

## Statistics Tools Chart A Path For AI Use In Expert Testimony

By **Chris Riper and Bilal Shah** (June 12, 2025, 4:16 PM EDT)

The integration of artificial intelligence and machine learning into various professional fields has brought significant advancements and efficiencies.

However, headlines continue to surface about attorneys filing briefs with AI-generated, hallucinated case law, and similar red flags are raised when expert witnesses use AI tools in legal proceedings.[1]

In U.S. federal courts, these concerns are especially relevant under the Rule 702 of the Federal Rules of Evidence, widely known as the Daubert standard.[2] This article discusses recent legal rulings and explores the risks expert witnesses face when relying on AI- and ML-based tools. It then outlines a potential path to court acceptance, much like the one established for statistical tools such as Statistical Analysis System and Stata, which are well accepted under Daubert.

### Rule 702 and Expert Testimony

Rule 702 governs the admissibility of expert testimony in federal courts. According to the rule, an expert witness may testify if several conditions are met, including that "the testimony is the product of reliable principles and methods." [3]

This clause is particularly pertinent when considering the use of AI and ML tools by expert witnesses. The reliability of these tools and how they are applied are key to meeting the Rule 702 standards.

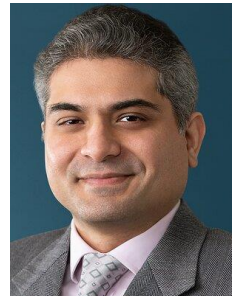
### Case Studies: The Use of AI in Expert Testimony

In *Kohls v. Ellison*, a Minnesota case on free speech limits originally filed in September 2024, the defendant's expert defined "deepfakes" as "highly realistic, AI-generated manipulations of digital content ... where a person's likeness, voice, or actions are convincingly altered or fabricated," and claimed they could "significantly influence political beliefs by presenting convincing false narratives that are difficult to refute." [4]

The expert, a communications professor focused on technology and misinformation, cited two academic articles that could not be verified.[5] The expert admitted to the use of AI to draft his declaration, and the defendant requested leave to amend.



Chris Riper



Bilal Shah

While the U.S. District Court for the District of Minnesota acknowledged in January that the defendant's expert was "qualified to render an expert opinion on AI and deepfakes," it found that his citation to fake, AI-generated sources "shatter[ed] his credibility with the court" and excluded his testimony.[6]

In another 2024 case titled *Matter of Weber*, an expert witness used Microsoft's Copilot AI chatbot to cross-check damages in a real estate dispute.[7]

Judge Jonathan Schopf in the New York Surrogate's Court expressed significant concerns about the reliability of AI-generated outputs, noting "[t]he court ha[d] no objective understanding as to how Copilot works," and emphasized the inherent reliability issues associated with AI.[8] The judge's skepticism was further fueled by the expert's inability to recall the exact prompts used or explain how Copilot arrived at its results.

This lack of transparency and technical understanding undermined the credibility of the expert's testimony, leading the judge to question the admissibility of AI-assisted expert opinions in court.[9]

In yet another example, U.S. District Judge Richard Andrews of the U.S. District Court for the District of Delaware issued an order in March excluding expert testimony in a patent dispute, *Jackson v. NuVasive Inc.* The expert's damages opinion was deemed unreliable under Rule 702 because it depended heavily on third-party data analysis tools, and the expert failed to have a clear understanding of their methodologies.

The expert relied on two tools to conduct a forward citation analysis and determine the value of the patents in question. However, the expert's limited understanding of how these tools calculated their metrics undermined the reliability of his testimony.

The court concluded that the expert's reliance on these AI-based tools failed to meet Rule 702 standards, as the principles and methods were not sufficiently reliable or transparent.[10]

### **Acceptance of Statistical Software in U.S. Courts**

Statistics have long underpinned expert opinion, meeting Daubert's "scientific knowledge" standard by delivering useful, reliable results. While statistical methods are well documented, expert judgment is still needed to select the right method and data and to interpret results.[11]

Software like SAS[12] and Stata[13] revolutionized data analysis in litigation, shifting U.S. courts from early skepticism to broad acceptance. Today, outputs from these tools are routinely presented by expert witnesses in cases from employment discrimination to antitrust, and are rarely challenged. Tools employing methods like multiple regression, analysis of variance, autocorrelation and frequency procedures show how statistical tools have earned courtroom acceptance.

With known error rates, transparent programming, peer-reviewed methods and company warranties, these tools are now trusted, with disputes focusing on evidentiary weight rather than admissibility.[14] Their acceptance offers a compelling road map for how AI and ML might gain similar traction in litigation.

