The New York Times | https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html

Here's What Happens When Your Lawyer Uses ChatGPT

A lawyer representing a man who sued an airline relied on artificial intelligence to help prepare a court filing. It did not go well.

By Benjamin Weiser

May 27, 2023

The lawsuit began like so many others: A man named Roberto Mata sued the airline Avianca, saying he was injured when a metal serving cart struck his knee during a flight to Kennedy International Airport in New York.

When Avianca asked a Manhattan federal judge to toss out the case, Mr. Mata's lawyers vehemently objected, submitting a 10-page brief that cited more than half a dozen relevant court decisions. There was Martinez v. Delta Air Lines, Zicherman v. Korean Air Lines and, of course, Varghese v. China Southern Airlines, with its learned discussion of federal law and "the tolling effect of the automatic stay on a statute of limitations."

There was just one hitch: No one — not the airline's lawyers, not even the judge himself — could find the decisions or the quotations cited and summarized in the brief.

That was because ChatGPT had invented everything.

The lawyer who created the brief, Steven A. Schwartz of the firm Levidow, Levidow & Oberman, threw himself on the mercy of the court on Thursday, saying in an affidavit that he had used the artificial intelligence program to do his legal research — "a source that has revealed itself to be unreliable."

Mr. Schwartz, who has practiced law in New York for three decades, told Judge P. Kevin Castel that he had no intent to deceive the court or the airline. Mr. Schwartz said that he had never used ChatGPT, and "therefore was unaware of the possibility that its content could be false."

He had, he told Judge Castel, even asked the program to verify that the cases were real.

It had said yes.

Mr. Schwartz said he "greatly regrets" relying on ChatGPT "and will never do so in the future without absolute verification of its authenticity."

Judge Castel said in an order that he had been presented with "an unprecedented circumstance," a legal submission replete with "bogus judicial decisions, with bogus quotes and bogus internal citations." He ordered a hearing for June 8 to discuss potential sanctions.

As artificial intelligence sweeps the online world, it has conjured dystopian visions of computers replacing not only human interaction, but also human labor. The fear has been especially intense for knowledge workers, many of whom worry that their daily activities may not be as rarefied as the world thinks — but for which the world pays billable hours.

Stephen Gillers, a legal ethics professor at New York University School of Law, said the issue was particularly acute among lawyers, who have been debating the value and the dangers of A.I. software like ChatGPT, as well as the need to verify whatever information it provides.

"The discussion now among the bar is how to avoid exactly what this case describes," Mr. Gillers said. "You cannot just take the output and cut and paste it into your court filings."

The real-life case of Roberto Mata v. Avianca Inc. shows that white-collar professions may have at least a little time left before the robots take over.

It began when Mr. Mata was a passenger on Avianca Flight 670 from El Salvador to New York on Aug. 27, 2019, when an airline employee bonked him with the serving cart, according to the lawsuit. After Mr. Mata sued, the airline filed papers asking that the case be dismissed because the statute of limitations had expired.

In a brief filed in March, Mr. Mata's lawyers said the lawsuit should continue, bolstering their argument with references and quotes from the many court decisions that have since been debunked.

Soon, Avianca's lawyers wrote to Judge Castel, saying they were unable to find the cases that were cited in the brief.

When it came to Varghese v. China Southern Airlines, they said they had "not been able to locate this case by caption or citation, nor any case bearing any resemblance to it."

They pointed to a lengthy quote from the purported Varghese decision contained in the brief. “The undersigned has not been able to locate this quotation, nor anything like it in any case,” Avianca’s lawyers wrote.

Indeed, the lawyers added, the quotation, which came from Varghese itself, cited something called Zicherman v. Korean Air Lines Co. Ltd., an opinion purportedly handed down by the U.S. Court of Appeals for the 11th Circuit in 2008. They said they could not find that, either.

Judge Castel ordered Mr. Mata’s attorneys to provide copies of the opinions referred to in their brief. The lawyers submitted a compendium of eight; in most cases, they listed the court and judges who issued them, the docket numbers and dates.

The copy of the supposed Varghese decision, for example, is six pages long and says it was written by a member of a three-judge panel of the 11th Circuit. But Avianca’s lawyers told the judge that they could not find that opinion, or the others, on court dockets or legal databases.

Bart Banino, a lawyer for Avianca, said that his firm, Condon & Forsyth, specialized in aviation law and that its lawyers could tell the cases in the brief were not real. He added that they had an inkling a chatbot might have been involved.

Mr. Schwartz did not respond to a message seeking comment, nor did Peter LoDuca, another lawyer at the firm, whose name appeared on the brief.

Mr. LoDuca said in an affidavit this week that he did not conduct any of the research in question, and that he had “no reason to doubt the sincerity” of Mr. Schwartz’s work or the authenticity of the opinions.

ChatGPT generates realistic responses by making guesses about which fragments of text should follow other sequences, based on a statistical model that has ingested billions of examples of text pulled from all over the internet. In Mr. Mata’s case, the program appears to have discerned the labyrinthine framework of a written legal argument, but has populated it with names and facts from a bouillabaisse of existing cases.

Judge Castel, in his order calling for a hearing, suggested that he had made his own inquiry. He wrote that the clerk of the 11th Circuit had confirmed that the docket number printed on the purported Varghese opinion was connected to an entirely different case.

Calling the opinion “bogus,” Judge Castel noted that it contained internal citations and quotes that, in turn, were nonexistent. He said that five of the other decisions submitted by Mr. Mata’s lawyers also appeared to be fake.

On Thursday, Mr. Mata’s lawyers offered affidavits containing their version of what had happened.

Mr. Schwartz wrote that he had originally filed Mr. Mata’s lawsuit in state court, but after the airline had it transferred to Manhattan’s federal court, where Mr. Schwartz is not admitted to practice, one of his colleagues, Mr. LoDuca, became the attorney of record. Mr. Schwartz said he had continued to do the legal research, in which Mr. LoDuca had no role.

Mr. Schwartz said that he had consulted ChatGPT “to supplement” his own work and that, “in consultation” with it, found and cited the half-dozen nonexistent cases. He said ChatGPT had provided reassurances.

“Is varghese a real case,” he typed, according to a copy of the exchange that he submitted to the judge.

“Yes,” the chatbot replied, offering a citation and adding that it “is a real case.”

