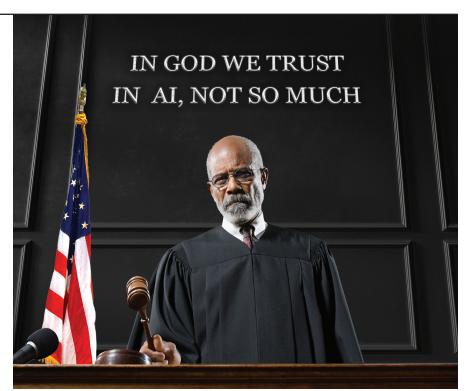Editor-in-Chief
Jeffrey I. Ehrlich
EHRLICH LAW FIRM, APC

AI in the courts

JUDICIAL DECISIONS DEALING WITH AI-RELATED ISSUES

When Open AI first released ChatGPT to the public in November 2022, it heralded a new age of generative AI. Lawyers and legal observers quickly saw that generative AI had the potential to substantially change how lawyers practiced law and how judges decided cases. Now, a bit more than two-and-a-half years later, judicial opinions that refer to or deal with AI-related issues tend to fall into five categories:

(1) judicial orders against litigants and lawyers who included "hallucinated" AI-generated legal citations or evidence in their submissions to the court, and are being sanctioned as a result;

(2) standing orders or rules requiring litigants and counsel to disclose when they have used AI tools to prepare some aspect of a filed document;

(3) judicial orders commenting on a litigant's use of AI-related evidence;

(4) judicial opinions that rely on or discuss the use of AI tools to analyze an issue in the case; and,

(5) cases involving copyright claims brought by copyright holders against AI companies;

In this article, I will be discussing the first four categories, where the issue involves the use of AI by litigants or judges.

The sanction cases – the cost of including hallucinated citations is climbing

I would wager that anyone reading this article is aware of the decision in *Mata v. Avianca, Inc.* (S.D.N.Y. 2023) 678 F.Supp.3d 443, the first case where a judge sanctioned a lawyer for including hallucinated case citations in a document. The case made nationwide news, and the lawyers involved were sanctioned $5,000. At the time, I thought it would be a cautionary tale that would cause lawyers to realize that they had to carefully check any citations provided by a generative-AI model like ChatGPT. I felt bad for the lawyers involved, but thought that they had inadvertently provided a public service.

I was wrong about the impact of the case. The message has not gotten through. Lawyers keep making the same mistake and the price they are paying is starting to climb as judges hope that increasingly severe sanctions will stem the tide of fake citations. (See, e.g., *United States v. Hayes* (E.D. Cal. Jan 17, 2025) ___ F.Supp.3d ___, 2025 WL 235531 at *10-15 [sanctioning criminal defense lawyer $1,500 for using AI; when questioned by the court, the lawyer's

response about the source of inaccurate legal citations "was not accurate and was misleading"]; *Saxena v. Martinez-Hernandez* (D. Nev. April 23, 2025, 2025 WL 1194003 at *2 and n.5 ["Saxena's use of AI generated cases – and his subsequent refusal to accept responsibility for doing so – is just another example of Saxena's abusive litigation tactics, and further explains why the court issued case-terminating sanctions"] (collecting cases); *Park v. Kim* (2d Cir. 2024) 91 F.4th 610, 614 [referring lawyer who used ChatGPT to supply citations in her reply brief, which included a non-existent case, the panel referred the lawyer to the Court's Grievance Panel and ordered the lawyer to serve her client with a copy of the opinion]; *Bunce v. Visual Technology Innovations, Inc.* (E.D. Pa., Feb. 27, 2025, No. CV 23-1740) 2025 WL 662398, at *1 [$2,500 sanction imposed and the attorney was ordered to attend a CLE course on AI]; *Kruse v. Karlen* (Mo. Ct. App. 2024) 692 S.W.3d 43, 51-52 [proper litigant who hired "consultant" attorney to draft reply brief which contained fake citations sanctioned $10,000 for filing a frivolous appeal, and the appeal was dismissed]; *People v. Crabill* (Colo. O.P.D.J., Nov. 22, 2023, No. 23PDJ067) 2023 WL 8111898, at *1 [attorney who used

ChatGPT to draft motion to set aside judgment, which contained fake citations, stipulated to a one-year suspension from the practice of law, with 90 days to be served and the balanced stayed upon his successful completion of a two-year probationary period].)

