AI Use In Class Actions Comes With Risks And Rewards

By **Simone Jones, Eric Mattson and Anna Shakotko** (April 22, 2025)

Simone Jones

Eric Mattson

Anna Shakotko

Recent advances in artificial intelligence are being hailed as potential game-changers by some observers, who suggest the new technology could offer solutions across a range of fields including healthcare, science, finance, and even art and design.

The announcements have been coming in a steady drumbeat: In January, DeepSeek released its large reasoning model.[1] OpenAI followed suit, revealing its new "agentic model," Deep Research, on Feb. 2.[2] And on March 31, OpenAI announced that it will be incorporating Deep Research into the free version of ChatGPT very soon.[3]

Some of these cutting-edge tools are now making their way into the courtroom, where they could hold promise for helping courts consider new types of evidence previously too cumbersome to analyze.

But AI use in litigation also comes with meaningful risks — especially now, with the advent of large language models, as made clear by a few well-publicized stories of attorneys submitting briefs containing AI-generated citations to nonexistent articles.

This article examines how experts — and the attorneys who retain them — can successfully use AI in class actions, but also details some of those risks. It concludes by discussing the potential future use of AI in litigation against the backdrop of courts' evolving views on the admissibility of AI-generated materials.

What exactly is artificial intelligence?

AI tools are computer systems that can perform tasks that typically require human intelligence, such as learning, problem-solving and decision-making.[4] The concept of AI has been evolving since the 1950s, building on the development of neural networks and other machine learning techniques that have been used for decades.

Recent breakthroughs in programming architectures, computing power and training data have enabled the creation of novel tools such as large language models and multimodal generative AI.[5]

How can AI be used in class actions?

Before the recent breakthroughs in AI, there was a long history of using more basic AI tools in class actions, particularly to collect and analyze evidence in support of experts' opinions. A key advantage of the most recent AI tools is that they allow experts to analyze large volumes of data that are not in numerical form, such as text, images and video.

With AI, a task that previously required hours or days of human review and many thousands of dollars in fees can now be completed in a fraction of the time at a fraction of the cost. AI tools also have the potential to uncover patterns in data that humans may miss,

unlocking new insights.

For instance, in the context of false advertising litigation, AI tools can be used to analyze a company's marketing materials to better understand the messaging conveyed to consumers and the frequency of alleged misrepresentations. AI-powered image analytics tools can classify images from marketing materials and identify the frequency of specific imagery that may be relevant to plaintiffs' allegations.

Similarly, language analytic AI tools could be applied to text on a company's marketing materials, including ad captions and social media posts, or the company's website content, to analyze the presence and frequency of relevant themes.

An expert witness might also leverage AI-powered tools to analyze consumer-generated data, such as online reviews, comments or social media posts. These analyses can provide insight such as what topics consumers are discussing over time, and consumer sentiment when discussing each of these topics.

AI tools may even be able to provide more nuanced assessments, such as consumers' favorite features in a product, or whether consumers are skeptical of certain marketing claims.

Further, AI tools can be applied to investigate consumer heterogeneity and causality claims. For example, AI can be used to analyze consumer data such as social media posts and identify segments of consumers who are more or less responsive to certain marketing claims.

AI techniques can also help identify complex patterns between the at-issue conduct and the alleged injury faster and more accurately than traditional tools.

Because class actions often deal with absent class members whose experiences with — or damages from — the challenged conduct is difficult or impossible to determine, AI that demonstrates differences in different consumers' interpretation of product marketing might help defendants win the denial of class certification.

An expert analysis that shows heterogeneity in consumer reactions to the challenged conduct could raise questions about standing, causation, reliance and damages that, depending on the jurisdiction and the specific claim, could thwart certification.

How can experts optimize the use of AI and mitigate drawbacks?

While AI tools offer unprecedented analytical capabilities, all tools — AI-powered or not — have limitations and potential drawbacks. By understanding the limitations and drawbacks of tools, expert witnesses can choose the most appropriate tool for the analysis at hand, and take steps to prevent or mitigate potential issues before they arise.

It has been widely publicized that imprudent use of AI tools can compromise confidentiality.[6] Some AI tools may use data submitted as part of a query for subsequent training.

Therefore, any confidential information provided as part of a query may be reflected in some form to another user of the tool. To avoid these confidentiality issues, certain AI tools can be downloaded locally and isolated from broader model training.[7]

Using a locally hosted model can also help address another potential issue when selecting

an analytical tool: replicability. This is particularly pertinent for expert work.

AI tools provided by third parties may evolve without notice, as the underlying models are updated with new training data. As a result, the tool may give a different answer to a query over time, which may raise questions about whether the analysis is truly replicable.[8] Locally hosted models provide control over how and when models may be updated.

Just as with most analytical tools, AI tools come with an error rate — the rate at which the AI gets the answer wrong. The challenge with AI is that, unlike in many traditional statistical analyses, the error rate is not known to the researcher.