Mr. Schwartz dug deeper.

“What is your source,” he wrote, according to the filing.

“I apologize for the confusion earlier,” ChatGPT responded, offering a legal citation.

“Are the other cases you provided fake,” Mr. Schwartz asked.

ChatGPT responded, “No, the other cases I provided are real and can be found in reputable legal databases.”

But, alas, they could not be.

Sheelagh McNeil contributed research.

Benjamin Weiser is a reporter covering the Manhattan federal courts. He has long covered criminal justice, both as a beat and investigative reporter. Before joining The Times in 1997, he worked at The Washington Post.

More from Benjamin Weiser

E. Jean Carroll Seeks New Damages From Trump for Comments on CNN
May 22, 2023

Bike-Path Killer Is Sentenced to Life as Victims Speak Out

May 17, 2023

A version of this article appears in print on , Section A, Page 1 of the New York edition with the headline: So, Have You Heard the One About the Lawyer Using A.I.?

The New York Times https://www.nytimes.com/2023/06/22/nyregion/lawyers-chatgpt-schwartz-loduca.html

ChatGPT Lawyers Are Ordered to Consider Seeking Forgiveness

Steven A. Schwartz and Peter LoDuca must pay a fine and send letters to judges named in a brief filled with fiction, a judge ordered.

By Benjamin Weiser

June 22, 2023

A Manhattan judge on Thursday imposed a $5,000 fine on two lawyers who gave him a legal brief full of made-up cases and citations, all generated by the artificial intelligence program ChatGPT.

The judge, P. Kevin Castel of Federal District Court, criticized the lawyers harshly and ordered them to send a copy of his opinion to each of the real-life judges whose names appeared in the fictitious filing.

But Judge Castel wrote that he would not require the lawyers, Steven A. Schwartz and Peter LoDuca, whom he referred to as respondents, to apologize to those judges, "because a compelled apology is not a sincere apology."

"Any decision to apologize is left to respondents," the judge added.

The discovery that ChatGPT had helped create the brief in an otherwise unremarkable lawsuit reverberated throughout the legal profession. The revelation also riveted the tech community, which has been debating the dangers of overreliance on artificial intelligence — even as a existential threat to humanity.

In the case involving Mr. Schwartz and Mr. LoDuca, Judge Castel made it clear they had violated a basic precept of the American legal system.

"Many harms flow from the submission of fake opinions," the judge wrote. "The opposing party wastes time and money in exposing the deception. The court's time is taken from other important endeavors."

The lawyers' action, he added, "promotes cynicism about the legal profession and the American judicial system. And a future litigant may be tempted to defy a judicial ruling by disingenuously claiming doubt about its authenticity."

Thursday's ruling followed a June 8 hearing at which Judge Castel grilled Mr. Schwartz and Mr. LoDuca about how they came to file the brief. In the suit, their client, Roberto Mata, sought to hold the airline Avianca responsible for an injury he says he sustained when a metal serving cart hit his knee during an August 2019

flight from El Salvador to New York.

After Avianca asked to dismiss the suit because the statute of limitations had expired, Mr. Schwartz prepared a 10-page brief citing more than a half-dozen court decisions with names like Martinez v. Delta Air Lines, Varghese v. China Southern Airlines and Zicherman v. Korean Air Lines, to argue that the litigation should be allowed to proceed.

Because Mr. Schwartz was not admitted to practice in federal court in Manhattan, his partner, Mr. LoDuca, became the lawyer of record and signed the brief, which was filed on March. 1.

Two weeks later, Avianca's lawyers, Bart Banino and Marissa Lefland, replied that they were "unable to locate most of the case law" cited in the brief.

When the judge then gave Mr. LoDuca a week to produce the cases mentioned, Mr. LoDuca responded that he was on vacation, and asked for another week. The judge agreed.

At the hearing on June 8, Mr. LoDuca admitted that he had not been on vacation but, because Mr. Schwartz was away, he wanted to give his colleague more time.

"The lie had the intended effect of concealing Mr. Schwartz's role," Judge Castel wrote.

In his opinion, Judge Castel examined the supposed decisions, demonstrating how they were clearly fabricated, and said the lawyers had acted in bad faith by submitting them.

The purported Varghese opinion was said to have been issued by a three-judge panel of the U.S. Court of Appeals for the 11th Circuit. But, Judge Castel noted, one of the judges the opinion listed actually sat on the federal appeals court for the 5th Circuit.

Beyond that, "the 'Varghese' decision shows stylistic and reasoning flaws that do not generally appear in decisions issued by United States Courts of Appeals," the judge said.

"Its legal analysis is gibberish," he wrote, adding, "The summary of the case's procedural history is difficult to follow and borders on nonsensical."

He said Mr. Mata's lawyers had abandoned their responsibilities, "then continued to stand by the fake opinions after judicial orders called their existence into question."

Had the matter ended with the lawyers "coming clean" earlier, the judge continued, "the record now would look quite different."

The judge noted that the lawyers' firm, Levidow, Levidow & Oberman, had arranged for outside lawyers to conduct a mandatory training program on technological competence and artificial intelligence programs. And he credited the lawyers' descriptions of their embarrassment and remorse in the face of widespread publicity about their actions.

The Levidow firm said in a statement that it had reviewed the judge's order and "fully intend to comply with it." But the firm disagreed with Judge Castel's finding that anyone at the firm acted in bad faith.

"In the face of what even the court acknowledged was an unprecedented situation, we made a good faith mistake in failing to believe that a piece of technology could be making up cases out of whole cloth," the firm said.

The firm said it was considering its options and had made no decision as to whether to appeal.

Ronald Minkoff, a lawyer for the firm and for Mr. Schwartz, declined to comment on Mr. Schwartz's behalf. A lawyer for Mr. LoDuca did not respond to a request for comment.

In a separate ruling on Thursday, Judge Castel dismissed Mr. Mata's lawsuit against Avianca on the statute-of-limitations grounds the airline had argued.

Mr. Banino, a lawyer for the airline, said, "Putting aside the lawyer's use of ChatGPT and submission of fake cases, the court reached the right conclusion in dismissing the underlying case."

In his ruling, the judge did not refer the lawyers for potential disciplinary action, but the disciplinary authorities could start their own investigation. Such inquiries can lead to a private reprimand or to public sanctions like suspension or disbarment.

