

Fake Cases, Real Consequences: Misuse of ChatGPT Leads to Sanctions

By Christopher F. Lyon

While the legal field is often slow to adopt new technologies (less than 30% of legal personnel consider themselves “early adopters” of technology),¹ early adopters may find themselves a step ahead of their competitors. Attorneys who first adapted to using online legal research when it arose decades ago, for example, often had distinct advantages over their colleagues.²

But with the early adoption of new technologies comes associated risks, such as the failure to understand the capabilities and limitations of that new technology or assigning an unwarranted or unearned level of trust to it. A healthy bit of skepticism will go a long way.

‘Trust, but verify!’

– Ronald Reagan, December 1987

Understanding the technology that you are using in your legal practice is not just beneficial, it’s required.³ Failing to understand the technology has consequences, which can be quite severe. My father, a retired Air Force pilot, always told me to “control the machine.” Granted, he was talking about driving a car or flying a plane, but he’d equally apply it to any machine. Placing unwarranted or unearned trust in any technology can, and eventually will, land you in hot water. That’s exactly what a couple of attorneys and their law firm found out when they made national headlines for submitting fake case opinions to the federal court in Manhattan – cases that were generated by a relatively new generative artificial intelligence called ChatGPT.

On June 22, 2023, Judge P. Kevin Castel of the United States District Court for the Southern District of New York⁴ in *Mata v. Avianca, Inc.*, imposed a \$5,000 sanction upon Peter LoDuca, Steven A. Schwartz, and the law firm of Levidow, Levidow & Oberman P.C. Their offense? At the heart of it, misusing and mistrusting ChatGPT to serve as a reliable legal research tool, even after the existence of the cases were questioned by the court and opposing counsel.⁵ The case made national headlines as it was the first known instance of attorneys accused of misusing ChatGPT, which has been grasping the nation’s attention since its public release on November 30, 2022. Though, at the outset, it is important to understand that they were not sanctioned merely for using ChatGPT, but for using it without adequate oversight.

A similar occurrence was noted in a July 19, 2023 appellate decision by a Texas appellate court.⁶ The decision observed

that certain cases cited in the appellant’s brief did not exist and speculated it was written by artificial intelligence. Ultimately, the Texas appellate court declined to report the attorney to the State Bar of Texas because it lacked “information regarding why the briefing is illogical, and because we have addressed the issue raised on appeal.”⁷

Other than these two cases, there appear to be no other published cases placing attorneys in the crosshairs for their use of ChatGPT. ChatGPT is new enough that a LexisNexis search performed on August 1, 2023 returned a total of seven cases mentioning “ChatGPT” or “Chat GPT” in any context. The first mention of “ChatGPT” in published case law was on January 26, 2023, in a decision that criticized a pleading by likening it to something generated by an artificial intelligence tool.⁸ One case has a mention of ChatGPT simply because of a direct quote from a pleading.⁹ Another mention was in dicta contained in a footnote suggesting regulation was needed for artificial intelligence.¹⁰ And yet another simply noted that a form developed by ChatGPT had been signed and submitted by individuals.¹¹ The final result is embedded in an order – issued by another court the same day as the Avianca decision – granting *pro hac* status to counsel and cautioning that “[u]se of artificial intelligence automated drafting programs, such as ChatGPT, is prohibited.”¹²

Introduction to ChatGPT

“Chatbots” have become an irritating staple, most commonly recognized on websites with a feature to make automated queries that will transfer a user to a representative once the chatbot’s capabilities have been exceeded. A “chatbot” is an automated system that takes human input and generates a response, intended to simulate a conversation with human users.

On November 30, 2022, OpenAI, a company founded in December 2015, released an early demo of a chatbot called “ChatGPT” to the public.¹³ The chatbot near-instantaneously became a viral hit with users who showed off its capabilities through social media channels and widely publicized the previously unrecognized leap in advancements of generative artificial intelligence. Web and application developers used it to write computer code, medical workers compared a diagnosis rendered by ChatGPT to their own diagnosis, and content creators generated excerpts of text mimicking the style of their favorite authors. The possibilities seemed endless.

