For The Defense™

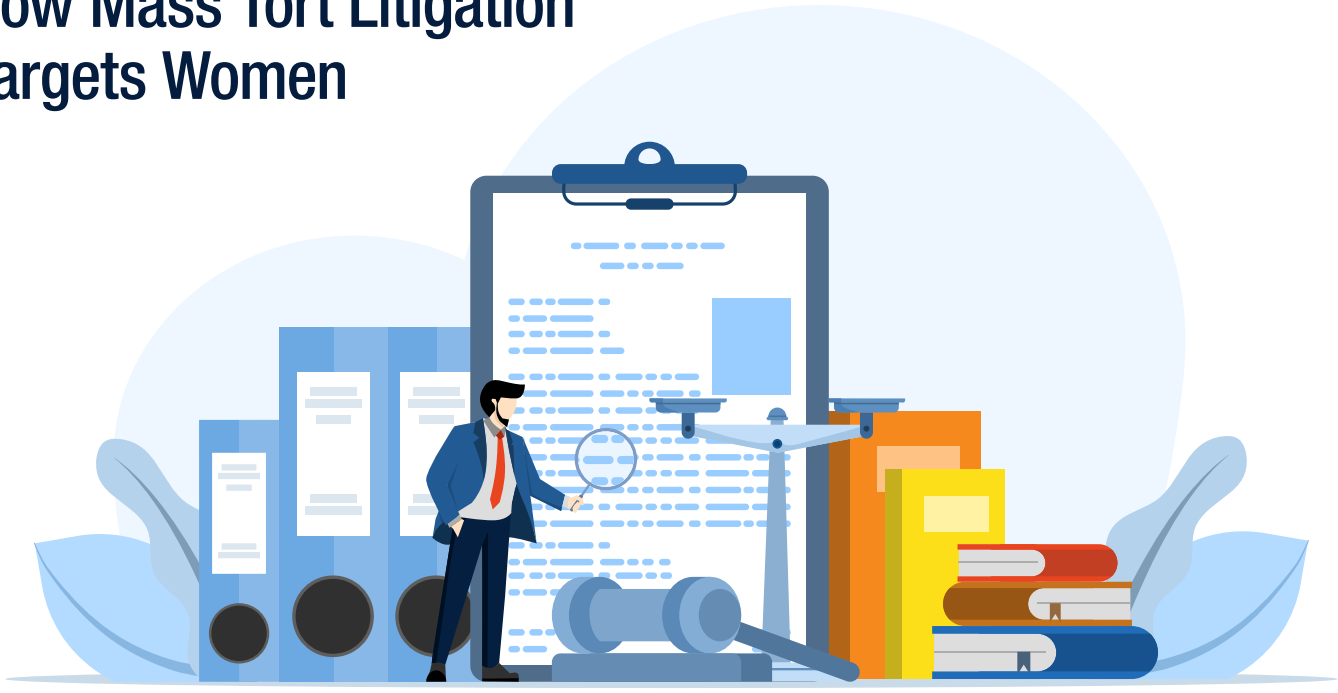
dri
The magazine for defense, insurance and corporate counsel

October 2025

Drug and Medical Device Law and Young Lawyers

Including . . .

Recklessness, Greed and Guinea Pigs: How Mass Tort Litigation Targets Women



Also in This Issue . . .

Different Rules for Defense Counsel: Navigating Discovery of Protected Health Information in Personal Injury Suits and Tips to Avoid Sanctions

And More!

Artificial Intelligence on Trial

Navigating and Challenging Improper Use of AI in Legal Proceedings

By Christian Castile, Jaclyn Setili Wood and Charlotte Flynn

While AI promises both gains in efficiency and reductions in costs, increasing reliance on AI in litigation also raises critical questions of reliability, transparency, and ethics.

The time of Artificial Intelligence (AI) as a distant futuristic eventuality has long ended—even within the legal profession. Once science fiction, AI has now become a practical, evolving tool that is already reshaping the ways we serve the interests of our clients. While AI promises both gains in efficiency and reductions in costs, increasing reliance on AI in litigation also raises critical questions of reliability, transparency, and ethics.

Legally, AI is defined with growing specificity. Under 15 U.S. Code § 9401, AI is defined as "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments." The Department of Defense, in 10 U.S. Code § 2358, offers a complementary and more technical view, identifying AI as "any artificial system that performs tasks... without significant human oversight," "learn[s] from experience," improves with data exposure, and uses "perception, planning, reasoning, learning, communicating, decision making, and acting" to achieve defined goals.

