Federal Judicial Conference to Revise Rules of Evidence to Address AI Risks

March 20, 2025

As the first quarter of 2025 draws to a close and we look ahead to the spring, important changes to the Federal Rules of Evidence ("FRE") regarding the use of AI in the courtroom are on the horizon. Specifically, the Federal Judicial Conference's Advisory Committee on Evidence Rules (the "Committee") is expected to vote on at least one AI-specific proposal at its next meeting on May 2, 2025. The Committee has been grappling with how to handle evidence that is a product of machine learning, which would be subject to Rule 702 if propounded by a human expert.

At the Committee's last meeting in November 2024, it agreed to develop a formal proposal for a new rule—which, if adopted, would become Rule 707 of the FRE—that would require federal courts to apply Rule 702's standards to machine-generated evidence. This means that the proponent of such evidence would, among other things, need to demonstrate that the evidence is the product of reliable principles and methods, and that those principles and methods were reliably applied to the facts of the case.

The Committee is also expected to continue its discussion of a second issue: how to safeguard against AI-generated deepfake audio or video evidence. For now, the Committee is likely to continue to take a wait-and-see approach because existing rules may be sufficiently flexible to deal with this issue. That being said, the Committee is likely to assess language for a possible amendment, so as to be able to respond if problems do arise.

Reliability of AI-Generated Evidence

Proposed new Rule 707 aims to address the reliability of AI-generated evidence that is akin to expert testimony—and therefore comes with similar concerns about reliability, analytical error or incompleteness, inaccuracy, bias, and/or lack of interpretability. *See Advisory Committee on Evidence Rules Agenda Book (Nov. 8, 2024)*, Tab 4 – *Memorandum Re: Artificial Intelligence, Machine-Learning, and Possible Amendments to the Federal Rules of Evidence* (Oct. 1, 2024), at 51-52 ("Reporter's Proposal"); *see also Committee on Rules of Practice and Procedure Agenda Book (Jan. 7, 2025)*, Tab 3A – Report of the Advisory

Committee on Evidence Rules (Dec. 1, 2024), at 3 ("Committee Dec. 24 Report"). Those concerns are heightened with respect to AI-generated content because it may be the result of complex processes that are difficult (if not impossible) to audit and certify. Examples of AI-generated evidence could include:

- In a securities litigation, an AI system analyzes stock trading patterns over the last ten years to demonstrate the relative magnitude of the stock drop as a percentage of the Dow Jones Industrial Average, or to assess how likely it is that the drop in price was caused by a particular event.

- An AI system analyzes keycard access records, iPhone GPS tracking, and Outlook calendar entries to demonstrate that an individual did not attend any of the senior management meetings over a period of time where alleged wrongdoing occurred.

- In a copyright dispute, an AI system analyzes image data to determine whether two works are substantially similar.

- An AI system assesses the complexity of an allegedly stolen software program in a trade secret dispute and renders an assessment of how long it would take to independently develop the code based on its complexity (and without the benefit of the allegedly misappropriated code).

Under the current rules, the methodologies that human expert witnesses employ and rely on are subject to Rule 702, which requires them to, among other things, establish that their testimony is based on sufficient facts or data; is the product of reliable principles and methods; and that those principles and methods are reliably applied to the facts of the case. See FRE Rule 702 (a)-(d). However, if machine or software output is presented on its own, without the accompaniment of a human expert, Rule 702 isn't obviously applicable, see Reporter's Proposal at 51. This leaves courts and litigants to craft case-by-case frameworks for deciding when and whether AI-driven software systems can be allowed to make predictions or inferences that can be converted into trial testimony.

As a result, at its May 2, 2025 meeting, the Committee is expected to vote on proposed new Rule 707, Machine-Generated Evidence, drafted by the Committee's Reporter, Professor Daniel J. Capra of Fordham School of Law. (If approved, the Rule will be published for public comment.) The text of the proposed Rule provides:

> Where the output of a process or system would be subject to Rule 702 if testified to by a human witness, the court must find that the output satisfies the requirements of Rule 702 (a)-(d). This rule does not apply to the output of basic scientific instruments or routinely relied upon

commercial software. Reporter's Proposal at 51; Committee Dec. 24 Report at 3.