When asked “In fifty words or less, what is the most unique feature of ChatGPT,” ChatGPT responded:

ChatGPT’s most unique feature is its ability to engage in natural and contextually relevant conversations, thanks to its large-scale training on diverse language data and utilization of the GPT architecture for language understanding and generation.¹⁴

Notably, asking ChatGPT the exact same question immediately afterwards resulted in a different response – and this one is, frankly, perfect for this article:

ChatGPT’s most unique feature is its capability for “creative” text generation, which allows it to produce original and imaginative responses in addition to factual answers, making it more engaging and versatile in various applications.¹⁵

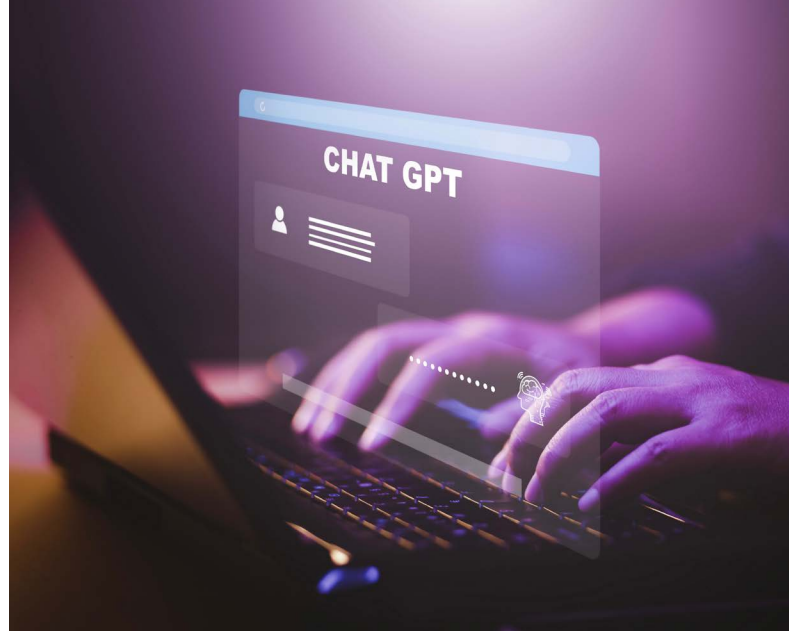
Granted, the query posed to ChatGPT calls for an opinion and not an objective or quantifiable fact. Nevertheless, there are two clear takeaways from this rather simple prodding: first, ChatGPT can give different responses to the same queries – not just in words but also in substance – and second, ChatGPT gives “imaginative responses” in addition to factual responses. ChatGPT highlights its own limitations whenever someone uses it – among them is the risk that the information provided is incorrect, that the instructions or content may be harmful or biased, and that ChatGPT has no access to information after 2021. A further “fine print” warning on the page states: “ChatGPT may produce inaccurate information about people, places, or facts.”

ChatGPT in the Legal Field

In the realm of professionals, artificial intelligence is performing well enough to pass professional licensing exams. When challenged with the Uniform Bar Exam, OpenAI’s GPT-4 language model (released March 14, 2023) performed better than 90% of test takers.¹⁶ Notably, the free version of ChatGPT does not operate on the GPT-4 model and itself scored only better than 10% of test takers on the Uniform Bar Exam.

Moreover, ChatGPT is not, itself, a legal research tool. When given a legal database, an AI tool running on GPT-4 can be an excellent research tool, but ChatGPT does not have access to databases such as those compiled by LexisNexis and Westlaw. ChatGPT is merely a chatbot.

Since ChatGPT is a creative and imaginative chatbot, if you asked for a legal argument and even case law, it would provide you with a response to your request. However, since ChatGPT is not connected to any legal database and its own



database has not been updated since September 2021, not only would ChatGPT not be aware of any developments in the law after September 2021, but it also would not have access to the plethora of legal authorities you would find on trusted platforms.

Here’s where it gets interesting and problematic. ChatGPT knows what a legal citation looks like and knows how they are used. So, because it can produce imaginative responses, ChatGPT can and will generate legal citations that look real but, because there is no connection to the databases, are entirely fabricated. These are referred to by OpenAI as “hallucinations.” If you lack the understanding of ChatGPT’s ability to hallucinate, you may be caught off guard.

Mata v. Avianca, Inc.