The most-expensive sanction seems to have been imposed in *Lacey v. State Farm*, Case No. CV 24-5205 FMO, on May 6, 2025. There, a special master appointed by Judge Olguin of the U.S. District Court for the Central District of California, the Hon. Michael A. Wilner (Ret.), imposed a $31,100 sanction on two large firms, Ellis George and K&L Gates, and also struck their briefs and declined to grant any further relief on the discovery issue they sought to litigate.

In *Lacey*, lawyers from one firm used AI tools to generate an "outline" for a supplemental brief. The document contained hallucinated case citations. That lawyer then sent the outline to co-counsel at the other firm, who incorporated the material into the brief. None of the lawyers or their staff cite-checked the document before it was filed. When the Special Master had difficulty locating two of the cited cases, he emailed counsel "to have them address the anomaly." Later that day, the firm re-submitted the brief without the two incorrect citations, but left the remaining false citations in the brief, uncorrected. An associate at the firm emailed the Special Master to thank him for catching the two cases that had been "inadvertently included" in the brief, and confirming that the citations in the revised brief had been "addressed and updated."

Here is the salient portion of the order:

17. . . . I conclude that the lawyers involved in filing the Original and Revised Briefs collectively acted in a manner that was tantamount to bad faith. Fink, 239 F.3d at 994. The initial, undisclosed use of AI products to generate the first draft of the brief was flat-out wrong. *Even with recent advances, no reasonably competent attorney should out-source research and writing to this technology – particularly without any*

attempt to verify the accuracy of that material. And sending that material to other lawyers without disclosing its sketchy AI origins realistically put those professionals in harm's way. Mr. [C.] candidly admitted that this is what happened, and is unreservedly remorseful about it.

18. *Yet, the conduct of the lawyers at K&L Gates is also deeply troubling. They failed to check the validity of the research sent to them. As a result, the fake information found its way into the Original Brief that I read. That's bad. But, when I contacted them and let them know about my concerns regarding a portion of their research, the lawyers' solution was to excise the phony material and submit the Revised Brief – still containing a half-dozen AI errors*. Further, even though the lawyers were on notice of a significant problem with the legal research (as flagged by the brief's recipient: the Special Master), there was no disclosure to me about the use of AI. Instead, the e-mail transmitting the new brief merely suggested an inadvertent production error, not improper reliance on technology. *Translation: they had the information and the chance to fix this problem, but didn't take it.*

19. *I therefore conclude that (a) the initial undisclosed use of AI, (b) the failure to cite-check the Original Brief, and (perhaps most egregiously), (c) the re-submission of the defective Revised Brief without adequate disclosure of the use of AI, taken together, demonstrate reckless conduct with the improper purpose of trying to influence my analysis of the disputed privilege issues*. The Ellis George and K&L Gates firms had adequate opportunities – before and after their error had been brought to their attention – to stop this from happening. Their failure to do so justifies measured sanctions under these circumstances.

20. Those sanctions are as follows. *I have struck, and decline to consider, any of the supplemental briefs that Plaintiff submitted on the privilege issue. From this, I decline to award any of the discovery relief (augmenting a privilege log, ordering production of materials, or requiring in camera review of items) that Plaintiff sought*

in the proceedings that led up to the bogus briefs. I conclude that these non-monetary sanctions will suffice to "deter repetition of the conduct or comparable conduct by others similarly situated." Fed. R. Civ. P. 11(c)(4). If the undisclosed use of AI and the submission of fake law causes a client to lose a motion or case, lawyers will undoubtedly be deterred from going down that pointless route. (Order in *Lacey v. State Farm*, pp. 7-9, emphasis added.)

