

To Appear in Vol. 107, Iss. 2 of *Judicature* (Oct. 2023)

## **IS DISCLOSURE AND CERTIFICATION OF THE USE OF GENERATIVE AI REALLY NECESSARY?**

Maura R. Grossman, Paul W. Grimm, and Daniel G. Brown\*

### **Abstract**

Stories about Generative AI (“GenAI”) applications such as ChatGPT have dominated the news for much of 2023. Whether this technology is a blessing or a curse is still open to debate; what is not, however, is that GenAI is here to stay and that billions of dollars are pouring into the development of new applications. While panic spread through the educational community over fear that students would use GenAI to complete their assignments and examinations, there were also some mortifying and highly publicized misuses of GenAI in the legal profession, including attorneys filing pleadings citing fictitious legal authority which was the product of GenAI hallucinations. Concerns about the misuse of GenAI in their own courtrooms prompted several judges in North America to issue individual standing orders requiring disclosure of and certifications related to the use of GenAI specifically, or AI more generally, in connection with legal filings. While an understandable reaction, these orders have lacked consistency in what they cover or require, have been over-broad in scope, and have the potential to cause uncertainty and confusion within the bar, as well to chill legitimate uses of GenAI to increase access to the courts by self-represented litigants, and to reduce the costs and burdens associated with legal research and writing.

This article addresses issues related to the use of GenAI in the justice system and the proactive efforts of individual judges to prevent its misuse in their courtrooms. We focus on the professional lapses that prompted the courts’ reactions, provide examples of the types of orders that judges have issued, explain how GenAI applications operate, why they can hallucinate, and discuss the potential problems, including confusion, increased costs, and the potential chilling effects that accompany such standing orders. We explain why ad-hoc orders may discourage appropriate use of GenAI to make the courts more accessible and the practice of law more efficient. We argue that existing rules of practice and procedure and rules of professional conduct already prohibit this misconduct, and that existing authority contains an equivalent if not stronger deterrent for the misuse of GenAI, without the concomitant downsides of the standing orders. As an alternative, we recommend that courts consider adopting local rules—enacted after public

notice and an opportunity for comment—that would apply court-wide, instead of the rapidly developing mosaic of individual standing orders for individual courtrooms. Local rules could address the problem in a more nuanced way without the unintended consequences. Finally, we explain how courts can address the public at large and *pro-se* litigants in particular, through their websites, to explain the proper and improper use of AI and GenAI applications in court cases.

## I. A Threat is on the Horizon

It is nearly impossible these days to read any news report without seeing multiple articles about the promises or perils of generative AI (“GenAI”) applications, such as ChatGPT and GPT-4. Depending on the author’s perspective, the use of large language models (“LLMs”) will either result in utopian or dystopian outcomes, with computers replacing humans in many activities. But, as is usually the case, the truth falls somewhere in the middle; there are both benefits and risks to these new tools. Regardless of one’s viewpoint, the genie is out of the bottle. GenAI applications are already in widespread use and billions of dollars are being invested in further development of this technology.

The legal profession is not immune from these developments. GenAI is presently being used for research and drafting purposes and is being implemented in eDiscovery tools. Its uses will only continue to proliferate in the future. Some welcome this development, while others dread it. Increasingly, judges are issuing individual standing orders for their courts that may intentionally or unintentionally curtail the use of GenAI in connection with court filings because they require litigants to disclose their use and to submit certifications about their efforts to verify the accuracy of any factual representations or case authority cited when using GenAI.

At first blush, this might seem like a welcome development. Little guidance has been offered on the use of GenAI to generate pleadings, and judges unquestionably have the inherent authority to issue orders and guidelines governing what lawyers and parties must do in cases pending before them. But while the impulse underlying the imposition of such orders is understandable—even commendable—there may be real disadvantages in doing so. For example, some of the orders have been overly broad—sweeping into their scope AI applications that do not produce final work product and that do not suffer from GenAI’s propensity to hallucinate. Such orders may infringe on attorney work product and can discourage the use of technology that will increase access to justice for unrepresented litigants and reduce costs for litigants who are represented. Other orders have been vague and ambiguous about the technologies they cover, leading to confusion among counsel and parties. And, given the speed with which judges are issuing such orders, there has been a lack of consistency, which can only add to confusion and impose additional burdens and costs on litigants who must—on pain of being sanctioned—make sure, in

each case, that they have checked to see whether such an order governs, and if so, to adhere to it.

In this article, we outline what led to this judicial response, describe the various standing orders that have been issued thus far, outline some of the concerns they raise, discuss the technical issue that is causing the problem and solutions that are currently available or on the horizon, and finally, propose what we believe to be a better alternative: public notice and/or consistent, court-wide rules, following publication and public comment.