Stephen Gillers, a legal ethics professor at New York University School of Law, said he believed the worldwide publicity about the case helped Mr. Schwartz and Mr. LoDuca avoid a worse fate.

"The lawyers will now and forever be known as 'the lawyers who got fooled by ChatGPT,' which Castel says is also a sanction," Professor Gillers said. "The case is the first, but not likely the last, warning to the bar not to get seduced by the siren call of generative A.I."

A correction was made on June 22, 2023*: Because of an editing error, an earlier version of the capsule summary with this article misstated the given name of a lawyer involved in the case. As the article correctly notes, he is Peter LoDuca, not Paul.*

Benjamin Weiser is a reporter covering the Manhattan federal courts. He has long covered criminal justice, both as a beat and investigative reporter. Before joining The Times in 1997, he worked at The Washington Post. More about Benjamin Weiser

A version of this article appears in print on , Section B, Page 6 of the New York edition with the headline: Lawyers Fined for Citing Cases A.I. Made Up

Published on *The National Law Review* https://www.natlawreview.com

California Bar to Set Guidelines for Using Generative AI to Practice Law, More States to Follow

Article By:
Seiko Okada, M.D., Ph.D.
Vincent Look
Dr. Christian E. Mammen

Can generative AI offer legal advice? On November 16, 2023, the State Bar of California approved guidelines to help lawyers navigate their ethical obligations when using generative artificial intelligence (AI). Titled "Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law," the guidance sets forth the initial recommendations of the Committee on Professional Responsibility and Conduct of the California State Bar regarding use of generative AI in practice of law. Guidance to California Lawyers is in line with the State Bar's Rules of Professional Conduct and the state's statutory authority and includes the following:

- Confidentiality – A lawyer must not input any confidential information of the client into generative AI solution unless the lawyer knows that the provider will not share the information with others or use the information for itself, such as to train or improve its AI product. In addition, the lawyer must anonymize the input so that it does not identify the client. In other words, be wary and treat the 'prompt' of a generative AI solution as you would the ears of a stranger.
- Competence – The lawyer must understand how the Generative AI solution works, including its limitations and potential use of client data. Further, the lawyer cannot simply trust that the output from the generative AI tool is correct, but must review and analyze these outputs to support 'the interests and priorities' of the client. Importantly, the guidelines state that the 'duty of competence requires more than the mere detection and elimination of false AI-generated results.' In other words, the lawyer cannot over-rely on the generative AI solution, because doing so would essentially result in a delegation of the lawyer's professional judgment to generative AI which should remain the lawyer's responsibility at all times.
- Communication regarding generative AI use – The lawyer should consider disclosure to their client that they intend to use generative AI in the representation, including how the technology will be used, and the benefits and risks of such use.
- Billing for AI work – A lawyer may use generative AI to more efficiently create work product and may charge for actual time spent on the legal work. While the time that is charged may include crafting or refining generative AI inputs, or reviewing and editing generative AI outputs, the lawyer must not charge hourly fees for the time saved by using generative AI. A fee agreement should explain the basis for all fees and costs, including those associated with the use of generative AI.

- Candor to the Tribunal – A lawyer must review all submissions that are made to the court for accuracy, including analysis and citation to case law. Generative AI already has a history of 'hallucinating' or making up non-existent case law and bogus quotations.
- Prohibition on discrimination, harassment, and retaliation – Some generative AI is trained on biased information, and a lawyer should be aware of possible biases and the risks they may create when using generative AI (e.g., to screen potential clients or employees).

California Bar's guidance marks a much needed first step for developing rules and regulations on AI use in law. Other state bars are also working on AI guidance – the Florida Bar's proposed opinion on lawyers' AI use is open for comment until January 2, 2024. In the Fall 2023 issue of the State Bar Journal, the North Carolina State Bar published an article by its ethics counsel listing key ethical considerations for the use of AI in the legal profession. Although the practice of law is governed by states, federal guidance may soon become imperative for consistency across the nation. Along the way, the fundamental questions to be addressed include identifying what activities constitute the practice of law and what activities if any, can leverage the use of AI. As AI increasingly becomes a necessary resource for lawyers to represent the clients competently and efficiently, it may make sense to require a license or certification for AI to participate in the practice of law, much like requiring a license or certification for human lawyers and paralegals. For example, for a generative AI solution to be allowed in the legal field, certification may require that the inputs and outputs of the AI solution are kept confidential and not shared. Similarly, certification may require additional cross-checking for AI-aided case law citations to address 'hallucinations.'

National Law Review, Volumess XIII, Number 334

Source URL: https://www.natlawreview.com/article/california-bar-set-guidelines-using-generative-ai-practice-law-more-states-follow

The State Bar *of California*

Ethics & Technology Resources

NEW! On November 16, 2023, the State Bar Board of Trustees approved the Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law, developed by the Committee on Professional Responsibility and Conduct to assist lawyers in navigating their ethical obligations when using generative artificial intelligence. The Practical Guidance will be a living document that is periodically updated as the technology evolves and matures, and as new issues are presented.

This page list resources addressing attorney professional responsibility issues that arise in connection with the use of websites, email, chat rooms, and other technologies. The resources include advisory ethics opinions, articles, and MCLE programs.

- Ethics Opinions
- Articles
- Online MCLE Programs

The above links organize items by the type of resources. The links below organize the same collection of resources by subject matter.

- Online Communication (email, chat, blogs, etc...)
- Electronic Files (cloud computing, ediscovery, metadata, virtual law office, etc...)
- Advertising (website and email content issues)
- Social Media
- Internet/Email Scams
- Miscellaneous

The links below are to the Rules of Professional Conduct that account for lawyer use of technology.

- Rule 1.1, Comment [1] - clarifies the duty of competence to include keeping abreast of technology in the practice of law (operative March 22, 2021)
- Rule 1.4, Comment [2] - updates the duty to provide copies of significant documents to expressly permit provision by "electronic or other means"
- Rule 1.16(e) - clarifies a lawyer's duty to release all client materials when terminating a representation to expressly include release of client materials created or held in "electronic or other form"
- Rule 4.4 - requires a lawyer who receives inadvertently produced materials that obviously appear to be subject to the attorney-client privilege or confidential and privileged to immediately notify the sender
- Rule 7.2(a) - clarifies the advertising rules to provide that a lawyer may advertise through "electronic means of communication, including public media"

- Rule 7.5, Comment - clarifies the scope of the rule governing a lawyer's professional designation to include logos and "URLs"

Los Angeles Times

POLITICS

Newsom wants to shape AI's future. Can California lead the way?