> **Do not let this happen to you!**
>
> I will repeat a modified version of the admonition I offered in my other article in this issue: "Do not, under any circumstances, rely on any type of historical information provided by any LLM, including legal-specific ones, unless you have personally verified that the cited information exists and says exactly what you are citing it for."

After the order in *Lacey* was filed, perhaps the most ironic of sanctions cases occurred in May 2025, in the ongoing copyright litigation between music publishers and Anthropic, the company that makes Claude. In a declaration filed in federal court in California, Anthropic's own data scientist cited an academic article to support her analysis. The citation was real – the article existed – but when a Latham & Watkins associate used Claude to format the citation, the AI hallucinated the title and authors while keeping the correct link, volume, and page numbers.

The error made it into a filed court document despite what the attorneys called their "manual citation check." When opposing counsel discovered the fabricated names, U.S. Magistrate Judge Susan van Keulen called it "a very serious and grave issue," noting there was "a world of difference between a missed citation and a hallucination generated by AI." After reviewing Latham's explanation that it was "an honest citation mistake," the judge struck the expert's declaration but did not otherwise sanction the firm or the lawyers.

The profound irony – Anthropic's own lawyers, using Anthropic's own AI, in a case about AI and copyright – demonstrates that no one is immune from

hallucination risks. Even partial use of AI (here, just for formatting a citation) can introduce errors. And even the most prestigious firms with sophisticated review processes can miss these fabrications.

And the false information provided by generative AI in lawsuits is not limited to citations to legal authorities. Sometimes, it can be fake to *evidence*. In a patent-infringement action pending in the Eastern District of Pennsylvania, *Magpul Industries Corp. v. Mission First Tactical, LLC.*, Civ.No. 24-5551-KSM, a San Francisco-based patent lawyer who was, at the time, lead counsel for the plaintiff, relied on AI tools to prepare a claim-construction chart that he filed with the court. He included 22 citations to and quotes from written descriptions and 26 to prosecution history to support his proposals. Of those, only two of the written descriptions were real and none of the prosecution history was real. At a hearing on the matter, the lawyer told the court, "I made a horrible mistake. I asked the AI to pull some supporting citations from the intrinsic record that I uploaded. [I] asked it to confirm that quotes were verbatim, what was in the documents, which it did. [I] asked it to double-check its work, which it did. I then submitted the results to a different AI, which did not catch these errors. Rather, it compounded the problem."

The court has not yet issued an order, but in a joint-status report filed by the parties, his client said it had retained new counsel and was deciding whether or not the attorney would remain on the case. At the hearing, the court urged the defense to file a request for sanctions under Rule 11.

Rules and orders requiring the disclosure of the use of AI tools – the list gets longer

Judges and courts across the country have made rules, policies, and standing orders that require the litigants and their counsel to disclose the use of generative AI tools in the preparation of documents filed in that court. The RAILS (Responsible AI in Legal Services) website has compiled a table listing these rules,

policies, and orders. It is available online at https://airtable.com/appKUCriCQDI1BxIV/shrfIAPpNKaNMnacR/tblNmp6mff8CzLuQD.

The table now shows 106 such orders, including nine federal and state judges in California. If you are litigating in federal court, you would be well-advised to check the standing orders of the judge and magistrate handling your case, to see if they have issued such a disclosure rule.

The Hon. Stanley Blumenfeld, Jr., a district judge in the Central District of California, includes the following disclosure requirement in paragraph 5 of his Standing Order for Civil Cases, which governs "Filing Requirements":

> c. **Artificial Intelligence**. Any party who uses generative artificial intelligence (such as ChatGPT, Harvey, CoCounsel, or Google Bard) to generate any portion of a brief, pleading, or other filing must attach to the filing a separate declaration disclosing the use of artificial intelligence and certifying that the filer has reviewed the source material and verified that the artificially generated content is accurate and complies with the filer's Rule 11 obligations.