#### **A. The Shot Heard 'Round the World**

Alarms went off on May 27, 2023, when The New York Times reported on a case<sup>1</sup> where the court issued an Order to Show Cause why plaintiff's counsel should not be sanctioned for papers they filed in opposition to a motion to dismiss that were "replete with citations to non-existent cases. . . . Six of the submitted cases appear[ed] to be bogus judicial decisions with bogus quotes and bogus internal citations."<sup>2</sup> It turned out that one of the attorneys in question had used ChatGPT to perform legal research, "a source that ha[d] revealed itself to be unreliable."<sup>3</sup> Possibly, an understatement.

In the immediate aftermath of that unfortunate brouhaha, several courts proactively issued standing orders to prevent such events from happening in their own courtrooms. Just three days later, Judge Brantley Starr of the U.S. District Court for the

---

\* Maura R. Grossman, J.D., Ph.D. and Daniel G. Brown, Ph.D., are professors in the David R. Cheriton School of Computer Science at the University of Waterloo. Dr. Grossman is also an adjunct professor at Osgoode Hall Law School of York University and an affiliate faculty member of the Vector Institute of Artificial Intelligence. Hon. Paul W. Grimm (ret.) is the Director of the Bolch Judicial Institute and the David F. Levi Professor of the Practice of Law at Duke Law School. Previously, he served as a District Judge (and before that as Magistrate Judge) in the U.S. District Court for the District of Maryland. Drs. Grossman and Brown's work is funded, in part, by the National Science and Engineering Council of Canada ("NSERC"). The authors wish to acknowledge Jason R. Baron and Amy Sellers for their thoughtful comments on a draft of this article. The views expressed in this article are the authors' own, and do not necessarily reflect the opinions of the institutions with which they are affiliated.

<sup>1</sup> Benjamin Weiser, *Here's What Happens When Your Lawyer Uses ChatGPT*, New York Times (May 27, 2023), available at <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>.

<sup>2</sup> *Mata v. Avianca, Inc.*, Case No. 22-cv-1461 (PKC), Order to Show Cause (S.D.N.Y. May 4, 2023), available at <https://www.documentcloud.org/documents/23826753-judgeaskingtheotherlawyerwhyhesubmittedafilingwithfakecases>.

<sup>3</sup> Benjamin Weiser, *supra*, n.1.

Northern District of Texas was the first to issue such a standing order, on May 30, 2023.<sup>4</sup> He requires attorneys and *pro-se* litigants appearing before him to file—on appearance in his court—a certificate indicating whether any portion of their filings will be drafted using GenAI tools. The standing order states in relevant part:

All attorneys and pro se litigants appearing before the Court, must, together with their notice of appearance, file on the docket a certificate attesting either that no portion of any filing will be drafted by generative artificial intelligence (such as ChatGPT, Harvey.AI, or Google Bard) or that any language drafted by generative artificial intelligence will be checked for accuracy, using print reporters or traditional legal databases, by a human being. . . . Any party believing a platform has the requisite accuracy and reliability for legal briefing may move for leave and explain why. Accordingly, the Court will strike any filing from a party who fails to file a certificate on the docket attesting that they have read the Court’s judge-specific requirements and understand that they will be held responsible under Rule 11 for the contents of any filing that they sign and submit to the Court, regardless of whether generative artificial intelligence drafted any portion of that filing.

A week later, on June 6, 2023, Judge Michael M. Baylson of the U.S. District Court for the Eastern District of Pennsylvania issued an order requiring attorneys and pro-se litigants to disclose the use of AI in drafting pleadings.<sup>5</sup> His order, however, was not limited in scope to the use of GenAI tools; rather, it referenced AI tools in general. The standing order stated:

If any attorney for a party, or a pro se party has used Artificial Intelligence (“AI”) in the preparation of any complaint, answer, motion, brief, or other paper files with the Court, and assigned to Judge Michael M. Baylson, **MUST**, in a clear and plain factual statement, disclose that AI has been used in any way in the

---

<sup>4</sup> Judge Brantley Starr, Judge Specific Requirements, *Mandatory Certification Regarding Generative Artificial Intelligence* (May 30, 2023), available at <https://www.txnd.uscourts.gov/judge/judge-brantley-starr>.

<sup>5</sup> *Standing Order Re: Artificial Intelligence (“AI”) Cases Assigned to Judge Baylson*, U.S. District Court for the Eastern District of Pennsylvania (June 6, 2023), available at <https://www.paed.uscourts.gov/documents/standord/Standing%20Order%20Re%20Artificial%20Intelligence%206.6.pdf>.

filing, and **CERTIFY**, that each and every citation to the law or the record in the paper, has been verified as accurate.<sup>6</sup>

Two days after that, on June 8, 2023, Magistrate Judge Gabriel A. Fuentes of the U.S. District Court for the Northern District of Illinois revised his standing order for civil cases,<sup>7</sup> to provide that:

The Court has adopted a new requirement in the fast-growing and fast-changing area of generative artificial intelligence (“AI”) and its use in the practice of law. The requirement is as follows: Any party using any generative AI tool to conduct legal research or to draft documents for filing with the Court must disclose in the filing that AI was used, with the disclosure including the specific AI tool and the manner in which it was used. Further, Rule 11 of the Federal Rules of Civil Procedure continues to apply, and the Court will continue to construe all filings as a certification by the person signing the filed document and after a reasonable inquiry, of the matters set forth in the rule, including but not limited to those in Rule 11(b)(2). . . . Just as the Court did before the advent of AI as a tool for legal research and drafting, the Court will continue to presume that the Rule 11 certification is a representation by filers, as living, breathing, thinking human beings, that they themselves have read and analyzed all cited authorities to ensure that such authorities exist and that the filings comply with Rule 11(b)(2). . . .