Early decisions, like the U.S. Supreme Court's 1977 rulings in *Teamsters v. U.S.*[15] and *Castaneda v.*

Partida,[16] showed that statistical analysis could credibly demonstrate discrimination and systemic issues.

The Supreme Court's 1986 ruling in *Bazemore v. Friday* confirmed that statistical evidence need not be perfect to be admissible, so long as the analysis is scientifically valid and addresses key variables. Imperfections affect weight, not admissibility.[17]

And its 1993 ruling in *Daubert v. Merrell Dow Pharmaceuticals Inc.*[18] transformed standards for expert evidence, emphasizing relevance and reliability, including testability, peer review and known error rates. Rather than limiting statistical tools, this formalized their role.

Courts began evaluating experts based on their methodology, not just the tool. In *McReynolds v. Sodexho Marriott Services Inc.* in 2004, the U.S. District Court for the District of Columbia accepted a Stata-generated regression analysis even though the expert did not write the code himself. His understanding of the method was key.[19]

Statistical software came to be seen as reliable in qualified hands. Today, Stata and SAS are treated as reliable scientific instruments. Their broad use depends on transparency: Experts share their code and explain the inner workings of their models, allowing others to replicate and assess their appropriateness.

This journey illustrates a path forward for AI and ML tools in litigation.[20] Like statistical tools, they must be used transparently by qualified experts applying peer-reviewed methodologies. Courts will evaluate algorithms as they do regression analyses — based on sound logic and scientific rigor.[21]

By meeting these standards, AI and ML tools can follow SAS and Stata into mainstream judicial acceptance.

## **Implications and Recommendations**

The cases above illustrate the significant risks associated with expert witnesses relying on AI and ML-based tools. These risks include:

- Lack of transparency — AI and ML algorithms often operate as black boxes, making it difficult for users to understand how outputs are generated. This opacity can undermine the credibility of the expert testimony.
- Reliability issues — AI tools can produce errors or hallucinations, generating false or misleading information. Without rigorous verification, these errors can compromise the integrity of legal proceedings.
- Insufficient understanding — Experts may lack the technical knowledge to fully comprehend and explain the AI tools they use. This gap can lead to unreliable testimony that fails to meet the standards of Rule 702(c).

To help mitigate these risks, the following suggestions could assist experts who are considering ways to incorporate AI and ML-based tools and methods into their processes:

- Expert witnesses and their teams should be thoroughly trained on the AI and ML tools they use, and able to confidently articulate the underlying principles and methodologies.
- All AI-generated data and outputs should be thoroughly verified against reliable sources before being presented in court.
- Experts using data derived from AI and ML methods or tools in their analysis should be transparent with their clients, and ensure they understand the benefits and risks of using AI- and ML-derived data.
- Experts should disclose their use of AI tools and provide detailed explanations of how these tools were applied to their analyses.
- Courts should rigorously scrutinize AI-assisted expert testimony, ensuring it meets the reliability standards set by Rule 702(c).
- Expert witnesses should demand that AI and ML vendors disclose the algorithms and formulas behind their outputs and provide accuracy guarantees, especially when the output is deterministic or analytical.

In conclusion, while AI- and ML-based tools offer significant potential benefits, their use by expert witnesses and the witnesses' teams in legal proceedings must be approached with caution. Unlike established tools such as SAS and Stata, which are based on widely accepted methods, AI and ML still have ground to cover before reaching the same level of courtroom trust.

Experts must fully understand these algorithms and be able to replicate the results using other methods or data sources. In other words, these tools can be used as a first pass for efficiency, but their output should not be treated as final without rigorous due diligence.

Ensuring the reliability and transparency of these tools is essential for maintaining the integrity of expert testimony and upholding the standards of the Federal Rules of Evidence.

---

*Chris Riper is a managing director and Bilal Shah is a director at Secretariat Advisors LLC.*

*The opinions expressed are those of the author(s) and do not necessarily reflect the views of their employer, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.*

[1] <https://www.reuters.com/technology/artificial-intelligence/ai-hallucinations-court-papers-spell-trouble-lawyers-2025-02-18/>; <https://www.reuters.com/legal/government/texas-lawyer-fined-ai-use-latest-sanction-over-fake-citations-2024-11-26/>.

[2] The authors acknowledge there are a wide range of rules governing the admissibility of expert testimony across many forums and international jurisdictions and look forward to the continued discussion as this issue evolves.

[3] [https://www.law.cornell.edu/rules/fre/rule\\_702](https://www.law.cornell.edu/rules/fre/rule_702).

[4] <https://storage.courtlistener.com/recap/gov.uscourts.mnd.220348/gov.uscourts.mnd.220348.23.0.pdf>.