California Gov. Gavin Newsom on Wednesday announced a new executive order on artificial intelligence. (Rich Pedroncelli / Associated Press)

BY QUEENIE WONG | STAFF WRITER

SEPT. 6, 2023 3:21 PM PT

California Gov. Gavin Newsom on Wednesday signed an executive order directing state agencies to examine the benefits and risks of artificial intelligence that can generate text, images and other content.

The executive order sets the stage for potential regulation around what's known as generative AI technology, which has already raised concerns about misinformation, plagiarism, bias and child safety. The governor and California lawmakers thus far have been cautious about regulating technology they might not fully understand and hindering business innovations that fuel the state's economy.

"We recognize both the potential benefits and risks these tools enable. We're neither frozen by the fears nor hypnotized by the upside," Newsom said in a statement. "We're taking a clear-eyed, humble approach to this world-changing technology."

Tech companies including Microsoft, Google and Facebook parent company Meta have been scrambling to incorporate generative AI in their products as firms, such as OpenAI, release popular tools such as ChatGPT. AI has the potential to transform various industries, including state government and politics.

The executive order outlines steps state agencies can take to better understand AI. Under the order, the California Department of Technology, the Office of Data and Innovation, and other state agencies must examine the most significant and beneficial ways generative AI can be used by the state within 60 days. The governor also ordered the agencies to look at potential AI risks to individuals, communities and state workers. They have until January 2024 to issue guidelines for the procurement, uses and training required for use of generative AI.

Darrell West, senior fellow at the Center for Technology Innovation within the Governance Studies program at the Brookings Institution, said the executive order was comprehensive and could influence action taken by other states.

"California has long been a trendsetter, and the governor's executive order is continuing that tradition on AI," West said. "If one large state makes a major movement it's going to force the tech companies to come along, whether they want to or not."

Child safety groups such as the nonprofit Common Sense Media say that they see the executive order as a first step but that lawmakers will need to pass legislation to combat AI's risks, including child sexual abuse imagery. These safety issues will only be "exacerbated by AI," said Jim Steyer, chief executive of Common Sense Media.

"We have to put major protections and guardrails in place and you have to do that legislatively," he said.

The executive order comes before tech executives including OpenAI CEO Sam Altman, Meta CEO Mark Zuckerberg, Google CEO Sundar Pichai, Tesla CEO Elon Musk and others are expected to head to Washington next week for an AI forum hosted by Senate Majority Leader Charles E. Schumer (D-N.Y.).

Newsom has previously signaled that he's cautious about AI regulation. At a Los Angeles conference in May, he said that "the biggest mistake" politicians can make is asserting themselves "without first seeking to understand." Newsom also participated in an AI roundtable in June with President Biden, whose administration has met with tech executives and released a "Blueprint for an AI Bill of Rights."

Newsom's office said the governor wasn't available for an interview. He told Bloomberg he thought there was a "Pandora's box" being opened with generative AI and the state wants "it done in a safe way."

Peter Leroe-Muñoz, general counsel and senior vice president for technology and innovation at the Silicon Valley Leadership Group, said he was pleased that the governor's office has been engaging with tech companies and other interested parties. AI, he said, has the potential to make government work more efficient and effective and services more accessible to Californians.

"This executive order really shows that the governor is placing California in the driver's seat as we road map AI's future in America," he said.

Queenie Wong

Queenie Wong is a state politics reporter covering tech and entertainment policy for the Los Angeles Times. Previously, she wrote about social media companies for CNET and the Mercury News. She also covered politics and education for the Statesman Journal in Salem, Ore. Growing up in Southern California, she started reading The Times as a kid and took her first journalism class in middle school. She graduated from Washington and Lee University, where she studied journalism and studio art.

EXECUTIVE DEPARTMENT
STATE OF CALIFORNIA

EXECUTIVE ORDER N-12-23

WHEREAS the State of California is a global leader in innovation, research, development, human capital, and entrepreneurship; and

WHEREAS Generative Artificial Intelligence ("GenAI") represents a significant leap forward in technology, by generating novel text, images, and other content, which will transform the way that the State and the world conduct business and serve the public; and

WHEREAS GenAI has the potential to catalyze innovation and the rapid development of a wide range of benefits for Californians and the California economy, such as advances in medicine, wildfire forecasting and prevention, and climate science, and to push the bounds of human creativity and capacity; and

WHEREAS California is leading the world in GenAI innovation and research, and is home to 35 of the world's top 50 Artificial Intelligence ("AI") companies; and

WHEREAS San Francisco and San Jose are dominating this technological revolution, accounting for a quarter of all AI patents, conference papers, and companies globally; and

WHEREAS the State of California endeavors to continue leading the world in the responsible development, adoption, and implementation of new technologies for the benefit of all Californians and the California economy; and

WHEREAS GenAI can enhance human potential and creativity but must be deployed and regulated carefully to mitigate and guard against a new generation of risks; and

WHEREAS the State of California is committed to accuracy, reliability, and ethical outcomes when adopting GenAI technology, engaging and supporting historically vulnerable and marginalized communities, and serving its residents, workers, and businesses in a transparent, engaged, and equitable way; and

WHEREAS the State of California seeks to realize the potential benefits of GenAI for the good of all California residents, through the development and deployment of GenAI tools that improve the equitable and timely delivery of services, while balancing the benefits and risks of these new technologies; and

WHEREAS the California state workforce is vital to California's continued prosperity and the State seeks to harness the potential of GenAI for the benefit of the state government workforce; and

WHEREAS California is home to the University of California, Berkeley, College of Computing, Data Science, and Society and Stanford University's Institute for Human-Centered Artificial Intelligence, which are world-leading research institutions in GenAI; and

WHEREAS the location of these institutions in California provides a unique opportunity for academic research and government collaboration; and

WHEREAS GenAI development is advancing and accelerating at an exponential pace; and

WHEREAS the unprecedented speed of innovation and deployment of GenAI technologies necessitates measured guardrails to protect against potential risks or malicious uses, including but not limited to, bioterrorism, cyberattacks, disinformation, deception,[1] and discrimination or bias;[2] and