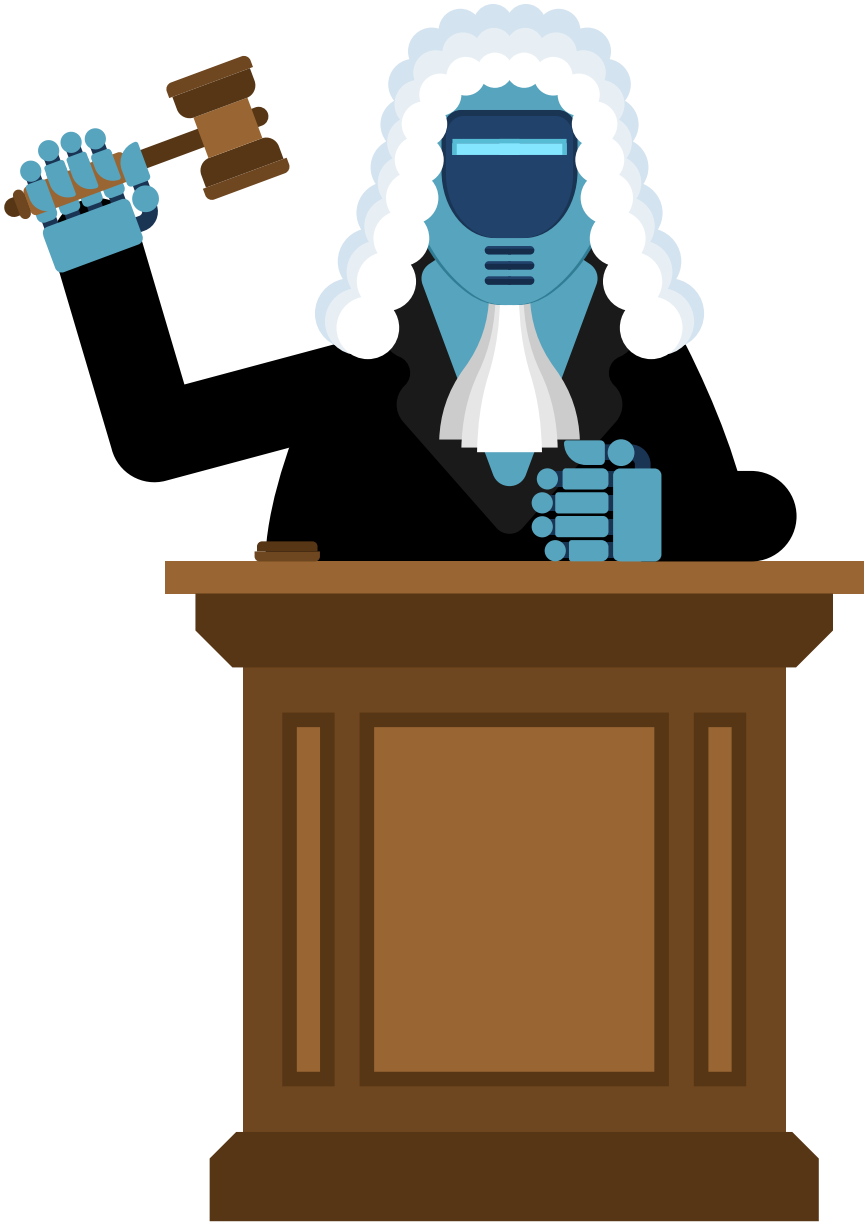
In litigation, AI can take many forms. Tools like Lexis+ AI, Casetext's CoCounsel, Harvey, Everlaw, and even general-purpose platforms like ChatGPT are increasingly being integrated into the day-to-day functions of legal teams. Some attorneys may choose to rely on AI to draft initial versions of pleadings or briefs, summarize deposition transcripts, or even identify inconsistencies across large volumes of evidence. Discovery in particular is already providing fertile grounds for AI innovation and application, and that trend shows no signs of slowing down. Advanced language models are continuing to grow in skill, capable of flagging and isolating potentially privileged content in a collection or synthesizing key information from massive production sets, in turn helping legal teams uncover patterns and insights more efficiently than ever before.

While these developments promise a more streamlined and cost-effective litigation process, they also introduce new challenges. This is particularly true when AI is used without oversight or disclosure, which can open the door to ethical violations and even sanctions. This article focuses on identifying those risks and outlining how defense attorneys can respond when opposing counsel crosses the still-evolving line.