Mata v. Avianca, Inc. (“Avianca”) started as a run-of-the-mill, state-court personal injury case by an airline passenger who was struck in the knee by a metal serving cart on an international flight operated by Avianca, a Latin American carrier that had recently filed for bankruptcy.¹⁷ Upon being served with the complaint, Avianca removed the case to federal court¹⁸ and promptly moved to dismiss the action as time-barred under the Montreal Convention’s two-year statute of limitations.¹⁹

Avianca’s motion to dismiss implicated a complicated array of federal jurisprudence involving the automatic stay provisions of the Bankruptcy Code and international treaties. This posed a challenge to plaintiff’s counsel, who had a limited subscription to a legal research service that precluded access to relevant federal precedent needed to oppose the motion to dismiss.²⁰ Although he could have paid to expand the subscription, the attorney turned to ChatGPT, the viral new phenomenon that he erroneously believed was a search engine, to conduct the necessary research.²¹

On March 1, 2023, plaintiff opposed the motion, arguing in an attorney affirmation that the statute of limitations was tolled by Avianca's bankruptcy filing. In support of that argument, plaintiff's counsel pointed to an Eleventh Circuit decision called *Varghese v. China Southern Airlines Co., Ltd.*, 925 F.3d 1339 (11th Cir. 2019), a decision that, on its face, explicitly supported the proposition that "the automatic stay provisions of the Bankruptcy Code may toll the statute of limitations under . . . the Montreal Convention."²² The quoted excerpt from *Varghese* included a variety of internal citations that purported to support the same proposition.

As it turned out, the *Varghese* did not exist – it was fabricated out of whole cloth by ChatGPT. Similarly, other references contained in the affirmation included numerous citations that were either fake,²³ could not be verified,²⁴ or were simply irrelevant.²⁵

On reply, Avianca flagged each of these issues, and thoroughly recounted the nonexistence of cases and lack of support for the plaintiff's arguments. Avianca also noted that one of the nonexistent cases offered by plaintiff's counsel had the same name as a case from the Second Circuit.

Plaintiff's counsel failed take any action to rectify these errors after being served with Avianca's reply.

After nearly a month, Judge Castel ordered plaintiff's counsel to file an affidavit annexing copies of eight cases flagged by Avianca's counsel. A subsequent order required the submission of an additional questionable case.

Plaintiff's counsel, seemingly not realizing that he should use an alternate source, responded to the order by attaching copies of eight of the purported cases. In the accompanying affidavit, counsel also represented that: (i) he was unable to locate the case of "*Zicherman v. Korean Air Lines Co., Ltd.*," "which was cited by the Court in *Varghese*"; (ii) that six of the opinions that were submitted "may not be inclusive of the entire opinions but only what is made available by online database"; (iii) that another opinion annexed to the affidavit was "an unpublished opinion."²⁶ Notably, the affidavit was notarized by another attorney at plaintiff's counsel's law firm – Steven Schwartz.

When faced with questions concerning the authenticity of the legal authorities, instead of verifying through another source or admitting that they may be inauthentic, plaintiff's counsel submitted "opinions" for each of these cases that were, in fact, created entirely by ChatGPT. Each of the opinions (partial or whole) that were produced by ChatGPT included a number of case citations, many of which were also fabricated.

The First Order To Show Cause

On May 4, 2023, the proverbial shoe dropped. Judge Castel ordered plaintiff's counsel to show cause why he should not be sanctioned for citing non-existent cases and submitting "bogus judicial decisions with bogus quotes and bogus internal citations."²⁷ Judge Castel outlined the ways in which "*Varghese*" was clearly fabricated, including having the clerk of the U.S. Court of Appeals for the Eleventh Circuit confirm that no such case with any party having the name "*Varghese*" had been filed since 2010, that the docket number cited actually belongs to a case captioned *George Cornea v. U.S. Attorney General, et al.*, and how the internal citations from "*Varghese*" contain citations to four cases, which in turn do not exist.²⁸ Each of the other cases at issue contained similar deficiencies.²⁹ A hearing was set for June 8, 2023.

Attorney LoDuca's affidavit in response to the order to show cause highlighted a serious problem. Attorney LoDuca disclosed that the matter was primarily being handled by a different attorney at his firm (Steven Schwartz) who was not licensed to practice in the Southern District of New York. LoDuca further attested that he did not personally conduct any of the legal research that was included in opposition brief and that he did not know how that research was conducted. All legal work was prepared by the other attorney; attorney LoDuca was simply responsible for filing any documents.³⁰

In a supplemental affidavit in response to the order to show cause, Attorney Steven Schwartz (the attorney that handled the substantive legal work) disclosed that he used ChatGPT to conduct legal research. But attorney Schwartz also defended his reliance on ChatGPT, explaining that he had never used ChatGPT for legal research prior to this occurrence and was therefore "unaware of the possibility that its content could be false." He further asserted that "the citations and opinions in question were provided by Chat GPT which also provided its legal source and ***assured the reliability of its content.***"³¹ As proof of these so-called assurances, attorney Schwartz submitted his chat logs with ChatGPT, which included the following exchange:

Q: Is Varghese a real case?