WHEREAS the promise of GenAI requires the creation of a risk-safety ecosystem and the investment of private companies and research institutions to foster talent dedicated to trust and safety in these new frontiers; and

WHEREAS research institutions face significant barriers in accessing the vast amounts of computing power necessary to use GenAI for research and education and there is a need for a public-private effort to overcome such barriers; and

WHEREAS thoughtful, responsive governance at the beginning of a technology's lifecycle can maximize equitable distribution of the benefits, minimize adverse impacts and abuse by bad actors, and reduce barriers to entry into emerging markets; and

WHEREAS the development of GenAI will necessitate united governance on issues of consumer data, financial services, healthcare, and innumerable other areas critical to our society, and my administration looks forward to engaging with the Legislature in furtherance of this aim; and

WHEREAS California is the most populous, diverse state in the nation and a global technology leader uniquely situated to lead the world in responsibly developing, implementing, and governing GenAI, by combining the strengths of California's world-class tech industry, universities, economy, and workforce.

NOW, THEREFORE, I, GAVIN NEWSOM, Governor of the State of California, in accordance with the authority vested in me by the State Constitution and statutes of the State of California, do hereby issue the following Order to become effective immediately:

IT IS HEREBY ORDERED THAT:

1) Within 60 days of issuance of this Order, the Government Operations Agency, the California Department of Technology, the Office of Data and Innovation, and the Governor's Office of Business and Economic Development, in collaboration with other State agencies and departments and their workforce, shall draft a report to the Governor examining the most significant, potentially beneficial use cases for deployment of GenAI tools by the State. The report shall also explain

---

[1] https://www.safe.ai/ai-risk#malicious-use.
[2] https://www.whitehouse.gov/ostp/ai-bill-of-rights/algorithmic-discrimination-protections-2/.

the potential risks to individuals, communities, and government and state government workers, with a focus on high-risk use cases, such as where GenAI is used to make a consequential decision affecting access to essential goods and services. Additionally, the report shall include but not be limited to: risks stemming from bad actors and insufficiently guarded governmental systems, unintended or emergent effects, and potential risks toward democratic and legal processes, public health and safety, and the economy. The report shall be regularly assessed for any significant developments or necessary updates and as appropriate, be done in consultation with civil society, academia, industry experts, and the state government workforce or organizations that represent state government employees.

2) No later than March 2024, the California Cybersecurity Integration Center and the California State Threat Assessment Center, both established within the Governor's Office of Emergency Services, and inclusive of the California Department of Technology, the California Military Department, and the California Highway Patrol, shall perform a joint risk analysis of potential threats to and vulnerabilities of California's critical energy infrastructure by the use of GenAI, including those which could lead to mass casualty events and environmental emergencies, and develop, in consultation with external experts as appropriate from civil society, academia, and industry, a strategy to assess similar potential threats to other critical infrastructure. Once this analysis is completed, these agencies shall provide a classified briefing to the Governor and, where appropriate and without divulging classified information, make public recommendations for further administrative actions and/or collaboration with the Legislature to guard against these potential threats and vulnerabilities. These recommendations shall address how to ensure systems are regularly tested and monitored to detect and avoid unintended behavior, and how to ensure they remain under effective human control. At a cadence deemed appropriate by the Governor's Office of Emergency Services, the analysis and public recommendations should be updated to reflect changes to the technology, its applications, and risk management processes and learnings.

3) To ensure State government fosters a safe and responsible innovation ecosystem that puts AI systems and tools to the best uses for Californians:

    a. By January 2024, the Government Operations Agency, the California Department of General Services, the California Department of Technology, and the California Cybersecurity Integration Center, shall issue general guidelines for public sector procurement, uses, and required trainings for use of GenAI, including for high-risk scenarios such as for consequential decisions affecting access to essential goods and services. The guidelines should build on guidance from the White House's Blueprint for an AI Bill of Rights and the National Institute for Science & Technology's AI Risk Management Framework, and shall address safety, algorithmic discrimination, data privacy,

and notice of when materials are generated by GenAI. The Government Operations Agency shall engage and consult with the state government workforce or organizations that represent state government employees, and industry experts, including but not limited to, trust and safety experts, academic researchers, and research institutions, in developing these guidelines. The Government Operations Agency, the California Department of General Services, the California Department of Technology, and the California Cybersecurity Integration Center shall thereafter evaluate at a cadence they deem appropriate any need to revise the guidelines and establish a consultative process by which to do so with civil society, academia, industry experts, and the state government workforce or organizations that represent state government employees. Nothing in this Order shall be construed to allow the application of state funds by localities in the procurement of GenAI technologies.

b. By July 2024, the Government Operations Agency, the California Department of Technology, and the Office of Data and Innovation, in consultation with other State agencies and departments, shall develop guidelines for State agencies and departments to analyze the impact that adopting a GenAI tool may have on vulnerable communities, including criteria to evaluate equitable outcomes in deployment and implementation of high-risk use cases. These guidelines and criteria shall inform whether and how a State agency or department deploys a particular GenAI tool. The Government Operations Agency shall engage and consult with the state government workforce or organizations that represent state government employees, and industry experts, including but not limited to, trust and safety experts, academic researchers, and research institutions, in developing these guidelines. The Government Operations Agency, the California Department of Technology, and the Office of Data and Innovation shall thereafter evaluate at a cadence they deem appropriate any need to revise the guidelines and criteria and establish a consultative process by which to do so with civil society, academia, industry experts, and the state government workforce or organizations that represent state government employees.

c. By January 2025, the Government Operations Agency, the California Department of General Services, and the California Department of Technology shall update the State's project approval, procurement, and contract terms, incorporating analysis and feedback obtained through the processes outlined in 3.a. and 3.b.

d. In order to assist the Government Operations Agency and the California Department of Technology in these efforts, all agencies and departments subject to my authority shall conduct and submit an inventory of all current high-risk uses of GenAI

within the agency or department to the California Department of Technology, which will administer the inventory. To effectuate this inventory, all agencies and departments shall appoint a senior level management personnel who will be responsible for conducting and reporting the results of the inventory to the California Department of Technology within 60 days of issuance of this Order. The senior management personnel shall be responsible for maintaining the inventory on an ongoing basis.