AI Risks and Concerns

The endless potential advantages of incorporating AI as a litigation tool are not without significant risks. For defense counsel, understanding how and when AI can go wrong is essential not only for guarding against personal missteps, but also in developing the skills to identify and

Christian Castile is an associate at Reed Smith whose practice involves a wide range of commercial litigation matters, with an emphasis on defending pharmaceutical and medical device manufacturers in products liability litigation. He has successfully represented companies in a variety of mass torts and MDLs, including litigation involving ranitidine, breast implants, and vaginal mesh. **Jaclyn Setili Wood** is Counsel at Reed Smith and is a member of the Life Sciences Health Industry Group. Jaclyn has experience in numerous types of product liability litigation, but the majority of her practice focuses on complex litigation for pharmaceutical and medical device manufacturers, including federal multi-district litigation and state coordinated proceedings. Jaclyn has experience in all phases of pre-trial litigation, including written and oral discovery, working with fact and expert witnesses, and dispositive and evidentiary motion practice. Jaclyn also has trial experience in state and federal courts. **Charlotte Flynn** is an associate at Reed Smith and concentrates her practice on commercial litigation, with a particular focus on product liability and person injury cases. She works primarily with clients in pharmaceutical, consumer health, and food spaces.



Another pressing concern is the phenomenon of AI "hallucinations," a term used to describe a situation where AI produces factually incorrect or made-up outputs and presents them as credible

respond when opposing counsel crosses ethical or procedural lines.

One key area of concern relates to data and information privacy. Not all AI is created the same, and one of the most critical distinctions between models is whether it is a public (or "open-source") system or a private (or "enterprise-secure") platform. Open-source AI, like ChatGPT, is broadly accessible in part because it is trained on large, publicly available datasets. These systems are constantly evolving, typically by incorporating data and input from user interactions to improve future performance. One can immediately see the significant legal risks this creates. Content entered into a public AI platform may be retained and used for training purposes or even inadvertently exposed to third parties, raising red flags for confidentiality and privilege. By contrast, private AI systems are designed with data security and access control in mind. These platforms are typically built for a specific organization or user base, such as an in-house legal team or law firm, and operate within closed environments that safeguard proprietary or sensitive information.

Another pressing concern is the phenomenon of AI "hallucinations," a term used to describe a situation where AI produces factually incorrect or made-up outputs and presents them as credible. In the legal context, depending on the question posed to the AI model, hallucinations can take the form of fake case citations, made up facts from otherwise real cases, nonexistent statutes, or incorrectly expressed legal principles, just to name a few. And the dangers of hallucinated content are not theoretical. In recent years, multiple courts have imposed sanctions against attorneys who submitted AI-generated briefs that turned out to be riddled with fabricated citations.

A recent and instructive example of the perils associated with AI-generated hallucinations can be found in *Shahid v. Esaam*, No. A25A0196, 2025 Ga. App. LEXIS 299 (Ct. App. June 30, 2025), where the Georgia Court of Appeals had to vacate a trial court's order on a divorce decree after it was revealed that both the trial court's order and the appellee's brief relied extensively on non-existent case law. The underlying appellate dispute centered on whether service by publication was proper, but the appellate court's review was fundamentally compromised by the presence of fake authorities in the record. The court not only vacated the judgment and remanded for a new hearing, but also imposed the maximum permissible penalty on counsel for filing frivolous motions, explicitly referencing the dangers of AI hallucinations and the ethical responsibilities of attorneys to verify the accuracy of their filings. *Shahid* thus starkly illustrates the systemic risks

posed by uncritical reliance on AI-generated legal research.

Judicial Response to AI

As we continue to explore the implications of AI in litigation, the judiciary has become a cautious, but proactive voice in the developing dialogue. While there is no uniform national rule governing the use of AI in legal practice, several federal courts have issued standing orders, proposed local rules, or otherwise expressed concern over the reliability and accountability of AI-generated content—a number expected to continually increase. Even in districts without court-wide guidance, many federal judges have adopted their own standing orders with regard to AI-generated content. Unsurprisingly, though, jurisdictions across the country have taken different approaches.

While FDA may not technically regulate advertising for unrestricted devices, it has sought to expand its jurisdiction to do just that.

