



NOT FOR REPRINT

COMMENTARY

## Courts Remain Hesitant to Regulate the Use of Generative AI in Litigation

While AI technology advances at breathtaking pace, the legal system—not exactly known as an early adopter of new technology—has continued to move slowly and cautiously in response, awaiting new developments rather than trying to anticipate them.

June 18, 2024 at 10:21 PM

Artificial Intelligence

By Natalie Shkolnik and Michael Van Riper | June 18, 2024 at 10:21 PM



Last summer, an otherwise routine personal injury action vividly demonstrated the pitfalls of generative AI. Facing an unfamiliar issue on a motion to dismiss, two New York attorneys turned to ChatGPT to assist their research, unaware of its propensity for the occasional “hallucination.” The chatbot then conjured up non-existent cases to support their position, which the attorneys relied on in their briefing.

In an order sanctioning the attorneys, Judge Kevin Castel of the federal Southern District of New York dryly noted that the chatbot’s fictitious work product “shows stylistic and reasoning flaws that do not generally appear in decisions issued by United States Courts of Appeals,” and less charitably described its legal analysis as “gibberish.” *Mata v. Avianca*, 678 F. Supp. 3d 443, 453 (S.D.N.Y. 2023). Yet these citations were facially passable enough to fool two seasoned litigators once they were outside their usual area of expertise. That is a problem.

Following the headline-grabbing debacle in *Mata*, then, one might have expected the courts to start reworking the rules to limit the risk that the increasing use of AI in litigation would introduce similar errors. Legal research and briefing, after all, is not the only part of litigation that AI will both benefit and disrupt. E-discovery platforms already make excellent use of non-generative AI to collect, sort, and search documents, and will increasingly incorporate generative AI to summarize and analyze their content in written work product as the technology improves. AI-enhanced imagery and digital analysis promises to enhance the quality of poorly recorded evidence.

Yet these tasks are also vulnerable to the “black box” nature of generative AI; that is, the inability of AI’s human interlocutors to fully understand and reproduce the methods through which it reaches its results. Indeed, AI data analysis and evidentiary enhancement is arguably more vulnerable to undetected error than AI-assisted briefing and argument. Unlike with briefing and argument, where any AI-generated output can be easily checked against the federal reporter system and the text of the cases themselves, there is no settled and universally available reference point against which an adversary, court, or even the user itself can check an AI’s interpretation of controverted data or imagery in a specific case. And the amount of time and effort required to manually check the output against a full case record could largely defeat the time-saving purpose of AI.

Instead, the courts have moved slowly. Only a relative handful of courts and judges have adopted rules governing the use of AI. Most of these focus on the use of generative AI in research and writing, requiring attorneys to disclose whether they used AI in preparing court filings. Few apply to the use of AI processes more generally, and almost none outright ban

the practice.

The courts are also far from consensus. A proposed rule requiring attorneys to disclose the use of generative AI in the Fifth Circuit Court of Appeals has sharply divided the legal community, with many public commenters doubting the need for newly crafted rules given the bar's existing ethical obligations, and others suggesting that the rule did not go far enough.

In a reportedly first-of-its-kind decision issued in late March, meanwhile, a trial court judge in Washington state excluded seven AI-enhanced iPhone recordings from evidence. Applying the venerable standard of expert reliability set forth in *Frye v. United States*, the court reasoned that there is no scientific consensus yet on the reliability of AI-enhancement techniques. More pointedly, the court observed that the machine learning algorithms used to produce such images are “not reproducible” by forensic experts—an apparent allusion to the “black box” problem. While it is hard to fault the court's reasoning given the current state of the art, applying *Frye* only postpones an eventual reckoning with that problem.

Earlier this month, the New York State Bar Association's Task Force on Artificial Intelligence also weighed in with its long-awaited report on the legal, social, and ethical implications of AI. The Task Force identified many potential benefits and perils from the use of AI in litigation and provided ethical AI guidelines that practicing lawyers would do well to read in full. But it stopped short of proposing extensive new court rules and procedures, instead reaching the stark conclusion that “this report offers no ‘conclusions.’”

Finally, on April 19, the Advisory Committee on Evidence Rules—one of the five principal rule-making committees of the federal courts—met to consider one of the few concrete rule revisions currently on the table for regulating generative AI across the court system. That proposal would add a new provision to Federal Rule of Evidence 901(b) requiring the proponent of AI-generated evidence to authenticate such evidence by describing the software used to generate the evidence and showing that it produced “valid and reliable” results. The proposal would also add a new Rule 901(c) to address the problem of so-called “deepfakes.” Under the new rule, a party may explicitly challenge electronic evidence by showing that it is “more likely than not” fabricated or altered, after which the proponent of the evidence would need to show that its probative value outweighs any likely prejudice in order to admit the item into evidence.

The eight-member panel of the Advisory Committee, which includes both well-respected practicing attorneys and experienced trial and appellate judges, expressed considerable skepticism about those revisions. The panel especially doubted whether the proposals were yet necessary, given the court system's existing toolkit for managing evidence and regulating the conduct of attorneys. Judge Richard Sullivan of the Second Circuit Court of Appeals summed up that view: “I'm not sure that this is the crisis that it's been painted as, and I'm not sure that judges don't have the tools already to deal with this.”

So while AI technology advances at breathtaking pace, the legal system—not exactly known as an early adopter of new technology—has continued to move slowly and cautiously in response, awaiting new developments rather than trying to anticipate them. But perhaps this posture is for the best. After all, as Justice Oliver Wendell Holmes Jr. famously put it: “The life of the law has not been logic; it has been experience.”

**Natalie Shkolnik** represents domestic and international companies in complex litigation at New York-based *Wilk Auslander*. She excels at helping foreign entities understand the intricacies of U.S. litigation rules, including interpreting the differences in discovery and attorney-client privilege. Her extensive background in complex commercial litigation and securities litigation encompasses class-action cases, governmental and regulatory investigations, and arbitrations. Fluent in both Russian and English, Natalie serves a diverse international client base.

**Michael Van Riper** practices in Wilk Auslander's commercial litigation group, focusing on cross-border judgment enforcement, securities, and complex commercial litigation, and regulatory matters before FINRA and the SEC. He principally represents clients in the securities and investment banking industries in state and federal courts across New York and New Jersey and in alternative dispute resolution.

NOT FOR REPRINT

---

Copyright © 2024 ALM Global, LLC. All Rights Reserved.