e. State agencies and departments subject to my authority shall consider procurement and enterprise use opportunities where GenAI can improve the efficiency, effectiveness, accessibility, and equity of government operations consistent with the Government Operations Agency, the California Department of General Services, and the California Department of Technology's guidelines for public sector GenAI procurement.

f. By March 2024, the California Department of Technology shall establish the infrastructure to conduct pilots of GenAI projects, including California Department of Technology approved environments, or "sandboxes," to test such projects. These environments will be available to State agencies and departments to help evaluate GenAI tools and services, to further safe, ethical, and responsible implementations, and to inform decisions to use GenAI, consistent with state guidelines.

g. By July 2024, all state agencies under my authority shall consider pilot projects of GenAI applications, in consultation with the state workforce or organizations that represent state government employees, and experts as appropriate from civil society, academia, and industry. Under a controlled setting, pilots shall measure 1) how GenAI can improve Californians' experience with and access to government services, and 2) how GenAI can support state employees in the performance of their duties in addition to any domain-specific impacts to be measured by the agency.

4) The Government Operations Agency, the California Department of Human Resources, the California Department of General Services, the California Department of Technology, the Office of Data and Innovation, and the California Cybersecurity Integration Center, shall engage with the Legislature and relevant stakeholders, including historically vulnerable and marginalized communities, and organizations that represent state government employees, in the development of any guidelines, criteria, reports, and/or training as directed by this Order.

5) State agencies and departments subject to my authority shall support California's state government workforce and prepare for the next generation of skills needed to thrive in the GenAI economy by:

a. No later than July 2024, the Government Operations Agency, the California Department of Technology, the California Department of Human Resources, and the Labor and Workforce Development Agency shall make available trainings for state government worker use of state-approved GenAI tools to achieve equitable outcomes, and to identify and mitigate potential output inaccuracies, fabricated text, hallucinations, and biases of GenAI, while enforcing public privacy and applicable state laws and policies. Where appropriate, the California Department of Technology and the California Department of Human Resources shall collaborate with the state government workforce or organizations that represent state government employees, and industry experts, on developing and providing training.

b. No later than January 1, 2025, the Government Operations Agency, the California Department of Human Resources, and the Labor and Workforce Development Agency, in consultation with the state government workforce or organizations that represent state government employees, shall establish criteria to evaluate the impact of GenAI to the state government workforce, and provide guidelines on how State agencies and departments can support state government employees to use these tools effectively and respond to these technological advancements.

6) The Governor's Office of Business and Economic Development, in consultation with the Government Operations Agency, is directed to pursue a formal partnership with the University of California, Berkeley, College of Computing, Data Science, and Society and Stanford University's Institute for Human-Centered Artificial Intelligence to consider and evaluate the impacts of GenAI on California and what efforts the State should undertake to advance its leadership in this industry. As part of this effort, beginning in the fall of 2023, those agencies are directed to work with the University of California, Berkeley, College of Computing, Data Science, and Society and Stanford University's Institute for Human-Centered Artificial Intelligence to develop and host a joint California-specific summit in 2024, to engage in meaningful discussions and thought partnership about the impacts of GenAI on California and its workforce and how all stakeholders can support growth in a manner that safeguards Californians.

7) Legal counsel for all State agencies, departments, and boards subject to my authority shall consider and periodically evaluate for any potential impact of GenAI on regulatory issues under the respective agency, department, or board's authority and recommend necessary updates, where appropriate, as a result of this evolving technology.

IT IS FURTHER ORDERED that, as soon as hereafter possible, this Order be filed in the Office of the Secretary of State and that widespread publicity and notice be given of this Order.

This Order is not intended to, and does not, create any rights or benefits, substantive or procedural, enforceable at law or in equity, against the State of California, its agencies, departments, entities, officers, employees, or any other person.

IN WITNESS WHEREOF I have hereunto set my hand and caused the Great Seal of the State of California to be affixed this 6th day of September 2023.

GAVIN NEWSOM
Governor of California

ATTEST:

SHIRLEY N. WEBER, Ph.D.
Secretary of State

SB-1047 Safe and Secure Innovation for Frontier Artificial Intelligence Systems Act. (2023-2024)

Date Published: 02/07/2024 09:00 PM

CALIFORNIA LEGISLATURE— 2023–2024 REGULAR SESSION

SENATE BILL NO. 1047

Introduced by Senator Wiener

February 07, 2024

An act to add Chapter 22.6 (commencing with Section 22602) to Division 8 of the Business and Professions Code, and to add Sections 11547.6 and 11547.7 to the Government Code, relating to artificial intelligence.

LEGISLATIVE COUNSEL'S DIGEST

SB 1047, as introduced, Wiener. Safe and Secure Innovation for Frontier Artificial Intelligence Systems Act.

Existing law requires the Secretary of Government Operations to develop a coordinated plan to, among other things, investigate the feasibility of, and obstacles to, developing standards and technologies for state departments to determine digital content provenance. For the purpose of informing that coordinated plan, existing law requires the secretary to evaluate, among other things, the impact of the proliferation of deepfakes, defined to mean audio or visual content that has been generated or manipulated by artificial intelligence that would falsely appear to be authentic or truthful and that features depictions of people appearing to say or do things they did not say or do without their consent, on state government, California-based businesses, and residents of the state.

Existing law creates the Department of Technology within the Government Operations Agency and requires the department to, among other things, identify, assess, and prioritize high-risk, critical information technology services and systems across state government for modernization, stabilization, or remediation.

This bill would enact the Safe and Secure Innovation for Frontier Artificial Intelligence Systems Act to, among other things, require a developer of a covered model, as defined, to determine whether it can make a positive safety determination with respect to a covered model before initiating training of that covered model, as specified. The bill would define "positive safety determination" to mean a determination with respect to a covered model, that is not a derivative model, that a developer can reasonably exclude the possibility that the covered model has a hazardous capability, as defined, or may come close to possessing a hazardous capability when accounting for a reasonable margin for safety and the possibility of posttraining modifications.

This bill would require that a developer, before initiating training of a nonderivative covered model, comply with various requirements, including implementing the capability to promptly enact a full shutdown of the covered model until that covered model is the subject of a positive safety determination.