For instance, if a party uses AI to calculate a damages amount without proffering a damages expert, then they would need to prove that adequate data were used as the inputs for the AI program; that the AI program used reliable principles and methods; and that the resulting output is valid and reflects a reliable application of the principles and methods to the inputs, among other things. If adopted, Rule 707 analysis could require a determination of whether the training data is sufficiently representative to render an accurate output; whether the opponent and independent researchers have been provided sufficient access to the program to allow for adversarial scrutiny and sufficient peer review; and whether the process has been validated in sufficiently similar circumstances. *See* Reporter's Proposal at 51-52.

That the Committee is likely to approve this proposal underscores the federal judiciary's concerns about the reliability of certain AI-generated evidence that litigants have already sought to introduce in courtrooms. For example, U.S. District Judge Edgardo Ramos of the U.S. District for the Southern District of New York admonished a law firm for submitting ChatGPT-generated responses as evidence of reasonable attorney hourly rates because "ChatGPT has been shown to be an unreliable resource." *Z.H. v. New York City Dep't of Educ.*, 2024 WL 3385690, at *5 (S.D.N.Y. Jul. 12, 2024). U.S. District Judge Paul Engelmayer similarly rejected AI-generated evidence because the proponent did "not identify the inputs on which ChatGPT relied" or substantiate that ChatGPT considered "very real and relevant" legal precedents. *J.G. v. New York City Dep't of Educ.*, 719 F. Supp. 3d 293, 308 (S.D.N.Y. 2024).

State courts also are beginning to grapple with the reliability of AI-generated evidence. For example:

- In *Washington v. Puloka*, No. 21-1-04851-2 (Super. Ct. King Co. Wash. March 29, 2024), a trial judge excluded an expert's video where AI was used to increase resolution, sharpness, and definition because the expert "did not know what videos the AI-enhancement models are 'trained' on, did not know whether such models employ 'generative AI' in their algorithms, and agreed that such algorithms are opaque and proprietary." *Id.* at Par. 10.

- In *Matter of Weber as Tr. of Michael S. Weber Tr.*, 220 N.Y.S.3d 620 (N.Y. Sur. Ct. 2024), a New York state judge rejected a damages expert's financial calculations in part because he relied on Microsoft Copilot—a large language model generative AI chatbot—to perform calculations but could not describe the sources Copilot relied upon or how the AI tool arrived at its conclusion. In doing so, the judge reran the expert's inquiries on Copilot getting different results each time, and queried Copilot

regarding its reliability, to which Copilot self-reported that it should be "check[ed] with experts for critical issues." *Id.* at 633-35.

- [Reports](#) indicate that a Florida state judge in Broward County recently donned a virtual reality headset provided by the defense to view a virtual scene of the crime from the perspective of the defendant who is charged with aggravated assault. The parties are likely to litigate the reliability of the technology before the judge decides if it can be used by a jury.

In both *Puloka* and *Weber*, the state courts emphasized that their respective jurisdictions follow the *Frye* standard, requiring scientific evidence to be generally accepted in its field, and found no evidence supporting the general acceptance of AI-generated evidence. These initial judicial reactions indicate that experts should be prepared to satisfy the jurisdiction-specific reliability standards for AI technologies they rely on when rendering their expert opinions.

Keeping Deepfakes Out of the Courtroom

A related but distinct concern involves rules for handling AI-generated deepfakes. Although some scholars have warned of a coming "perfect evidentiary storm" due to the difficulty for even computers to detect deepfakes, *see* Reporter's Proposal at 5, the Committee—at least for now—is [unconvinced](#) that the existing Rules need to be immediately amended (or new ones introduced) to deal with this issue. Those expressing skepticism recalled that, when social media and texting first became popular, there were similar concerns about a judicial quagmire arising from parties routinely challenging admission of their texts/social media posts on the grounds that the accounts had been hacked and the texts/posts were not, in fact, their own. But the feared flood of litigation never arrived and FRE's Rule 901 proved up to the task of adjudicating the relatively few challenges that did come up.