And, on the same day, Judge Stephen Alexander Vaden of the U.S. Court of International Trade issued a standing order<sup>8</sup> requiring the disclosure of any GenAI program used for drafting and extended the requirement further to demand a representation that the use of such an application had not resulted in the disclosure of any confidential or proprietary information to any unauthorized party. The relevant language of his order provides that:

---

<sup>6</sup> *Id.* (emphases in original).

<sup>7</sup> *Standing Order for Civil Cases Before Magistrate Judge Fuentes*, U.S. District Court for the Northern District of Illinois (May 31, 2023), at 2, available at [https://www.ilnd.uscourts.gov/assets/documents/forms/judges/Fuentes/Standing%20Order%20For%20Civil%20Cases%20Before%20Judge%20Fuentes%20rev%27d%205-31-23%20\(002\).pdf](https://www.ilnd.uscourts.gov/assets/documents/forms/judges/Fuentes/Standing%20Order%20For%20Civil%20Cases%20Before%20Judge%20Fuentes%20rev%27d%205-31-23%20(002).pdf).

<sup>8</sup> *Order on Artificial Intelligence*, U.S. Court of International Trade, The Honorable Stephen Alexander Valden, Judge (June 8, 2023), available at <https://www.cit.uscourts.gov/sites/cit/files/Order%20on%20Artificial%20Intelligence.pdf>.

Because generative artificial intelligence programs challenge the Court’s ability to protect confidential and business proprietary information from access by unauthorized parties, it is hereby:

**ORDERED** that any submission in a case assigned to Judge Vaden that contains text drafted with the assistance of a generative artificial intelligence program on the basis of natural language prompts, including but not limited to ChatGPT and Google Bard, must be accompanied by:

- (1) A disclosure that identifies the program used and the specific portions of text that have been so drafted;
- (2) A certification that the use of such program has not resulted in the disclosure of any confidential business proprietary information to any unauthorized party; and it is further

**ORDERED** that, following the filing of such notice, any party may file with the Court any motion provided for by statute or the Rules of the Court of International Trade seeking any relief the party believes the facts disclosed warrant.<sup>9</sup>

Not long thereafter, several Canadian courts followed suit. On June 23, 2023, the Court of King’s Bench of Manitoba issued a Practice Direction on the Use of Artificial Intelligence in Court Submissions,<sup>10</sup> advising that “when artificial intelligence has been used in the preparation of materials filed with the court, the materials must indicate how artificial intelligence was used.” Three days later, the Supreme Court of Yukon issued Practice Direction General-29 on the Use of Artificial Intelligence Tools,<sup>11</sup> which directed that “if any counsel or party relies on artificial intelligence (such as ChatGPT or any other artificial intelligence platform) for their legal research or submission in any matter and in any form before the Court, they must advise the Court of the tool used and for what purpose.” Law360 Canada has reported that the Supreme Court of Canada “is among the courts mulling whether and what practice direction to issue to counsel and litigants about

---

<sup>9</sup> *Id.* (emphases in original).

<sup>10</sup> Available at [https://www.manitobacourts.mb.ca/site/assets/files/2045/practice\\_direction\\_-\\_use\\_of\\_artificial\\_intelligence\\_in\\_court\\_submissions.pdf](https://www.manitobacourts.mb.ca/site/assets/files/2045/practice_direction_-_use_of_artificial_intelligence_in_court_submissions.pdf).

<sup>11</sup> Available at <https://www.yukoncourts.ca/sites/default/files/2023-06/GENERAL-29%20Use%20of%20AI.pdf>.

the use of artificial intelligence (AI) tools in the preparation of Supreme Court materials . . . .”<sup>12</sup>

## **B. Bringing a Cannon to a Sword Fight**

We can certainly appreciate why courts throughout North America reacted swiftly and decisively to the GenAI mishap in the Southern District of New York—which, regrettably, was subsequently repeated in a filing in the Tenth Court of Appeals in Waco, Texas, where an appellate brief contained “fabricated and non-existent citations.”<sup>13</sup> No judge wants to be faced with bogus anything, no less to discover that “[n]one of the three published cases cited actually exist in [a] Reporter,” and that “[e]ach citation provide[d] the reader a jump-cite into the body of a different case that ha[d] nothing to do with the proposition” for which it was cited.<sup>14</sup> But, we suggest that the solution proposed—a mosaic of inconsistent, individual standing orders—is not the best means to address the problem, especially when existing rules address the conduct at issue, and other institutions are better positioned to develop a more nuanced response.