Cases where litigants have openly relied on generative AI

I have seen very few cases where courts have commented on a litigant's use of generative AI to produce evidence. But in a string of cases brought by the Cuddy Law Firm in New York against the New York City Department of Education, that firm openly relied on ChatGPT to state what the proper billing rate was for lawyers in New York bringing IDEA cases. (See, e.g., *S.C. v. New York City Department of Education* (S.D.N.Y., Apr. 2, 2024, No. 23-CV-1266 (LGS) (JLC)) 2024 WL 1447331, at *5, *report and recommendation adopted* (S.D.N.Y., July 24, 2024, No. 23 CIV. 1266 (LGS)) 2024 WL 3518522.) This was not well received. As the court put it:

> Breaking ground on a novel tactic, CLF also asked artificial intelligence

system ChatGPT-4 about proper rates in IDEA fee cases. Kopp Decl. ¶ 3. ChatGPT, with careful prompting by CLF, managed to produce 14 pages of answers containing no probative information.

(*Id*. at *5.)

The court in *S.C.* cited a prior order from another court that was equally unimpressed with this tactic: "It suffices to say that the Cuddy Law Firm's invocation of ChatGPT as support for its aggressive fee bid is utterly and unusually unpersuasive." (*Ibid*.)

Cases where judges have relied on and cited generative AI in their legal analysis

The landscaper

I have thus far found only two cases where courts have openly relied on generative AI to aid their legal analysis in a case. The first was a concurring opinion by Judge Kevin Newsome of the 11th Circuit, in *Snell v. United Specialty Insurance Company* (11th Cir. 2024) 102 F.4th 1208.

Snell was a declaratory-relief action by Snell, a landscaper, against his liability insurer, seeking to establish that the insurer owed him a defense in an action against him for negligently installing an in-ground trampoline. The district court granted summary judgment for the insurer, on the ground that the accident did not arise from Snell's "landscaping" work within the meaning of his commercial general liability policy. The 11th Circuit affirmed on a different ground; that in his application, Snell answered "no" to the following question: "Do you do any recreational or playground equipment construction or erection?"

But as Judge Newsome explained in his concurring opinion, before the basis for that decision occurred to the court, he had spent hours with dictionaries seeking to ascertain whether the installation of an in-ground trampoline could reasonably be considered to be "landscaping."

The opinion is well-written and worth your time. Here is how the opinion begins:

I write separately (and I'll confess this is a little unusual) simply to pull back the curtain on the process by which I thought through one of the issues in this case – and using my own experience here as backdrop, to make a modest proposal regarding courts' interpretations of the words and phrases used in legal instruments.

Here's the proposal, which I suspect many will reflexively condemn as heresy, but which I promise to unpack if given the chance: Those, like me, who believe that "ordinary meaning" is *the* foundational rule for the evaluation of legal texts should consider – *consider* – whether and how AI-powered large language models like OpenAI's ChatGPT, Google's Gemini, and Anthropic's Claude might – *might* – inform the interpretive analysis. There, having thought the unthinkable, I've said the unsayable.

Now let me explain myself. (*Id.*, 102 F.4th at p. 1221, Newsome, J., concurring.)

Here is why Judge Newsome concluded that generative-AI tools might be worth consulting when a court examines the "ordinary meaning" of a disputed term. He first explains that, before the issue with the application became clear to the court, he attempted to determine, using conventional means, whether a trampoline could be "landscaping." He explains:

I spent hours and hours (and hours) laboring over the question whether Snell's trampoline-installation project qualified as "landscaping" as that term is ordinarily understood. And it was midway along that journey that I had the disconcerting thought that underlies this separate writing: *Is it absurd to think that ChatGPT might be able to shed some light on what the term "landscaping" means?* Initially, I answered my own question in the affirmative: *Yes, Kevin, that is positively absurd.* But the longer and more deeply I considered it, the less absurd it seemed.

