ALM | LAW.COM

Corporate Criminal Liability for Artificial Intelligence

By Leo Tsao, Robert Luskin and Corinne Lammers

May 21, 2024

The evolution of artificial intelligence (AI) continues at a torrid pace, with each new iteration exponentially more powerful than the last. What is now clear is that AI is much more than just another new technology, but is, instead, a transformative tool that promises fundamentally to change the way people live and work. Not wanting to be left behind, companies from across a wide range of industries are racing to find ways to adopt AI to make their businesses more productive, more efficient, and more profitable. At the same time, government regulators have taken note of not only the great potential benefits of AI, but also the tremendous damage that AI can cause. Several U.S. regulators have already started issuing policies and rules governing the safe use of AI to help prevent these potential negative consequences, such as fraud, discrimination, and data intrusions.

Questions remain, however, on how criminal law enforcement will be able to police the improper use of AI. That task will be particularly challenging where the legal violations result from AI-driven decision-making rather than intentional human actions.

The Department of Justice's Focus on Artificial Intelligence

Courtesy photos

L-R: Leo Tsao, Robert Luskin and Corinne Lammers of Paul Hastings.

The Department of Justice (DOJ) has already begun discussing how it will approach crimes committed with AI. In a February 2024 speech at Oxford University, Deputy Attorney General (DAG) Lisa Monaco explained that AI "may well be the most transformational technology we've confronted yet," but stated this did not mean that the DOJ needed to change its approach to investigating and prosecuting such crimes. As she explained, "new technologies don't necessarily demand new structures." The DAG analogized the current threats posed by AI to those associated with the arrival of the Internet, explaining that as criminals adopted the Internet, prosecutors had "evolved with the threat, applied and

adjusted existing legal tools, and added needed technology and expertise."

Yet, the DAG could not ignore the potential for the harm that criminal use of AI could cause. According to the DAG, AI was already "accelerating risks to our collective security" and "changing how crimes are committed and who commits them." As a first step, the DAG put criminals on notice that, going forward, DOJ prosecutors would "seek stiffer sentences for offenses made significantly more dangerous by the misuse of AI." One month later, at the American Bar Association's 39th National Institute on White Collar Crime, the DAG expanded on her prior remarks to address the use of AI by corporations. As she explained, where corporations deploy AI tools, the DOJ will expect them to adopt compliance programs that mitigate the risks posed by AI. As the DAG explained:

When our prosecutors assess a company's compliance program—as they do in all corporate resolutions—they consider how well the program mitigates the company's most significant risks. And for a growing number of businesses, that now includes the risk of misusing AI. That's why, going forward and wherever applicable, our prosecutors will assess a company's ability to manage AI-related risks as part of its overall compliance efforts.

The DAG tasked the DOJ's Criminal Division with updating its guidance on evaluation of corporate compliance programs to include an assessment of the risks associated with AI and other disruptive technologies.

Violations of Law Caused by AI

While there is no universally accepted definition of AI, at its most basic level, AI refers to a machine's independent ability to perform creative human-like cognitive functions, including reasoning, learning, interacting with the new information, and problem-solving. Thus, the primary value of AI is to be able to make its own decisions based on general programming that are otherwise free from human control. Of course, AI will not result in perfect decision-making, and the freedom given to AI systems to act without direct human control will invariably result in negative actions and consequences not intended by AI's developers. Even in the less sophisticated AI programs that have been deployed, we have already seen examples of AI decision-making causing violations of law. As companies entrust AI-based systems to handle increasingly complex and important functions, the risks posed by AI—and the potential damage that it can cause—will certainly grow. It is only a matter of time before the DOJ will be under pressure to prosecute corporations for violations of the law caused by their AI systems. What is less clear, however, is how the DOJ will be able to do so.

At the outset, it is important to make clear that this article is not focused on cases where AI is intentionally used as an instrumentality to carry out a criminal violation. In other words, where a human uses AI as a tool to commit a crime—just like a gun, computer, or telephone—the analysis of criminal liability does not present any novel issues because the focus remains on the actions of the individual. Indeed, the DAG explained as much in her speech this past March: "Fraud using AI is still fraud. Price fixing using AI is still price fixing. And manipulating markets using AI is still market manipulation." In this scenario, the existing enforcement framework for traditional crimes can easily be adapted to such AI-facilitated crimes. The much harder question is who is may be prosecuted where AI causes a violation of the law, but no human actor intended to commit that

offense, that is, where an AI system violates the law despite its developers and users having no intention to do so. This article is focused on that question.