We do not believe the courts that issued standing orders and practice directives intended to sow chaos, but the result has been a lack of clarity. There are many different GenAI and other AI technologies and some of the orders are not explicit about what technology use needs to be reported. For example, if a lawyer drafts a brief and uses Grammarly<sup>15</sup> to edit and revise their prose, does this need to be disclosed? Many online legal research databases already employ AI features for natural language querying.<sup>16</sup> Must the use of these tools be reported, even though there is no risk of fake citations? And, at what point does this reporting requirement begin to infringe on attorney work product and legal strategy?

---

<sup>12</sup> Cristin Schmitz, *SCC considers possible practice direction on use of AI in top court as more trial courts weigh in*, Law360 Canada (July 7, 2023), available at <https://www.law360.ca/articles/48377/scc-considers-possible-practice-direction-on-use-of-ai-in-top-court-as-more-trial-courts-weigh-in>.

<sup>13</sup> Lauren Berg, *Texas Appeals Court Calls Out Seemingly AI-Generated Cites*, Law360 (July 26, 2023), available at <https://www.law360.com/pulse/articles/1704217/texas-appeals-court-calls-out-seemingly-ai-generated-cites>.

<sup>14</sup> *Ex Parte Allen Michael Lee*, No. 10-22-00281-CR (10th Ct. App. TX July 19, 2023), at 2, available at <https://law.justia.com/cases/texas/tenth-court-of-appeals/2023/10-22-00281-cr.html>.

<sup>15</sup> See Grammarly Home Page at <https://www.grammarly.com/> (“generative AI writing assistant”).

<sup>16</sup> See, e.g., Westlaw Edge Home Page at <https://legal.thomsonreuters.com/en/products/westlaw-edge> (“powered by AI-enhanced capabilities that can help you research more effectively and be more strategic”).

Most search engines<sup>17</sup> and word-processing systems<sup>18</sup> will soon embed LLMs that will render GenAI ubiquitous in the daily tools that all litigators use. Rules of civil procedure should be technology neutral and should not have to be revised with the introduction of each new technological development. No one can predict what the legal technology landscape will look like two years from now.

Finally, a likely unintended consequence of these standing orders and practice directives is to impede legal innovation and access to justice. The legal profession is already sufficiently risk averse and technologically backward. These orders will serve to chill the use of technology that could not only enable unrepresented parties to access the justice system but also reduce the time and cost for those who can afford representation. We need a solution that is better tailored to the problem it is seeking to solve.

## II. We Have Seen the Enemy and It is Us

From the perspective of the courts, the most important new developments in GenAI are LLM-based tools that, in response to a prompt, generate text to fulfill the demands of the prompt. For example, a litigant might request that a tool “draft a complaint about a neighbor’s noisy dog,” or “find me a dog-noise case cite from Tennessee.” These tools will respond with text that appears akin to a pleading or case citation that could be filed in court. However, the goal of these LLMs is neither accuracy nor logical forms of argument per se, wherein lies the problem. Below we briefly discuss the history and underlying technology of GenAI systems and some of their limitations; the interested reader is referred to our [earlier] [forthcoming] article, *The GPT Judge: Justice in a Generative AI World*,<sup>19</sup> for more detail.

### A. What is the Peril and Where Did It Come From?

GenAI systems use deep-learning algorithms based on neural networks<sup>20</sup> to model written language, speech, music, or other pattern-based media. Typically, these systems

---

<sup>17</sup> See, e.g., Will Knight, *Google Just Added Generative AI to Search*, WIRED (May 18, 2023), available at <https://www.wired.com/story/google-io-just-added-generative-ai-to-search/>.

<sup>18</sup> See, e.g., Jared Spataro, *Introducing Microsoft 365 Copilot – your copilot for work*, Official Microsoft Blog (Mar. 16, 2023), available at <https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/>.

<sup>19</sup> Maura R. Grossman, Paul W. Grimm, Daniel G. Brown, and Molly (Yiming) Xu, *The GPTJudge: Justice in a Generative AI World*, 23 Duke L. & Tech. Rev. \_\_\_ (2023), Authors’ Copy available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4460184](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4460184).