The error rate can, however, be approximated. One way to do so is by comparing the AI tool's performance on a task against some benchmark. There are a number of publicly available, domain-specific benchmark datasets that can be used for this purpose.

If a public benchmark does not exist, researchers can generate their own, by having humans perform the same task that the AI performed, usually for a small subset of the data. Then, the human responses can be compared with the AI's.

Of course, humans make errors too, but comparing the agreement rate between the AI and humans can provide an understandable benchmark of the error rate.

One highly publicized type of error that can be made by large language models is when they hallucinate. AI tools hallucinate when their responses to queries make no sense or are completely inaccurate.[9] For instance, AI tools have fabricated academic papers or legal citations as part of their output.[10]

These errors can be difficult for the researcher to identify, because they appear plausible and require independent research to disprove. To reduce the likelihood of hallucinations, a researcher can ask the AI tool to provide links to underlying source documents that can be consulted to verify information.

The researcher can also instruct the tool to confine its search for solutions within a defined set of documents — a process known as retrieval-augmented generation.

How have courts responded to the evolving use of AI tools?

Courts have had diverging reactions to the advent of AI tools. In response to widely publicized instances of lawyers submitting briefs with fake citations generated by large language models, courts struck the papers, issued sanctions and reminded lawyers of their obligations under the Federal Rules.[11]

Some courts also have adopted new local rules relating to AI.[12] Courts have been more receptive to well-tested applications of AI tools.

One example is technology assisted review — a tool that uses human coding of seed sets of documents to predict the responsiveness of the remaining universe of documents for review.

Courts have not only accepted this technology, but encouraged its use, finding it cheaper and more accurate than keyword searches, and thus able to serve litigants well in cases that involve heavy-duty document review.[13]

Courts have had fewer opportunities to assess the reliability of the most cutting-edge AI

tools. But there are signs that AI will appear more often and in novel ways.

In Bertuccelli v. Universal City Studios LLC, a 2020 case in the U.S. District Court for the Eastern District of Louisiana involving alleged violations of the Copyright Act, the plaintiffs' expert used facial recognition software trained on images of the New Orleans Pelicans mascot — the King Cake Baby — to assess the likelihood that consumers would confuse the mascot with the killer's mask in the "Happy Death Day" films.[14]

The court denied a motion to exclude the expert's testimony after finding the AI-assisted analysis reliable enough to predict whether a human might view the images as "substantially similar."[15]

In other cases, AI tools used were not challenged, and the courts therefore did not assess their reliability. Miyoko's Kitchen v. Ross, a 2021 case in the U.S. District Court for the Northern District of California involving allegedly deceptive product labels, is one example.

There, experts conducted a study where participants were presented with images of a product and then asked the meaning of its slogan: "revolutionizing dairy with plants."[16] The responses were fed into a machine learning algorithm to classify them into topics.

Using topic modeling, the experts concluded that participants understood the phrase to mean changing the way dairy is made or producing alternatives to dairy.[17] The court found that there were "reasonable inferences" arising from the report, but that the report did not reach the propriety of the experts' modeling, given the defendant's concession the slogan was not misleading.[18]

Experts' use of AI has met judicial criticism when the expert did not reasonably justify its application and provide transparency around its use. For example, in In re: Weber, a 2024 case in the New York Surrogate's Court, an expert used an AI tool to cross-check his damage calculations without explaining how the tool worked or whether the tool was reliable for the particular application.[19]

While the court acknowledged the absence of any bright-line rule in assessing the reliability of AI tools, it highlighted the need for appropriate human oversight when using those tools, and the need for proof that the tool generates reliable output.[20]

What might the future hold for the use of AI tools?

AI tools are new and evolving, but the tests for admissibility have not changed. Lawyers should ensure that experts' use of AI tools complies with Rule 702 of the Federal Rules of Evidence, or related state rules on the admissibility of expert testimony.

This may include ensuring the expert is using AI tools consistently with how they are applied in rigorous research, or sufficiently addressing the error rate.

Exactly how admissibility lines will be drawn in the AI context is still unclear. But as is true with other complex forms of expert evidence, lawyers should be prepared to explain in plain language why the methodology is reliable.

A lawyer's shrug and pointing to the magic of AI will not suffice.[21] At the same time, courts have allowed the use of AI tools where an expert report explains why the data is reliable.[22]

The recent flurry of fake citations in briefs may have left some judges with a bad impression

of what AI has to offer. But the potential for enhancing the truth-seeking function of litigation, as well as reducing litigation costs in at least some cases, remains a tantalizing possibility.

J. Simone Jones is a partner at Sidley Austin LLP.

Eric Mattson is a partner and co-leader of the consumer class actions and ERISA class actions practice at the firm.

Anna Shakotko is a vice president at Cornerstone Research.

Sidley associate Sudeep Dhanoa and Cornerstone Research senior manager Lucia Yanguas contributed to this article.

[1] "DeepSeek-R1 Release," DeepSeek API Docs, Jan. 20, 2025, https://api-docs.deepseek.com/news/news250120.