One example of the judicial response to generative AI in litigation comes from Judge Michael Baylson of the Eastern District of Pennsylvania. His standing order mandates that any party using AI in the preparation of court filings must include "a clear and plain factual statement" disclosing that use and must certify that all legal and factual citations have been independently verified. A similar approach has been adopted by Judge Padin in the District of New Jersey: her standing policy requires attorneys to identify both the AI tool used and the specific portions of the filing that were AI-generated. Counsel must also certify that a human reviewed the AI tool's output for accuracy and relevance. These types of orders reflect the foundational nature of an attorney's duty of diligence under Rule 11 of the Federal Rules of Civil Procedure.

In the Eastern District of Texas, by contrast, Local Rule AT-3(m) does not mandate the disclosure of AI use in briefing, but does warn attorneys of the "factually or legally inaccurate content" often produced by AI and explicitly reaffirms that filings created with the assistance of AI remain subject to the obligations of Rule 11. The Eastern District of Michigan has gone a step further, proposing Local Rule 5.1(a)(4), which would require affirmative disclosure of any use of generative AI in drafting a court filing. The proposed rule defines "generative AI" broadly and mandates that attorneys attest that they personally verified all legal citations and have ensured the accuracy of the submitted content.

With this sort of piecemeal approach, of course, there is bound to be disagreement and divergence, which inevitably leads to outliers. The Western District of North Carolina, at one end of the spectrum, has adopted one of the most restrictive approaches we have seen. The Western District's standing policy effectively bans the use of generative AI in legal filings, requiring a dual certification from filing attorneys that: (1) no generative AI was used in researching or drafting the document, and (2) a human has reviewed and verified "every statement and every citation" for accuracy. See W.D. N.C., Dkt. No. 3:24-mc-104, June 18, 2024 (Standing Order "In Re: Use of Artificial Intelligence"). At the other end of the spectrum, the Illinois Supreme Court has declined to impose AI-specific disclosure requirements in state court litigation altogether. In doing so, the Court emphasized that existing ethical and procedural safeguards like Rule 11 already provide adequate oversight mechanisms. Illinois Supreme Court Policy on Artificial Intelligence, eff. Jan. 1, 2025.

This jurisdictional divergence serves to underscore the need for counsel to familiarize themselves with local standing orders and proposed rules regarding AI use, even if only to help understand how to respond when your opponents incorporate it into their practice.

Practice Tips for Preempting and Responding to Improper Use of AI

Conferring with Counsel and Striking Problematic Filings

With the increasing use of and reliance on AI tools, litigators are bound to come across a filing that appears to rely on the use of AI. The first step, of course, is to determine whether the applicable jurisdiction/judge applies any rules or standing orders. If it appears that the filing does violate an applicable rule (or if there are no AI-specific rules), an affirmative response may be appropriate.

First, however, courts and local rules often require parties to confer in good faith before seeking judicial intervention regarding deficient or problematic filings. In such courts, failure to timely correct the mistake may result in dire consequences. Indeed, even in jurisdictions where there is no meet and confer requirement, it may be beneficial to send a deficiency letter or notice to the offending counsel, alerting them that their improper or incorrect use of AI did not go unnoticed and giving them an opportunity to withdraw the offending filing. For example, in *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443, 462 (S.D.N.Y. 2023), the court lambasted plaintiffs' counsel when they "abandoned their responsibilities [by submitting] non-existent judicial opinions with fake quotes and citations created by the artificial intelligence tool ChatGPT, then continued to stand by the fake opinions after judicial orders called their existence into question." However, if plaintiffs' counsel had "com[e] clean about their actions shortly after they received the defendant's March 15 brief questioning the existence of the cases," the court clarified, then the outcome might have been different – before imposing a penalty of $5,000 jointly and severally imposed on the attorneys and law firm involved in the faulty filing and ordering, among other things, that they "shall send via first-class mail a letter individually addressed to each judge falsely identified as" an author of a fake case.

When a conference does not resolve the issue, or is unnecessary, a party may escalate the issue by moving to strike the offending filing on the grounds of Federal Rule of Civil Procedure 11 and 37, as well as a court's inherent authority to strike

submissions that contain false or fabricated content.