<sup>20</sup> Deep learning consists of a series of machine-learning algorithms made up of multiple layers: an input layer, one or more hidden layers, and an output layer. The method is referred to as “deep learning” because, unlike previous approaches, one layer can feed its output to the next layer.



are trained on vast collections of human-generated work—often scraped from the Internet—that then generate new work using the properties identified in the training dataset. GenAI systems can also be tuned to specific tasks; for example, one can fine-tune an LLM on the available artwork of a single artist and then generate thousands of new works in the style of that creator, potentially flooding the market with synthetic competition to the creations of the living artist. Or the fine-tuning can be to a particular goal: one could, for example, train an LLM to write newspaper editorials from a particular political perspective. Some researchers and commercial entities have already developed special-purpose GenAI for conducting legal research or generating legal pleadings.<sup>21</sup>

Recent advances have allowed much faster training of these models, as has the availability of larger training datasets, which explains what has appeared to be the sudden emergence of this technology. In fact, ChatGPT, which incorporates OpenAI’s GPT 3.5 model, is simply the latest in a series of generative pre-trained (“GPT”) LLMs that were introduced in early 2018. Similarly, visual models like Dall-E 2, Midjourney, and Stable Diffusion, are built upon previous models dating back to the early 2010s. Perhaps, the primary reason for the recent emergence of so many such models is commercial: Corporations like Microsoft, Google, OpenAI, and Meta are all trying to claim market dominance and therefore, have been rushing GenAI products to market in the past year.

### **B. Why is GenAI So Good at Camouflage?**

GenAI is so hard to detect because the primary goal of its creators is to model the style of ordinary language, and because the models on which it is based have gotten better and massively more complex in a very short period of time. In particular, GenAI systems are trained on larger and larger datasets—of largely unknown provenance—which include many different types of writing, many different languages, and many different levels of fluency. LLM training datasets typically include publicly available news sources, Wikipedia articles, government documents, Reddit posts, and much more. Since this training data includes so many different styles of writing, the models learn the various common and distinctive patterns of these various forms, and on the surface, can convincingly mimic human-generated content.

GenAI systems also make use of humans to identify when they create unconvincing (or unacceptable) outputs. This approach, Reinforcement Learning with Human Feedback (“RLHF”) allows the parameters—a special kind of variable that is set during the training process—of a model to be tuned so that it will create more believable (or acceptable)

---

Each layer processes data in a manner inspired by the human brain, using interconnected nodes, hence the reason why they are often referred to as “neural networks.”

<sup>21</sup> See, e.g., *Meet Co-Counsel – the world’s first AI legal assistant*, Co-Counsel Home Page (Mar. 1, 2023), available at <https://casetext.com/>; *Generative AI for Elite Law Firms*, Harvey.ai Home Page, available at <https://www.harvey.ai/>.

outcomes. A similar approach, called a Generative Adversarial Network (“GAN”), models a game between two AI participants, where one player generates new material, and another player attempts to distinguish the generated material from authentic material by giving mathematical feedback to the generator, which updates and improves the generator’s output. This iterative process continues until the generator no longer improves. As a result, the better the distinguisher gets, the better the content generator gets, which explains why GenAI content is so hard to distinguish from human-generated content.

Some automated tools built by both LLM creators and by other researchers or providers have sought to identify whether a text is the output of an LLM or a human. LLMs often provide text that is more “unsurprising,” in a mathematical sense, than text generated by humans (that is, the individual words that appear in sentences are each on average more likely ones, according to the model, to occur in text written by humans). That property can be used to detect AI-generated text.<sup>22</sup> However, in a recent experiment, one such tool incorrectly identified text written by Non-Native English speaking students as having been crafted by GenAI: The smaller vocabularies and simpler sentence structure used by ESL students were flagged as hallmarks of AI generation.<sup>23</sup> Even OpenAI recently withdrew its ChatGPT detection tool (GPTZero) for lack of accuracy.<sup>24</sup>

Some other innovations have been suggested as ways of identifying the products of GenAI; for example, watermarking (*i.e.*, hiding an identifying marker in GenAI-produced text) that could allow one to later search for such an indicator in the text. However, since most LLMs do not watermark their output, one could simply use such an LLM as a last step in the creation process, asking the unmarked LLM to paraphrase the output of a watermarking LLM; this request might remove the watermark. The fact is, those intent on mischief will find ways to circumvent watermarks. Unfortunately, the arms race between creators and detectors will continue, and there is no reason to believe that the detectors (who are typically less well resourced) will win.

---

<sup>22</sup> However, there are now also AI-to-human text converters that will take AI-generated text and add variety, uniqueness, and complexity to the content to bypass AI content detectors. *See, e.g.*, Paraphrasing Tool AI’s AI To Human Text Converter, available at <https://paraphrasingtool.ai/ai-content-bypass-tool/>.

<sup>23</sup> Weixin Liang et al., *GPT detectors are biased against non-native English writers*, 4 Patterns 1-4 (July 2023), available at [https://www.cell.com/patterns/fulltext/S2666-3899\(23\)00130-7#%20](https://www.cell.com/patterns/fulltext/S2666-3899(23)00130-7#%20).

<sup>24</sup> Benj Edwards, *Unsafe at any seed—OpenAI discontinues its AI writing detector due to “low rate of accuracy,”* Ars Technica (July 26, 2023), available at <https://arstechnica.com/information-technology/2023/07/openai-discontinues-its-ai-writing-detector-due-to-low-rate-of-accuracy/>.