The DOJ Will Face Significant Hurdles in Prosecuting AI-Based Crimes

Consider the following hypothetical: Bank A seeks to improve its transaction monitoring for compliance with U.S. economic sanctions laws by adopting an AI screening process. As developed and deployed, the AI screening system is intended to collect relevant information about the transaction (such as the parties, jurisdiction, and amounts), compare the information to applicable sanctions lists, conduct due diligence, make an assessment of the sanctions risk, and then decide whether to allow the transaction to proceed. Unbeknownst to—and unintended by any persons at Bank A—the AI screening system does an inadequate job of recognizing flags associated with certain North Korean entities, and allows thousands of transactions with those entities to be processed before the error is discovered and corrected. To establish a criminal violation of U.S. sanctions laws, the DOJ would have to prove that Bank A acted willfully (i.e., that the Bank knew about the law and intended to violate it). Because corporations typically can be held liable only vicariously—that is, through the criminal conduct of individual employees—the DOJ historically has investigated whether the individual employees and managers involved in the sanctions screening process were aware of red flags, and what actions they took as a result, to determine whether the evidence supports that those individuals responsible for the violations acted willfully. Liability could then be imputed to a corporation through the doctrine of respondeat superior—which allows corporations to be held criminally liable for the criminal acts of its agents where the agent acts within the scope of their authority and acts at least in part to benefit the corporation. Under respondeat superior, if any corporate agent acted with criminal intent and the other elements of respondeat superior liability are satisfied, a corporation is liable for a criminal violation of the sanctions laws.

But where the responsible decision-maker is a machine-based AI system—and the DOJ cannot prove that any individual acted with criminal intent—bringing a corporate prosecution for that charge becomes unsupportable under current law. That is because an AI system cannot act with the required criminal intent, or even if that were possible, it is hard to imagine how the DOJ could prove it. Moreover, even if those two hurdles could be overcome—under current law—corporate criminal liability cannot be based on the actions of an agent that is an artificial entity, but requires the crime to be committed by the actions of a human agent. Thus, while there may be corporate civil liability for a legal violation carried out by AI, it does not appear that the DOJ would have a viable legal basis to assert criminal liability for a corporation.

As companies continue to deploy AI systems to replace humans in making increasingly complex and important business decisions, this may potentially lead to a reduced role for the DOJ in fighting corporate misconduct. Indeed, the DOJ has already acknowledged that the current legal framework may not fully support the prosecution of AI-based crimes. Unless and until the law is changed—either through an act of Congress or by the courts in revisiting respondeat superior liability—the DOJ will face significant hurdles in prosecuting corporations for AI violations for

many, if not all, of the substantive offenses that it uses to prosecute corporations today.

How Will the DOJ Prosecute Companies for AI-Based Crimes?

That is not to say, however, that DOJ is left entirely without any options. Under the current enforcement framework, it appears that the DOJ can still pursue two potential avenues for prosecuting corporations for AI-based offenses: charging corporations with compliance-based crimes for failing to implement an AI system based on inadequate diligence and controls; charging corporations with strict liability offenses and or using the Park Doctrine to charge corporate officers responsible for the operation of the AI system. It is important for companies to understand these theories of prosecution, so they can avoid liability in the first place, and defend against a potential criminal charge when it arises.

A Compliance-Based Charge

Where the DOJ is unable to prove that an individual acted with criminal intent to violate a specific law, the most likely alternative theory of prosecution will be a compliance-based charge, such as a violation of the Foreign Corrupt Practice Act's (FCPA) internal controls provisions or the Bank Secrecy Act's (BSA) anti-money laundering (AML) provisions. For such a charge, the DOJ would only need to prove that an individual employee acted with criminal intent in failing to put legally required controls in place to ensure that the AI worked as intended, but not that anyone intended to use the AI to violate the law. Notably, this is a strategy that the DOJ already appears to use with success. For example, where the DOJ does not bring an anti-bribery charge under the FCPA—which requires the DOJ to prove that an employee intended to pay a bribe—it will often seek to charge a corporation with a violation of the FCPA's internal controls or books and records provisions. Such a charge could be available, for example, if the decision to approve a corrupt payment to a intermediary was made by an AI-based system, making an anti-bribery charge unavailable. Similarly, where the DOJ does not charge a financial institution with engaging in money laundering offenses—which requires the DOJ to prove that an employee knew the transactions were connected to illegal activity—it has charged the financial institution with a violation of the BSA for failing to have adequate AML provisions in place. Such a charge could be available, for example, if an AI-based system made decisions to allow certain illegal transactions to proceed.

In her most recent speech, the DAG signaled that the DOJ would use this compliance-based focus as a tool for addressing AI crimes. As previously noted, the DAG tasked the DOJ's Criminal Division with drafting guidance on the expectations for effective compliance programs related to AI. If a company falls short of having effective compliance programs in place to prevent an AI system from violating the law, it is reasonable to expect that the DOJ's investigation will include such compliance failures. While helpful for the DOJ, the scope of compliance-based crimes is limited for corporations. For example, the FCPA's internal control provisions only apply to U.S. public companies, and the BSA only applies to financial institutions. Congress could consider expanding the scope of such compliance violations to account for AI violations.