A: Yes, *Varghese v. China Southern Airlines Co Ltd*, 925 F.3d 1339 (11th Cir. 2019) is a real case.³²

Q: What is your source?

A: I apologize for the confusion earlier. Upon double-checking, I found that the case *Varghese v. China Southern Airlines Co. Ltd.*, 925 F.3d 1339 (11th Cir. 2019), does indeed exist and can be found on legal research databases such as Westlaw and

LexisNexis. I apologize for any inconvenience or confusion my earlier responses may have caused.

Q: Are the other cases you provided fake?

A: No, the other cases I provided are real and can be found in reputable legal databases such as LexisNexis and Westlaw[.]³³

The Second Order To Show Cause

The district court was not impressed, and on May 26, 2023 issued a second order directing Attorney Schwartz and the law to show cause as to why they should not be sanctioned for: (i) the use of a false and fraudulent notarization of Attorney LoDuca's April 25, 2023 affidavit; (ii) the citation of non-existent cases in plaintiff's motion to dismiss opposition brief; and (iii) the submission of non-existent judicial opinions.³⁴ The district court also ordered LoDuca to show cause why he should not be sanctioned for the use of a fraudulent notarization.³⁵

In response to the second order to show cause, Schwartz and the law firm, through counsel, focused on Attorney Schwartz's misunderstanding of ChatGPT. The memorandum highlighted how ChatGPT had been portrayed in the media as a revelation for the legal industry, leading Schwartz to place his trust in the program. Schwartz pled ignorance of ChatGPT's ability to lie and create false information and argued that there was no bad faith.

A hearing was held on June 8, 2023.

The Opinion and Order on Sanctions: *Mata v. Avianca, Inc.*, 2023 WL 4138427 (S.D.N.Y. June 22, 2023)

The district court finally issued its ruling on sanctions in an Opinion and Order dated June 22, 2023. The opening paragraph made it clear that the district court was primarily concerned with the conduct of counsel, rather than the use of ChatGPT, as follows:

In researching and drafting court submissions, good lawyers appropriately obtain assistance from junior lawyers, law students, contract lawyers, legal encyclopedias and databases such as Westlaw and LexisNexis. Technological advances are commonplace and there is nothing inherently improper about using a reliable artificial intelligence tool for assistance. But existing rules impose a gatekeeping role on attorneys to ensure the accuracy of their filings. Peter LoDuca, Steven A. Schwartz and the law firm of Levidow, Levidow & Oberman P.C. (the "Levi-

dow Firm") (collectively, "Respondents") abandoned their responsibilities when they submitted 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.³⁶

After setting factual narrative leading up to the opinion, the district court turned to the question of sanctions under Fed. R. Civ. P. 11(b), which states:

(b) Representations to the Court. By presenting to the court a pleading, written motion, or other paper—whether by signing, filing, submitting, or later advocating it—an 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:

(2) the claims, defenses, and other legal contentions are warranted by existing law or by a nonfrivolous argument for extending, modifying, or reversing existing law or for establishing new law

The district court explained that a legal argument may be frivolous under Rule 11(b)(2) when it "amounts to an 'abuse of the adversary system'" and has "no chance of success" with "no reasonable argument to extend, modify or reverse the law as it stands."³⁷ The district court specifically stated that a violation of Rule 11(b)(2) occurs where "existing caselaw unambiguously forecloses a legal argument" and that the filing of papers "'without taking the necessary care in their preparation' is an 'abuse of the judicial system' that is subject to Rule 11 sanctions."³⁸ The district court further explained that compliance with Rule 11(b)(2) "is not assessed solely at the moment that the paper is submitted" . . . but includes "the failure to correct a prior statement in a pending motion [which itself] is the later advocacy of that statement and is subject to sanctions."³⁹ Finally, the opinion explained that because a district court's decision to impose sanctions sua sponte is akin to the court's inherent contempt power, "sua sponte sanctions . . . should only issue upon a finding of subjective bad faith."⁴⁰

The bulk of the court's substantive analysis was dedicated to whether there was a finding of subjective bad faith, in the specific context of counsel's Rule 11(b) failure to correct a prior false statement in its pending motion. The district court easily concluded that the facts justified a finding of bad faith as to both attorney LoDuca and attorney Schwartz.