### C. How Does GenAI Sabotage the Truth?

One basic goal of GenAI is to model a style or a genre, like writing new poems in the style of Walt Whitman, or creating a satisfying werewolf romance story. These systems are not designed with accuracy as a goal, and they are not designed to engage in logical reasoning; indeed, their primary purpose is to create new content. GPT methods simply sample from a probability distribution of relevant words and phrases, and while there may be some bias toward truthful results, in some cases—because the truth may be more common—the model itself is unable to separate fact from fiction. Newer LLMs are attempting to create more trustworthy content, but building in proper legal reasoning and accurate citations is a tall order.

This inability becomes especially problematic when one attempts to perform legal research using GenAI on a tool that was not purpose built for that. ChatGPT 3.5 routinely cites irrelevant or non-existent cases, alongside relevant or real ones, because it is trying to fit the pattern of how one writes about the law; it is not necessarily trying to tell a true story. For example, in response to the prompt “find me a dog-noise case cite from Tennessee,” ChatGPT 3.5 provided a response that claimed to be based on a Tennessee dog-noise case, but actually mis-cited to a 2018 Texas Supreme Court medical-malpractice case (*Benge MD PLLC v. Williams*, Case No. 14-1057 (Tex. 2018)). And, when asked to write about the *Benge* case in the style of a newspaper article, ChatGPT 3.5 continued the incorrect pattern (“In a recent legal ruling, the Tennessee Court of Appeals addressed a contentious dispute between neighbors over incessant dog barking. The case of *Benge v. Williams* shed light on the complex issue of noise nuisances caused by pets and their potential impact on neighbors' quiet enjoyment of their property. . .”).

The phenomenon at issue here is referred to as “hallucinations,” and is to be expected of LLMs; indeed, many consider it a feature not a bug. Recall that the training goal of LLMs is to emulate the style of the text in the training dataset. Adding the word “not,” or removing “only,” for example, does not much change the overall fluency and apparent reasonableness of an LLM-generated sentence, but obviously changes the legal meaning dramatically. Similarly, a sentence in an GenAI-drafted legal brief may still fit the general structure of the text upon which the model was trained, regardless of whether the citations found in it are related to the subject under consideration. The reason why ChatGPT 3.5 consistently correctly associates *Obergefell v. Hodges*<sup>25</sup> with the topic of same-gender marriage is because the case is repeatedly mentioned in thousands of sentences about that subject in its training data; citations to less important or less well-known cases are less likely to be properly cited.<sup>26</sup> Even cases referenced in Wikipedia

---

<sup>25</sup> 576 U.S. 644 (2015).

<sup>26</sup> Another limitation of GenAI has to do with the date on which training of the system ceased. For example, ChatGPT 3.5 had a cutoff date of September 2021, so it cannot possibly cite to more

articles (such as Judge Grimm’s *Mancia v. Mayflower Textile Services Co.* opinion),<sup>27</sup> can be mis-analyzed by ChatGPT 3.5. It claims that briefs citing that opinion focus on overtime pay and labor standards (the overall subject of the *Mancia* litigation), when, in fact, Judge Grimm’s ruling was focused on the parties’ failure to cooperatively engage in the discovery process, in violation of Fed. R. Civ. P. 26(g), the proposition for which the case has been frequently cited.

Newer GenAI systems may eventually ameliorate this concern, for example, they can be trained to detect when a user is seeking a case citation and add a verification step to ensure that the output is valid and appropriate, and, as we previously mentioned, GenAI systems are being built specifically for the purpose of legal research. For the time being, however, *pro-se* filers will likely not have access to such specialized systems and will instead turn to free GenAI systems (like ChatGPT 3.5).

### III. Are There Already Sufficient Weapons in the Arsenal?

#### A. Rule 11

Part of our concern about the use of individual standing orders to regulate GenAI use is that they impose on parties and litigants obligations that already apply under existing rules of civil practice and procedure and/or ethical obligations presently imposed on lawyers by state rules of professional responsibility. Most notably, Fed. R. Civ. P. 11 requires that all pleadings motions, and other papers filed with the court in civil cases be signed by a lawyer, or if the party is not represented by counsel, by the party themselves. Failure to sign a pleading obligates the court to strike the filing unless the omission is “promptly corrected after being called to the attorney’s or party’s attention.”<sup>28</sup> The individual’s signature on the pleading makes several specific representations to the court. Namely, that “whether by signing, filing, submitting, or later advocating” what the pleading discusses, the “attorney or unrepresented party certifies that to the best of the person’s knowledge, information, and belief, formed after an inquiry reasonable under the circumstances,”<sup>29</sup> that (1) it is not being presented for any improper purpose, such as “to harass, cause unnecessary delay, or needlessly increase the cost of litigation”; (2) that the claims, defense, and legal contentions are “warranted by existing law or by a nonfrivolous argument for extending modifying or reversing existing law or for establishing new law”; (3) the factual contentions have evidentiary support or, if specifically so identified, will likely have evidentiary support after a reasonable

---

recent cases. While newer, purpose-built tools will have more up-to-date cutoffs, unless the training is continually refreshed, there will always be issues involving (in)completeness.