This bill would require a developer of a nonderivative covered model that is not the subject of a positive safety determination to submit to the Frontier Model Division, which the bill would create within the Department of Technology, an annual certification of compliance with these provisions signed by the chief technology officer, or a more senior corporate officer, in a format and on a date as prescribed by the Frontier Model Division. By expanding the scope of the crime of perjury, this bill would impose a state-mandated local program. The bill would also require a developer to report each artificial intelligence safety incident affecting a covered model to the Frontier Model Division in a manner prescribed by the Frontier Model Division.

This bill would require a person that operates a computing cluster, as defined, to implement appropriate written policies and procedures to do certain things when a customer utilizes compute resources that would be sufficient to train a covered model, including assess whether a prospective customer intends to utilize the computing cluster to deploy a covered model. The bill would punish a violation of these provisions with a civil penalty, as prescribed, to be recovered by the Attorney General.

This bill would also create the Frontier Model Division within the Department of Technology and would require the division to, among other things, review annual certification reports from developers received pursuant to these provisions and publicly release summarized findings based on those reports. The bill would authorize the division to assess related fees and would require deposit of the fees into the Frontier Model Division Programs Fund, which the bill would create. The bill would make moneys in the fund available for the purpose of these provisions only upon appropriation by the Legislature.

This bill would also require the Department of Technology to commission consultants, as prescribed, to create a public cloud computing cluster, to be known as CalCompute, with the primary focus of conducting research into the safe and secure deployment of large-scale artificial intelligence models and fostering equitable innovation that includes, among other things, a fully owned and hosted cloud platform.

The California Constitution requires the state to reimburse local agencies and school districts for certain costs mandated by the state. Statutory provisions establish procedures for making that reimbursement.

This bill would provide that no reimbursement is required by this act for a specified reason.
Vote: majority Appropriation: no Fiscal Committee: yes Local Program: yes

THE PEOPLE OF THE STATE OF CALIFORNIA DO ENACT AS FOLLOWS:

SECTION 1. This act shall be known, and may be cited, as the Safe and Secure Innovation for Frontier Artificial Intelligence Systems Act.

SEC. 2. The Legislature finds and declares all of the following:

(a) California is leading the world in artificial intelligence innovation and research, through companies large and small, as well as through our remarkable public and private universities.

(b) Artificial intelligence, including new advances in generative artificial intelligence, has the potential to catalyze innovation and the rapid development of a wide range of benefits for Californians and the California economy, including advances in medicine, wildfire forecasting and prevention, and climate science, and to push the bounds of human creativity and capacity.

(c) If not properly subject to human controls, future development in artificial intelligence may also have the potential to be used to create novel threats to public safety and security, including by enabling the creation and the proliferation of weapons of mass destruction, such as biological, chemical, and nuclear weapons, as well as weapons with cyber-offensive capabilities.

(d) The state government has an essential role to play in ensuring that California recognizes the benefits of this technology while avoiding the most severe risks, as well as to ensure that artificial intelligence innovation and access to compute is accessible to academic researchers and startups, in addition to large companies.

SEC. 3. Chapter 22.6 (commencing with Section 22602) is added to Division 8 of the Business and Professions Code, to read:

CHAPTER 22.6. Safe and Secure Innovation for Frontier Artificial Intelligence Systems

22602. As used in this chapter:

(a) "Advanced persistent threat" means an adversary with sophisticated levels of expertise and significant resources that allow it, through the use of multiple different attack vectors, including, but not limited to, cyber,

physical, and deception, to generate opportunities to achieve its objectives that are typically to establish and extend its presence within the information technology infrastructure of organizations for purposes of exfiltrating information or to undermine or impede critical aspects of a mission, program, or organization or place itself in a position to do so in the future.

(b) "Artificial intelligence model" means a machine-based system that can make predictions, recommendations, or decisions influencing real or virtual environments and can use model inference to formulate options for information or action.

(c) "Artificial intelligence safety incident" means any of the following:

(1) A covered model autonomously engaging in a sustained sequence of unsafe behavior other than at the request of a user.

(2) Theft, misappropriation, malicious use, inadvertent release, unauthorized access, or escape of the model weights of a covered model.

(3) The critical failure of technical or administrative controls, including controls limiting the ability to modify a covered model, designed to limit access to a hazardous capability of a covered model.

(4) Unauthorized use of the hazardous capability of a covered model.

(d) "Computing cluster" means a set of machines transitively connected by data center networking of over 100 gigabits that has a theoretical maximum computing capacity of 10^20 integer or floating-point operations per second for training artificial intelligence.

(e) "Covered guidance" means any of the following:

(1) Applicable guidance issued by the National Institute of Standards and Technology and by the Frontier Model Division.

(2) Industry best practices, including relevant safety practices, precautions, or testing procedures undertaken by developers of comparable models, and any safety standards or best practices commonly or generally recognized by relevant experts in academia or the nonprofit sector.

(3) Applicable safety-enhancing standards set by standards setting organizations.

(f) "Covered model" means an artificial intelligence model that meets either of the following criteria:

(1) The artificial intelligence model was trained using a quantity of computing power greater than 10^26 integer or floating-point operations in 2024, or a model that could reasonably be expected to have similar performance on benchmarks commonly used to quantify the performance of state-of-the-art foundation models, as determined by industry best practices and relevant standard setting organizations.

(2) The artificial intelligence model has capability below the relevant threshold on a specific benchmark but is of otherwise similar general capability.

(g) "Critical harm" means a harm listed in paragraph (1) of subdivision (n).

(h) "Critical infrastructure" means assets, systems, and networks, whether physical or virtual, the incapacitation or destruction of which would have a debilitating effect on physical security, economic security, public health, or safety in the state.

(i) (1) "Derivative model" means an artificial intelligence model that is a derivative of another artificial intelligence model, including either of the following:

(A) A modified or unmodified copy of an artificial intelligence model.

(B) A combination of an artificial intelligence model with other software.

(2) "Derivative model" does not include an entirely independently trained artificial intelligence model.

(j) (1) "Developer" means a person that creates, owns, or otherwise has responsibility for an artificial intelligence model.

(2) "Developer" does not include a third-party machine-learning operations platform, an artificial intelligence infrastructure platform, a computing cluster, an application developer using sourced models, or an end-user of an artificial intelligence model.

(k) "Fine tuning" means the adjustment of the model weights of an artificial intelligence model that has been previously trained by training the model with new data.