In any event, this theory of prosecution would still not allow the DOJ to charge a corporation where the AI system is designed, implemented, and tested in good faith, but nevertheless fails

to work as intended and violates the law in some unintended fashion. As is the case with other compliance-based crimes, the DOJ would still be required to prove that some individual possessed the requisite mens rea with respect to the compliance failures. Negligence or even incompetence in drafting or implementing an ineffective compliance program would not support a criminal prosecution.

Strict Liability Crimes

Another possible theory of prosecution is for the DOJ to focus on strict liability criminal offenses that do not require any proof that a corporate defendant acted with a criminal mens rea. For example, under the Food, Drug, and Cosmetic Act (FDCA), a company can be found guilty of certain violations even if it was unaware that it was doing so. Strict liability for criminal offenses also exist for other so-called public welfare laws, such as environmental laws. Such offenses are typically misdemeanors, but for companies, the potential forfeiture and other penalties can still be substantial. To the extent a violation of these strict liability statutes results from the failure of an AI system, a corporation could be prosecuted notwithstanding the lack of any criminal intent.

Certain individuals may also be subject to prosecution under a strict liability standard. In *United States v. Park*, the Supreme Court held that with respect to public welfare offenses, such as the FDCA, a corporate officer who has the power to prevent or correct violations may be prosecuted when violations occur, even where that officer lacks any affirmative knowledge and did not personally participate in the violations. Under the Park Doctrine—also called the Responsible Corporate Officer Doctrine—the DOJ can seek to hold certain "responsible" officers accountable for compliance failures even where that individual did not intend for those failures to occur. With respect to legal violations caused by AI, where the Park Doctrine is available, it may be possible for the DOJ to prosecute a responsible corporate officer even if the officer was not personally involved in or aware of the violation.

While strict liability offenses and the Park Doctrine generally have not been applied outside of the above-mentioned federal public welfare laws, Congress could consider expanding the scope of strict liability offenses to account for AI-based crimes.

What Are the Takeaways?

Corporations may be found criminally liable under the doctrine of respondeat superior only if an individual human agent commits a crime. As more corporations use AI to make decisions that were previously handled by humans, the DOJ may struggle to find individual actors upon whom to base corporate liability. Courts have already expressed concern that the use of AI may result in gaps in liability for violations of the law. Other solutions are on the table: Respondeat superior is a judicially imposed framework to address when criminal liability may be imputed to corporations based on the conduct of its employees. Just as the courts in the 19th century struggled to define when it is appropriate to impute the conduct of individual employees to their employers, the courts now might revisit the doctrine to address the challenges posed by AI actors. Alternatively, there has been discussion that Congress might act by either making individual actions unnecessary or recognizing AI as a legal person. Even that approach remains problematic for the DOJ under the traditional methods of finding mens rea, unless the DOJ can somehow prove that the AI acted with criminal intent, or unless Congress

dispenses with the need to prove criminal intent for AI crimes.

Under current law, the DOJ may have options to prosecute AI crimes by focusing on compliance-based crimes or strict liability offenses, where those offenses are applicable. In this uncertain environment, companies seeking to implement AI solutions to replace human decision-making should—at the outset—ensure that they have done sufficient diligence and testing of their AI systems, and that they have effective controls to ensure that their AI systems are working as intended.

As a starting point, just as any compliance program should be based on a thorough assessment of relevant risks, companies should consider the risks presented by the AI systems in use within their organization. Companies should evaluate, for all business and compliance-related functions that rely upon AI, the probability that a particular risk (e.g., the AI system will allow an illegal transaction to proceed) will materialize and the resulting impact. Once the risks have been identified and assessed, companies can consider what controls are already in effect that mitigate risk and evaluate the residual risk in light of their risk tolerance. This exercise will not only foster a better understanding of the risk presented by a company's AI-reliant systems, but also enable the company to design additional controls as needed. On the back end, once a company has implemented an AI system, the company will need to institute appropriate mechanisms for testing and monitoring to verify that the system is working as intended. This may include running a parallel program to verify the accuracy of and any unintended outputs (e.g., bias) from an AI-driven process.

Given the DOJ's limited tools to prosecute AI crimes where no one intended for the AI to violate the law, effective compliance likely will be the best defense for companies to avoid criminal charges for AI-based crimes.

Leo Tsao *is a partner in the investigations and white collar defense and Fintech and payments practices at Paul Hastings based in the firm's Washington, D.C. office. His practice focuses on internal corporate investigations, defense of criminal and regulatory enforcement actions, and compliance counseling.*

Robert Luskin *is a partner in the Investigations and white collar defense practice and based in the firm's Washington, D.C. office. He is one of the best known and most highly-regarded litigators in Washington, D.C., concentrating in complex criminal litigation at both the trial court and appellate level.*

Corinne Lammers *is a partner in the litigation department and the investigations and white collar practice group at the firm, based in the firm's Washington, D.C. office. Her practice focuses on anti-corruption compliance counseling and strategy, including compliance program development and enhancement, risk assessments, and testing and monitoring.*

—Braddock Stevenson, of counsel with the firm and Natasha Nicholson Gaviria, an associate with the firm, contributed to the preparation of the article.