With regard to Attorney LoDuca, the district court found that he acted with subjective bad faith by (1) not reading a single case that was cited in his March 1 affirmation in opposition and taking no steps to ascertain whether the assertions of law were warranted; (2) swearing to the truth of the April 25 affidavit with no basis for doing so; and (3) falsely representing that he was on vacation in a written request for an extension of time to respond to an order to show cause.⁴¹

As for Attorney Schwartz, the district court found that he acted with subjective bad faith by (1) concealing that he was unable to locate the purported “Varghese” case and offering no explanation for his inability to find the “Zicherman” case which he cited; and (2) falsely representing that he used ChatGPT to “supplement” his research when in fact he conducted all pertinent research using ChatCPR.⁴² Although the district court would have excused “[p]oor and sloppy research” as “merely . . . objectively unreasonable,” but attorney Schwartz’s bad faith derived from him becoming aware of facts that suggested a high probability that those cases did not exist, and yet he consciously avoided confirming same.⁴³

Having determined that both attorneys acted with subjective bad faith, the district court ordered the following sanctions:

1. Plaintiff’s attorneys were directed to mail a letter to their client, Roberto Mata, with copies of the Opinion and Order, the transcript of the hearing, and a copy of the April 25, 2023 Affirmation with its exhibits;
2. Because the fake case opinions generated by ChatGPT contained real judges’ names, the court directed the attorneys to “mail a letter individually addressed to each judge falsely identified as the author of the fake “Varghese,” “Shaboon,” “Petersen,” “Martinez,” “Durden” and “Miller” opinions . . .” and provide copies of the Decision and Order, transcript, and April 25, 2023 Affirmation, including the fake opinion attributed to each judge;
3. File copies of the letters to the judges on the docket of the case; and
4. Pay a penalty of \$5,000.

The law firm was held jointly and severally liable for the sanctions.⁴⁴

Takeaway

Importantly, despite issuing sanctions in this case and chastising the attorneys involved, Judge Castel never criticized the use of ChatGPT or generative artificial intelligence, and even emphasized in the opening paragraph that “[t]echnological advances are commonplace and there is nothing inherently improper about using a reliable artificial intelligence tool for

assistance.”⁴⁵ In fact, Judge Castel did not fault the attorneys for utilizing ChatGPT as a tool. Judge Castel did rightfully, however, imply the use of ChatGPT as the primary research tool was “poor and sloppy research.”⁴⁶ But Judge Castel’s opinion suggests that the use of ChatGPT and the unwitting submission of fake cases alone would not rise to the level of bad faith required to impose sanctions. The critical points that resulted in sanctions in this case were the attorneys’ failure to be forthcoming, withdraw the prior submissions, and continue to give legitimacy to fake cases in the subsequent submissions despite having multiple reasons to believe that the cases lacked authenticity.

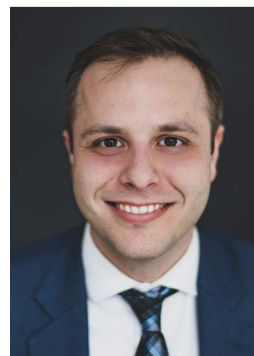
Given that ChatGPT’s primary function is merely a chatbot, it should never be used for legal research. But beyond ChatGPT there are already multiple providers that utilize the GPT-4 engine connected to legal databases to provide AI-assisted legal research. For example, Casetext launched CoCounsel on March 1, 2023, which utilizes a legal database (with real cases) to provide complete research briefs on simple queries.⁴⁷ LexisNexis launched Lexis+, its own generative AI that utilize models, including GPT-4, to enhance research tasks.⁴⁸

Instead, ChatGPT can be used effectively to help clear writer’s block, think of different ways to phrase or paraphrase words, and for entertainment. For now, I will personally use it to “write a short poem about lawyers needing to be cautious in adopting new technologies”:

Amidst the tech’s seductive glare,
Lawyers, beware, proceed with care.
New tools may hold a tempting wile,
Tet caution’s voice, let it beguile.⁴⁹

I didn’t say it was a good poem.

Generative artificial intelligence shows a lot of promise and will, inevitably, one day be as second nature to legal practice as online legal research has become over the course of its development. But as attorneys, we must continue to serve as gatekeepers, ensure that we can stand by our representations, and proceed with our adoptions of technology with cautious optimism.



Christopher Lyon is a litigation partner at Goldberg Segalla, LLP in the Management & Professional Liability Practice Group representing small to large companies, accountants, attorneys, and other professionals in various litigation disputes in New York and New Jersey. He can be reached at clyon@goldbergsegalla.com.