<sup>27</sup> 253 F.R.D. 354 (D. Md. 2008).

<sup>28</sup> Fed. R. Civ. P. 11(a).

<sup>29</sup> Fed. R. Civ. P. 11(b).

opportunity for further investigation or discovery”; and (4) the denials of factual contentions are warranted on the evidence, or if specifically so identified, are reasonably based on belief or a lack of information.”<sup>30</sup>

Lawyers or *pro-se* litigants who blindly rely on factual contentions taken from GenAI applications, or who rely—without independently confirming—on cases cited by such applications clearly have failed to conduct a reasonable inquiry, are filing a pleading that likely will cause unnecessary delay or increase litigation costs, are stating facts that are not based on existing law, and are presenting factual arguments without evidentiary support. The consequences of violating Rule 11 can be severe. If the court determines that Rule 11 has been violated, it may sanction any lawyer, law firm, or party that violated the rule or is responsible for it having been violated.

Thus, the standing orders that we have described above appear to be redundant. If the consequences of failing to comply with Rule 11 do not adequately deter the type of conduct that courts have criticized regarding the use of GenAI, it is hard to imagine what additional deterrence a judge’s individual standing order would add.

#### **B. Rule 26(g)**

On its face, Fed. R. Civ. P. 11 applies only to pleadings, motions, and other “papers,” and is inapplicable to discovery.<sup>31</sup> But this does not mean that there are no procedural impediments to a lawyer improperly using GenAI during the discovery phase of a civil case. Indeed, Rule 26(g)(1), which applies to “disclosures and discovery requests, responses, and objections” also requires that every discovery related disclosure, discovery request, response, or objection must be signed by an attorney or party, if unrepresented. As with Rule 11, the Rule 26(g)(1) signature “certifies that to the best of the person’s knowledge, information, and belief formed after a reasonable inquiry” that: the disclosure is complete and correct as of the time it was made; and that a discovery request, response, or objection is (a) consistent with the discovery rules and warranted by existing law (or a nonfrivolous argument for extending, modifying, or reversing existing law, or establishing new law), (b) is not interposed for an improper purpose (such as harassing an opponent, imposing unnecessary delay, or needless increase in the cost of litigation), and (c) is neither unreasonable nor unduly burdensome or expensive, considering the needs of the case, the amount in controversy in the case, and the importance of the issues at stake in the litigation.<sup>32</sup> If a party or attorney omits the required signature, the opposing counsel or party is under no duty to act on the unsigned discovery matter until it is signed, and the court must strike the unsigned discovery matter

---

<sup>30</sup> Fed. R. Civ. P. 11(b) (1)-(4).

<sup>31</sup> Fed. R. Civ. P. 11(d) “[Rule 11] does not apply to disclosures and discovery requests, responses, objections, and motions under Rules 26 through 37 [which deal with discovery].”

<sup>32</sup> Fed. R. Civ. P. 26(g)(1)(A)-(B)(i)-(iii).

unless the signature is promptly supplied when called to the attention of the lawyer or party. If a certification violates Rule 26(g), the offending lawyer and or party may be sanctioned.<sup>33</sup>

Accordingly, between Fed. R. Civ. P. 11 and 26(g), lawyers or parties who violate these rules in connection with their use of GenAI in civil litigation are already subject to sanctions that can be strong medicine—depending on the extent of the violation—regardless of whether the presiding judge has issued their own standing order concerning the use of GenAI. Moreover, if widespread public humiliation over being sanctioned by a court for committing this kind of error is insufficient disincentive, the Rules of Professional Conduct also impose independent ethical obligations on attorneys who use GenAI applications to refrain from the types of misconduct that have led courts to adopt standing orders prohibiting or regulating the use of these applications.

**C. American Bar Association Model Rules of Conduct 1.1 Comment [8], 3.3, and 1.6 (and their State-Law Equivalent)**

All attorneys are required to be licensed by the states or provinces in which they practice, and each jurisdiction has adopted rules of professional conduct that lawyers must follow, lest they be sanctioned or have their license suspended or revoked. While each jurisdiction has adopted its own ethics code, almost all of them (at least in the U.S.) follow or are guided by the American Bar Association’s (“ABA’s”) Model Rules of Professional Conduct. Three of the Model Rules impose ethical duties on attorneys that are specifically implicated by the behavior we have referenced in this article, resulting from the improper use of GenAI.

First, Model Rule 1.1 requires lawyers to provide their clients with competent representation, which requires the legal knowledge, skill, thoroughness, and preparation reasonably necessary for the representation. Comment [8] to Rule 1.1 provides that, to maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology. GenAI is clearly a relevant technology to the practice of law today, and lawyers must understand its strengths and weaknesses in order to provide competent representation.