* * *

[F]ollowing the district court's lead, I did here what any self-

respecting textualist would do when trying to assess the ordinary meaning of a particular word, here "landscaping": I went to the dictionaries. In his brief, Snell had served up a buffet of definitions. [Citing examples.] As occasionally happens, the dictionaries left a little something to be desired. From their definitions alone, it was tough to discern a single controlling criterion. Must an improvement be natural to count as "landscaping"? Maybe, but that would presumably exclude walkways and accent lights, both of which intuitively seemed (to me, anyway) to qualify. Perhaps "landscaping" work has to be done for aesthetic reasons? That, though, would rule out, for instance, a project to regrade a yard, say away from a house's foundation to prevent basement flooding. I once regraded my own yard, and while my efforts did nothing to improve my house's curb appeal, they served an important functional purpose – and for what it's worth, I definitely thought I was engaged in "landscaping."

(*Id.*, 102 F.4th at p. 1223.)

Judge Newsome then looked at photos of the project in the record, and they did not strike him as "particularly 'landscaping'-y." He then explained:

The problem, of course, was that I couldn't articulate *why*. And visceral, gut-instinct decisionmaking has always given me the willies – I definitely didn't want to be *that* guy. So in a way, I felt like I was back to square one.

And that's when things got weird. Perhaps in a fit of frustration, and most definitely on what can only be described as a lark, I said to one of my clerks, "I wonder what ChatGPT thinks about all this." So he ran a query: "What is the ordinary meaning of 'landscaping'?" Here's what ChatGPT said in response:

"Landscaping" refers to the process of altering the visible features of an area of land, typically a yard, garden or outdoor space, for

aesthetic or practical purposes. This can include activities such as planting trees, shrubs, flowers, or grass, as well as installing paths, fences, water features, and other elements to enhance the appearance and functionality of the outdoor space.

Interesting, for two reasons. First, ChatGPT's explanation seemed more sensible than I had thought it might – and definitely less nutty than I had feared. Second, it squared with my own impression – informed by my own experience writing, reading, speaking, and listening to American English in the real world – that ordinary people might well use the word "landscaping" (1) to include more than just botanical and other natural improvements and (2) to cover both aesthetic and functional objectives. In fact, several of the examples that ChatGPT flagged – "paths, fences, [and] water features" – jibed with the sorts of things that had sprung to mind when I first started thinking about the case.

Suffice it to say, my interest was piqued. But I definitely didn't want to fall into the trap of embracing ChatGPT's definition just because it aligned with my priors. (Bad.) So, in what might have been a mistake – more on that later – we went ahead and asked it the ultimate question: "Is installing an in-ground trampoline 'landscaping'"?

(*Id.*, 102 F.4th at p. 1225.)

ChatGPT answered "yes" because, in its view, "Landscaping involves altering the visible features of an outdoor area for aesthetic or practical purposes, and adding an in-ground trampoline would modify the appearance and function of the space. It's a deliberate change to the outdoor environment, often aimed at enhancing the overall landscape and usability of the area." (*Ibid.*)

Judge Newsome then explained that, although the panel's resolution of the issue did not require him to decide whether the installation of a trampoline could qualify as "landscaping," as used in its ordinary way, the process he undertook to explore that issue with

generative AI brought him to "what was to me a previously unimaginable possibility: Might LLMs be useful in the interpretation of legal texts? Having initially thought the idea positively ludicrous, I think I'm now a pretty firm 'maybe.' At the very least, it seems to me, it's an issue worth exploring." (*Ibid.*)

He then lays out what he sees as the pros and cons of that approach. It's a lengthy discussion that I won't repeat or summarize here, given space constraints, but it's worth a read.

Animal cruelty

The second case I found where courts openly used generative AI in legal analysis were the majority and dissenting opinions of the District of Columbia Court of Appeals (the appellate court for the District of Columbia, not to be confused with the D.C. Circuit), in *Ross v. United States* (DC Ct. App. 2025) __ A.3d __, 2025 WL 561532.