[2] "Open AI Unveils A.I. Tool That Can Do Research Online," The New York Times, Feb. 2, 2025, https://www.nytimes.com/2025/02/02/technology/openai-deep-research-tool.html.

[3] "OpenAI says Deep Research is coming to ChatGPT free 'very soon'," Bleeping Computer, March 31, 2025, https://www.bleepingcomputer.com/news/artificial-intelligence/openai-says-deep-research-is-coming-to-chatgpt-free-very-soon/.

[4] Stuart J. Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, Fourth Edition (Pearson Education, 2021), p. 19.

[5] See, e.g., Mustafa Suleyman, "Humans and AI Will Understand Each Other Better Than Ever," Wired, Dec. 28, 2022, https://www.wired.com/story/artificial-intelligence-neural-networks/; "The history of AI," IBM, Oct. 21, 2024, available at https://www.ibm.com/think/topics/history-of-artificial-intelligence.

[6] See, e.g., Rich Vibert, "Think AI Agents Are Safe? These 3 Risks Could Put Your Business In Jeopardy," Forbes, Dec. 23, 2024, https://www.forbes.com/councils/forbestechcouncil/2024/12/23/think-ai-agents-are-safe-these-3-risks-could-put-your-business-in-jeopardy/; Jodi Daniels, "How Generative AI Can Affect Your Business' Data Privacy," Forbes, May 1, 2023, https://www.forbes.com/councils/forbesbusinesscouncil/2023/05/01/how-generative-ai-can-affect-your-business-data-privacy/.

[7] See, e.g., "Microsoft's new toolkit makes running AI locally on Windows easier," TechCrunch, Nov. 15, 2023, https://techcrunch.com/2023/11/15/microsofts-new-toolkit-makes-running-generative-ai-locally-on-windows-easier/.

[8] See, e.g., Lauren Leffer, "Yes, AI Models Can Get Worse over Time," Scientific American, Aug. 2, 2023, https://www.scientificamerican.com/article/yes-ai-models-can-get-worse-over-time/.

[9] "What Are AI Hallucinations?," IBM, https://www.ibm.com/think/topics/ai-hallucinations, Sept. 1, 2023.

[10] Douglas Laney, "Artificial Irony: Misinformation Expert's Testimony Has Fake AI Citations," Forbes, Dec. 6, 2024, https://www.forbes.com/sites/douglaslaney/2024/12/06/artificial-irony-misinformation-experts-testimony-has-fake-citations/.

[11] See, e.g., Mata v. Avianca Inc., 678 F. Supp. 3d 443, 466 (S.D.N.Y. 2023) (sanctions); In re: Celsius Network LLC, 655 B.R. 301, 308–09 (S.D.N.Y. 2023) (striking expert report); Kohls v. Ellison, No. 24-cv-3754, 2025 WL 66514, at *3-5 (D. Minn. Jan. 10, 2023) (excluding expert declaration about dangers of AI and misinformation, and admonishing counsel to validate the papers provided by their witnesses before filing).

[12] Several courts have issued standing orders that require certifications by counsel that they have personally ensured the accuracy of what they are submitting or forbid the use of artificial intelligence in drafting entirely. See, e.g., Standing Order of Judge Araceli Martínez-Olguín, Sec. H. 4. (N.D. Cal., Rev. Nov. 22, 2023) (requiring certification by attorney when utilizing AI); Standing Order of Judge Sharon Johnson Coleman, Mem. of Law Reqs. (N.D. Ill.) (banning use of AI tools for brief drafting).

[13] Hyles v. New York City, No. 10 Civ. 3119, 2016 WL 4077114, at *2 (S.D.N.Y. Aug. 1, 2016).

[14] Bertuccelli v. Universal City Studios LLC, No. 19-1304, Dkt. 149-2 at 8-9 (E.D. La. June 9, 2020).

[15] Bertuccelli v. Universal City Studios LLC, No. 19-1304, 2020 WL 6156821, at *2 (E.D. La. Oct. 21, 2020).

[16] Miyoko's Kitchen v. Ross, No. 20-cv-893, Dkt. 60-3 at 7 (N.D. Cal. Mar. 15, 2021).

[17] Id.

[18] Miyoko's Kitchen v. Ross, No. 20-cv-893, 2021 WL 4497867, at *3, *6 (N.D. Cal. Aug. 10, 2021).

[19] In re: Weber, 220 N.Y.S.3d 620, 633–35 (N.Y. Surrogate's Ct. 2024).

[20] Id.

[21] See id. (expert utilized large language model to cross-check his calculations, but could not state the sources the model relied on, how it worked, or how it arrived at a given output).

[22] E.g., People v. Wakefield, 38 N.Y.3d 367, 380–83 (N.Y. 2022) (noting that the source code of the program was not needed to validate whether the methodology was accepted by the relevant scientific community); Bertuccelli v. Universal City Studios LLC, No. 19-1304, Dkt. 149-2 at 5-9 (describing tools used in depth, including as to their reliability compared to humans).