Shapiro Snyder has helped health care clients successfully serve their communities and protect their investments for over 40 years at Epstein Becker Green. She deftly navigates clients through regulatory, reimbursement, compliance, enforcement, and policy challenges, including advising clients on establishing compliance programs to manage the enterprise risk of creating and using artificial intelligence tools in health care.