Ross was convicted of animal cruelty for leaving her dog in her parked car on a 98-degree day. The car was parked in the shade of a tree with the windows rolled down several inches. Happily, passersby alerted the authorities to the situation and they opened the car and removed the dog, who seemed unharmed. They arrested Ross when she arrived at the scene about an hour after the police arrived. The Court of Appeal reversed her conviction, finding that the government had not proved all the elements of the offense. In particular, it failed to prove how hot it was in the car, or the presence of any heat-related distress in the dog.

In dissent, Judge Deahl, was offended, as any dog lover would be. The trial court had found that it was "common knowledge" that leaving a dog in a parked car on a 98-degree day was a harmful, potentially deadly situation, even with the windows rolled down a few inches. She explained why, at length. But part of her dissent stated:

Let me nonetheless briefly scrutinize what I have claimed to be common knowledge, using both the relevant facts of this case and those presented in Jordan. I have asked ChatGPT, "Is it

harmful to leave a dog in a car, with the windows down a few inches, for an hour and twenty minutes when it's 98 degrees outside?" Its response, with my emphasis added, follows:

Yes, leaving a dog in a car under these conditions is very harmful. Even with the windows slightly open, the temperature inside a car can quickly rise to dangerous levels. In fact, on a 98-degree day, the temperature inside a car can escalate to over 120°F (49°C) in just a matter of minutes, which can cause heatstroke or even be fatal to a dog.

Dogs can suffer from heatstroke when they are exposed to extreme temperatures, as they do not regulate heat as efficiently as humans. They rely on panting to cool down, but this becomes less effective in a hot, confined space like a car. Symptoms of heatstroke in dogs include excessive panting, drooling, weakness, vomiting, and even collapse.

For your dog's safety and well-being, it's important never to leave them in a hot car, even with the windows cracked. If you need to run errands or be in a hot environment, it's best to leave your dog at home in a cooler, safer space.

Now compare that unequivocal affirmative answer to how it responds to the facts of Jordan. I asked ChatGPT, "Is it harmful to leave a German shepherd outside in 25 degree temperature for five hours? The first paragraph of its far lengthier response . . . boils down to "it depends."
(*Id*. at p. *14.)

In a footnote, the majority noted that ChatGPT was not always reliable. And in a concurrence, another judge, who sits on the D.C. Courts AI Task Force, to address some points that were not addressed in Judge Newsome's concurring opinion in *Snell*, including that AI systems can be susceptible to bias, and could inadvertently disclose confidential information. He noted that the use of AI in responsible ways could lighten courts' workload and make them more efficient, and he assured the public that "cautious and proactive thought is being directed

by our judges and D.C. Courts team members, toward the beneficial, secure, and safe use of AI technology."

Judicial opinion error

On June 30, 2025, U.S. District Judge Julien Xavier Neals of the District of New Jersey withdrew his decision and order in a biopharma securities case after lawyers complained that his decision included made-up quotations and misstated case outcomes. The court stated that "the opinion and order were entered in error." Although there is no acknowledgement that AI was the culprit, these errors bear the hallmarks of AI hallucination.

A closing observation

In January and February 2025, the Thomson Reuters Institute surveyed 1,700 legal professionals, mostly in the U.S., Canada, and the U.K., and complied a report titled "2025 Generative AI in Professional Services Report." Forty-one percent of the respondents said that they used publicly available tools like ChatGPT in their practices, and 17 percent said that they used legal-specific Generative AI tools. But in the same survey, 71 percent of law-firm clients said that they were not aware of whether their lawyers were using Generative AI tools.

In other words, it appears that the lawyers are using the tools but not telling the clients. Do we really think that busy judges and law clerks are behaving any differently? I therefore applaud judges like Judge Newsome and Judge Deahl for their transparency. I hope we see more of it by lawyers and judges alike.

Jeffrey I. Ehrlich is the principal of the Ehrlich Law Firm, APC, *in Claremont, California. He is certified by the State Bar of California as an Appellate Specialist. He is the editor-in-chief of Advocate Magazine, is an Emeritus Member of the CAALA Board of Governors, and is the Chair of the California Academy of Appellate Lawyers' Task Force on Generative AI.*