[5] <https://stanforddaily.com/2024/12/02/jeff-hancock-court-declaration/>.

[6] <https://storage.courtlistener.com/recap/gov.uscourts.mnd.220348/gov.uscourts.mnd.220348.46.0.pdf>.

[7] <https://arstechnica.com/tech-policy/2024/10/judge-addresses-expert-witness-who-used-ai/>.

[8] [https://www.nycourts.gov/reporter/3dseries/2024/2024\\_24258.htm](https://www.nycourts.gov/reporter/3dseries/2024/2024_24258.htm).

[9] [https://www.nycourts.gov/reporter/3dseries/2024/2024\\_24258.htm](https://www.nycourts.gov/reporter/3dseries/2024/2024_24258.htm) ("Whether or not [the expert] was retained and/ or qualified as a damages expert in areas other than fiduciary duties, [the expert's] testimony shows that [the expert] admittedly did not perform a full analysis of the problem, utilized an incorrect time period for damages, and failed to consider obvious elements into [the expert's] calculations, all of which go against the weight and credibility of [the expert's] opinion.") and ("The use of artificial intelligence is a rapidly growing reality across many industries. The mere fact that artificial intelligence has played a role, which continues to expand in our everyday lives, does not make the results generated by artificial intelligence admissible in Court.").

[10] [https://www.ded.uscourts.gov/sites/ded/files/opinions/21-053\\_2.pdf](https://www.ded.uscourts.gov/sites/ded/files/opinions/21-053_2.pdf), pp. 15-18 ("[Plaintiff] moves to exclude the damages opinion of [], NuVasive's damages expert, on the grounds that it is unreliable under FRE 702. I agree with [Plaintiff]." (citations omitted)); ("[The damages expert's] use of [two data analytics tools] are both unreliable... It is impossible to test [one tool's] conclusion as to the value of the asserted patents if [the damages expert] himself has no understanding of how [the tool] works.").

[11] Reference Manual on Scientific Evidence, Reference Guide on Statistics, Third Edition, pp. 214 - 215. <https://www.fjc.gov/sites/default/files/2015/SciMan3D01.pdf>.

[12] Introduced in the 1970s - [https://www.sas.com/en\\_us/company-information/history.html](https://www.sas.com/en_us/company-information/history.html).

[13] Introduced in mid-1980s - <https://www.stata.com/support/faqs/resources/history-of-stata/>.

[14] Reference Manual on Scientific Evidence, Reference Guide on Statistics, Third Edition, p. 214. <https://www.fjc.gov/sites/default/files/2015/SciMan3D01.pdf>.

[15] *Teamsters v. United States*, 431 U.S. 324 (1977) (<https://supreme.justia.com/cases/federal/us/431/324/>)

[16] *Castaneda v. Partida*, 430 U.S. 482 (1977) (<https://supreme.justia.com/cases/federal/us/430/482/>)

[17] *Bazemore v. Friday*, 478 U.S. 385 (1986) (<https://supreme.justia.com/cases/federal/us/478/385/>)

[18] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993) (<https://supreme.justia.com/cases/federal/us/509/579/>) (An expert may testify about scientific knowledge that assists the jury in understanding the evidence or determining a fact in issue in the case. Factors that a judge should consider include whether the theory or technique in question can be and has been tested, whether it has been subjected to peer review and publication, its known or potential error

rate, the existence and maintenance of standards controlling its operation, and whether it is widely accepted in the relevant scientific community.).

[19] *McReynolds v. Sodexho Marriott Services, Inc.* 1:01-cv-00510 | U.S. District Court for the District of District of Columbia. Opinion and Order (denying defendant's motion to exclude the testimony of plaintiffs' statistical expert) (Dec. 20, 2004).

[20] Other tools used in litigation have undergone a similar journey. For example, the use of "technology assisted review" for discovery of electronically stored information is not only widely accepted but encouraged by the courts because of the benefits it conveys. See *The Sedona Conference, TAR Case Law Primer*, Second Edition, 24 SEDONA CONF. J. 1 (2023).

([https://thesedonaconference.org/publication/TAR\\_Case\\_Law\\_Primer](https://thesedonaconference.org/publication/TAR_Case_Law_Primer)).

[21] See, e.g., *Tyson Foods, Inc. v. Bouaphakeo*, 577 U.S. 442 (2016)

(<https://supreme.justia.com/cases/federal/us/577/442/>) (A representative sample may be sufficient to establish liability in a class action if it is the only feasible way to establish liability, and each member of the class could have relied on the sample to establish liability if each of them had brought an individual action.).