In light of that history, the Committee has developed—but does not yet plan to vote on—text that would amend Rule 901 to add a subsection (c) as follows:

> If a party challenging the authenticity of computer-generated or other electronic evidence demonstrates to the court that a jury reasonably could find that the evidence has been fabricated, in whole or in part, by artificial intelligence, the evidence is admissible only if the proponent demonstrates to the court that it is more likely than not authentic. Committee Dec. 24 Report at 4.

This addition would constitute a proactive approach to addressing the potential misuse of AI-generated deepfakes in the courtroom, which would allow an opponent of the evidence to challenge the authenticity of an alleged deepfake and would cover all evidentiary deepfake disputes. But, as some Committee members have pointed out, creating a distinct "right-to-challenge" could itself invite unnecessary sparring among litigants and encourage them to refuse to enter into otherwise routine stipulations. Nor is it clear how far litigants could push any new rule in challenging other types of AI-generated materials as "inauthentic" even if they are not intentionally deceptive including, for example:

- Unofficial transcripts or summaries of meetings produced by AI that are largely, but not entirely, accurate.

- AI-simulated or altered evidence such as a video that recreates a crime scene for a jury to demonstrate how dark it was and how difficult it could have been for a witness to view the crime from a certain distance.

- AI-enhancements to otherwise unaltered videos or photographs to increase their resolution.

- Evidence that was altered by AI for some reason that is not material for the purpose for which it is being offered (e.g., a photo that was altered to remove someone in the background, that later becomes relevant in a litigation).

Because of the potential for increased evidentiary disputes stemming from the proposed amendment, the Committee has also discussed whether to address bad-faith evidentiary challenges by potentially issuing guidance to courts regarding the issuance of sanctions for such bad-faith challenges. This is another area to watch at the upcoming May meeting.

Practical Considerations

Even if new Rule 707 is approved for public comment in May, formal adoption of the Rule is still likely years away. That being said, even now litigators can begin thinking through steps to ensure they can appropriately leverage potentially powerful AI tools in courtroom presentations, including:

- **Conducting Robust Diligence Before Attempting to Admit AI-Generated Evidence**. Litigants who want to rely on AI-generated evidence should consider how to establish that the AI generates reliable, consistently accurate results when applied

to similar facts and circumstances, and that the methodology underlying those results is reproduceable, including by opponents and peer reviewers.

- **Preparing to Disclose AI Systems for Adversarial Scrutiny**. The draft Committee Note to proposed Rule 707 implies an expectation that proponents of AI-generated evidence will provide their opponents and independent researchers with access to the AI technology for adversarial scrutiny—the validation studies conducted by the developer or related entities are unlikely to suffice. Litigants should think carefully now about the legal, commercial, and reputational implications of having to disclose their AI technologies both before significantly investing in them and before seeking to admit AI-generated evidence.

- **Developing Methods to Efficiently Authenticate Audio, Video, or Photographic Evidence**. In light of the federal judiciary's concern with possible use of deepfakes in litigation and the potential for increased evidentiary disputes over AI-generated evidence, litigants should consider developing strategies and capabilities to authenticate evidence that could have, but has not been, altered or fabricated by AI. Examples could include chain of custody record-keeping, use of software to detect image or audio manipulation, as well as retaining qualified forensic experts that can identify AI-generated alterations (or, conversely, testify to their absence).

To subscribe to the Data Blog, please [click here](#).

* * *

Please do not hesitate to contact us with any questions.

Avi Gesser
Partner, New York
+1 212 909 6577
agesser@debevoise.com

Jim Pastore
Partner, New York
+1 212 909 6793
jjpastore@debevoise.com

Matt Kelly
Counsel, New York
+1 212 909 6990
makelly@debevoise.com

Jackie Dorward
Associate, New York
+1 212 909 7406
jmdorward@debevoise.com

Gabriel A. Kohan
Associate, Washington, D.C.
+1 202 383 8036
gakohan@debevoise.com