In *Lacey v. State Farm Gen. Ins. Co.*, 2025 U.S. Dist. LEXIS 90370, at *1, 10 (C.D. Cal. May 6, 2025), the Special Master granted defendant's motion to strike the plaintiff's brief in a privilege dispute after determining that the inclusion of "bogus AI-generated research" constituted reckless conduct "with the improper purpose of trying to influence [the] analysis of the disputed privilege issues." Beyond that, the court also denied the underlying discovery relief sought by the plaintiff, illustrating that submitting—or knowing when and how to challenge—AI-generated briefing can have potentially dispositive consequences. Similarly, in *Gonzalez v. Texas Taxpayers & Research Association*, 2025 U.S. Dist. LEXIS 16801 (W.D. Tex. Jan. 29, 2025), the court granted defendant's motion to strike the plaintiff's brief after finding that the brief contained numerous nonexistent case citations generated by an AI tool. The court emphasized that, regardless of whether the errors were the result of AI or administrative mistakes, the submission of a brief "that contained an abundance of technical and substantive errors" warranted striking the filing and imposing monetary sanctions.

Protective Orders

The rise of generative AI in litigation has, unfortunately, but unsurprisingly, outpaced the development of uniform procedural safeguards. At present, no nationwide rule or standardized protocol exists to regulate how AI systems may be used to store, process, or analyze discovery materials. Yet the ethical and practical implications of incorporating AI into the discovery process are substantial, especially when it comes to safeguarding confidential and privileged client information.

The American Bar Association's Model Rules of Professional Conduct provide a useful starting point. Rule 1.6(c) imposes a duty of technological competence, requiring that lawyers "make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client." This obligation applies regardless of the medium in which the information is stored or transmitted, such that it would extend to digital tools like AI. The use of open-source AI in discovery, therefore, implicates this rule directly.

Open-source generative AI platforms, such as ChatGPT, Claude, and Gemini, are typically trained on broad internet datasets and are accessible to the general public. Although some of these tools include usage restrictions or privacy disclaimers, they are not designed to protect the confidentiality of litigation materials. Uploading sensitive discovery documents into such platforms likely would result in the loss of privilege or waiver of confidentiality protections. It may also expose the underlying data to potential retention by the model's training algorithm, raising additional risks around future use and third-party access. To illustrate, imagine the Coca-Cola recipe being exposed to the public by way of discovery documents in a patent violation lawsuit being fed into ChatGPT.

To mitigate these dangers, protective orders may be an effective means of identifying guardrails for AI use in discovery. Targeted language in a protective order can limit the use of open-source generative AI, while AI-savvy counsel can maintain provisions allowing for the use of secure, closed-universe AI tools that comply with professional standards and preserve client confidences.

For example, a protective order may include provisions such as the following:

Prohibition on Open-Source AI: The order may bar any receiving party from uploading or inputting confidential material into an open-source or publicly accessible generative AI system, regardless of whether that information has been redacted or anonymized. A provision of this nature should make clear that restrictions on AI use apply even when the underlying discovery materials have been anonymized or redacted, as even a party's good-faith effort to obscure identifying details does not, in itself, immunize the disclosure.

Permitted Use of Secure AI Tools: The order may carve out exceptions for document review platforms that employ artificial intelligence within a limited, private, and secure data environment, like those used or provided by e-discovery vendors or in-house litigation support software.

By proactively negotiating such provisions during discovery planning conferences, counsel can more closely safeguard client confidentiality.

In the event of an ongoing case in which the governing protective order does not explicitly address the use of AI, but AI subsequently becomes a concern, a party may still argue that the general language of a protective order to prevent the use of generative AI in confidential proceedings. Most standard protective orders include language prohibiting the disclosure or use of protected material for any purpose other than the prosecution or defense of the action and require that such material be accessed only by individuals authorized under the order. These provisions can provide a strong basis for arguing that uploading confidential documents into an open-source AI platform violates the letter and spirit of the protective order. For example, in *Black v. City of San Diego*, 2025 U.S. Dist. LEXIS 84355 (S.D. Cal. May 2, 2025), a court determined that the use of open-source generative AI would violate a protective order that merely ruled that proceedings be "Confidential," as the data would not be kept confidential if fed into an open-source AI.

Practically speaking, if AI-related concerns arise after a protective order has been entered, parties can and should raise the issue early, either through meet-and-confer efforts or by seeking a clarifying order from the court. However, even without formal modification, cases like Black highlight that litigants have a viable argument that even threadbare protective orders bar the provision of confidential materials to open-source AI platforms.

Conclusion

Emerging jurisprudence on AI misuse in litigation underscores the importance of remaining informed and up to speed on technological advances in the legal profession. Courts have demonstrated a willingness to strike deficient filings and impose sanctions where AI-generated content undermines the reliability of submissions, and savvy defense counsel should be prepared to invoke these remedies when necessary to safeguard the adversarial process and advocate for the rights of their clients.