Second, Model Rule 3.3 imposes an ethical obligation on attorneys to demonstrate candor to courts and other tribunals and prohibits a lawyer from making a false statement of fact or law to a tribunal or failing to correct a false statement of material fact or law previously made to the tribunal. Citing non-existent case law or misrepresenting the holdings of a case in a brief is making a false statement to a court. It does not matter if GenAI told you so.

---

<sup>33</sup> Fed. R. Civ. P. 26(g)(3).

And third, Model Rule 1.6 prohibits lawyers from revealing information relating to the representation of a client unless the client gives informed consent. Entering confidential client information into a publicly available, third-party chatbot is not consistent with this duty. Lawyers should not need to be told this.

Now that counsel have been warned in two highly publicized recent cases, a lawyer who does not adequately understand the risks inherent in using GenAI to produce either factual or legal content to be included in a court filing, and who fails to independently verify the accuracy of factual matters and/or legal authority obtained from GenAI has failed to represent their client competently. Any competent attorney should now know that GenAI can hallucinate. Moreover, a lawyer who obtains factual information or legal authority from GenAI and uses it in a pleading without independently confirming the accuracy of their representations is failing to adhere to the obligation of candor to the court if those representations turn out to be false. Similarly, a lawyer who discloses information about the representation of a client by using such information to prompt a search using GenAI without first having explained to their client the risks and obtained the client's consent to such disclosure, has failed to properly maintain the confidentiality of client information. None of these duties should require a separate certification; they are already basic requirements to practice law in virtually every jurisdiction.

A judge who determines that a lawyer has used GenAI in a manner that fails to conform with their ethical duties can refer the lawyer to the bar authority or law society of the jurisdiction where the lawyer is licensed, and that body will likely initiate an ethics investigation that could result in sanctions against the lawyer, up to and including loss of their license to practice. When state bar authorities or provincial law societies receive complaints against an attorney from a court, they characteristically investigate them with the utmost seriousness. Therefore, a lawyer who engages in the type of misconduct that judges fear will happen with GenAI are risking more than the wrath of a single judge—they are putting their ability to practice law at risk. Bar associations and law societies should provide guidance and education to their members and remove this burden from the shoulders of individual judges.

Viewed both individually and collectively, existing rules of civil practice and procedure and ethical codes of conduct already provide adequate deterrence to the misuse of GenAI in litigation, and, if violated, provide sanctions that are at least as severe—if not more so—than can be imposed for failing to comply with a court's individual standing order regarding use of GenAI.

#### **IV. Why Not Offer an Olive Branch?**

We believe that individualized standing orders are unnecessary, create unintended confusion, impose unnecessary burden and cost, and deter the legitimate use of GenAI applications that could increase productivity and access to justice. We do not, however,

suggest that judges and courts should idly sit by on the sidelines and avoid engaging with issues regarding the use of GenAI in the justice system. Rather, if district courts feel the need to address this issue, they can issue local rules that apply court-wide.<sup>34</sup> These rules are enacted after having been published, with the public given a chance to comment on the proposed rules. A well-crafted local rule governing the use of GenAI tools, adopted after publication and public comment, is more likely to address definitional and scope issues in a more nuanced way and to expose any unintended adverse consequences implicated by such a rule.

It may be the case that a new rule is not necessary at all and that notice would be the better approach for the time being. There is certainly no harm in individual judges including in their standing orders a warning to litigants of the risks inherent in the use of GenAI for conducting research and generating pleadings, and the consequences of misrepresentations to the court, but, as we mentioned above, there may already be sufficient deterrence to attorneys from repeating this error. For the benefit of *pro-se* litigants, courts can give notice to the public in general (*e.g.*, on their websites) that the use of GenAI tools in connection with court filings must be consistent with the obligation to verify the accuracy of factual and legal representations, including validating all citations, and explain the potential sanctions that can be imposed for failure to do so. Additionally, we see no problem with requiring *pro-se* litigants to disclose whether they have had any assistance in drafting their court filings. This would be similar to the requirement that already is imposed by ABA Formal Opinion 07-447 (2007), which requires an attorney who has provided assistance to a party in drafting a court filing, but who has not entered an appearance as counsel for that party, to disclose to the court the assistance they provided.

It is evident that the use of AI applications—and GenAI in particular—will be increasingly common in the court system. However, we urge caution and restraint in imposing additional disclosure and certification obligations—particularly when the scope of such requirements may be ambiguous—which impose unnecessary and inconsistent burdens on litigants. It is possible, in this instance, that honey may work better than vinegar.

---

<sup>34</sup> See 28 U.S.C. §2071(b) (“Any rule prescribed by a court, other than the Supreme Court, under subsection (a) shall be prescribed only after giving appropriate public notice and an opportunity for comment.”); Fed. R. Civ. P. 83(a)(1) (“After giving public notice and an opportunity for comment, a district court, acting by a majority of its district judges, may adopt and amend rules governing its practice.”).