(l) "Frontier Model Division" means the Frontier Model Division created pursuant to Section 11547.6 of the Government Code.

(m) "Full shutdown" means the cessation of operation of a covered model, including all copies and derivative models, on all computers and storage devices within custody, control, or possession of a person, including any computer or storage device remotely provided by agreement.

(n) (1) "Hazardous capability" means the capability of a covered model to be used to enable any of the following harms in a way that would be significantly more difficult to cause without access to a covered model:

(A) The creation or use of a chemical, biological, radiological, or nuclear weapon in a manner that results in mass casualties.

(B) At least five hundred million dollars ($500,000,000) of damage through cyberattacks on critical infrastructure via a single incident or multiple related incidents.

(C) At least five hundred million dollars ($500,000,000) of damage by an artificial intelligence model that autonomously engages in conduct that would violate the Penal Code if undertaken by a human.

(D) Other threats to public safety and security that are of comparable severity to the harms described in paragraphs (A) to (C), inclusive.

(2) "Hazardous capability" includes a capability described in paragraph (1) even if the hazardous capability would not manifest but for fine tuning and posttraining modifications performed by third-party experts intending to demonstrate those abilities.

(o) "Machine-learning operations platform" means a solution that includes a combined offering of necessary machine-learning development capabilities, including exploratory data analysis, data preparation, model training and tuning, model review and governance, model inference and serving, model deployment and monitoring, and automated model retraining.

(p) "Model weight" means a numerical parameter established through training in an artificial intelligence model that helps determine how input information impacts a model's output.

(q) "Open-source artificial intelligence model" means an artificial intelligence model that is made freely available and may be freely modified and redistributed.

(r) "Person" means an individual, proprietorship, firm, partnership, joint venture, syndicate, business trust, company, corporation, limited liability company, association, committee, or any other nongovernmental organization or group of persons acting in concert.

(s) "Positive safety determination" means a determination, pursuant to subdivision (a) or (c) of Section 22603, with respect to a covered model that is not a derivative model that a developer can reasonably exclude the possibility that a covered model has a hazardous capability or may come close to possessing a hazardous capability when accounting for a reasonable margin for safety and the possibility of posttraining modifications.

(t) "Posttraining modification" means the modification of the capabilities of an artificial intelligence model after the completion of training by any means, including, but not limited to, initiating additional training, providing the model with access to tools or data, removing safeguards against hazardous misuse or misbehavior of the model, or combining the model with, or integrating it into, other software.

(u) "Safety and security protocol" means documented technical and organizational protocols that meet both of the following criteria:

(1) The protocols are used to manage the risks of developing and operating covered models across their life cycle, including risks posed by enabling or potentially enabling the creation of derivative models.

(2) The protocols specify that compliance with the protocols is required in order to train, operate, possess, and provide external access to the developer's covered model.

22603. (a) Before initiating training of a covered model that is not a derivative model, a developer of that covered model shall determine whether it can make a positive safety determination with respect to the covered model.

(1) In making the determination required by this subdivision, a developer shall incorporate all covered guidance.

(2) A developer may make a positive safety determination if the covered model will have lower performance on all benchmarks relevant under subdivision (f) of Section 22602 than either of the following:

(A) A non-covered model that manifestly lacks hazardous capabilities.

(B) Another model that is the subject of a positive safety determination.

(3) Upon making a positive safety determination, the developer of the covered model shall submit to the Frontier Model Division a certification under penalty of perjury that specifies the basis for that conclusion.

(b) Before initiating training of a covered model that is not a derivative model that is not the subject of a positive safety determination, and until that covered model is the subject of a positive safety determination, the developer of that covered model shall do all of the following:

(1) Implement administrative, technical, and physical cybersecurity protections to prevent unauthorized access to, or misuse or unsafe modification of, the covered model, including to prevent theft, misappropriation, malicious use, or inadvertent release or escape of the model weights from the developer's custody, that are appropriate in light of the risks associated with the covered model, including from advanced persistent threats or other sophisticated actors.

(2) Implement the capability to promptly enact a full shutdown of the covered model.

(3) Implement all covered guidance.

(4) Implement a written and separate safety and security protocol that does all of the following:

(A) Provides reasonable assurance that if a developer complies with its safety and security protocol, either of the following will apply:

(i) The developer will not produce a covered model with a hazardous capability or enable the production of a derivative model with a hazardous capability.

(ii) The safeguards enumerated in the policy will be sufficient to prevent critical harms from the exercise of a hazardous capability in a covered model.

(B) States compliance requirements in an objective manner and with sufficient detail and specificity to allow the developer or a third party to readily ascertain whether the requirements of the safety and security protocol have been followed.

(C) Identifies specific tests and test results that would be sufficient to reasonably exclude the possibility that a covered model has a hazardous capability or may come close to possessing a hazardous capability when accounting for a reasonable margin for safety and the possibility of posttraining modifications, and in addition does all of the following:

(i) Describes in detail how the testing procedure incorporates fine tuning and posttraining modifications performed by third-party experts intending to demonstrate those abilities.

(ii) Describes in detail how the testing procedure incorporates the possibility of posttraining modifications.

(iii) Describes in detail how the testing procedure incorporates the requirement for reasonable margin for safety.

(iv) Provides sufficient detail for third parties to replicate the testing procedure.

(D) Describes in detail how the developer will meet requirements listed under paragraphs (1), (2), (3), and (5).

(E) If applicable, describes in detail how the developer intends to implement the safeguards and requirements referenced in paragraph (1) of subdivision (d).

(F) Describes in detail the conditions that would require the execution of a full shutdown.

(G) Describes in detail the procedure by which the safety and security protocol may be modified.

(H) Meets other criteria stated by the Frontier Model Division in guidance to achieve the purpose of maintaining the safety of a covered model with a hazardous capability.

(5) Ensure that the safety and security protocol is implemented as written, including, at a minimum, by designating senior personnel responsible for ensuring implementation by employees and contractors working on a covered model, monitoring and reporting on implementation, and conducting audits, including through third parties as appropriate.

(6) Provide a copy of the safety and security protocol to the Frontier Model Division.

(7) Conduct an annual review of the safety and security protocol to account for any changes to the capabilities of the covered model and industry best practices and, if necessary, make modifications to the policy.

(8) If the safety and security protocol is modified, provide an updated copy to the Frontier Model Division within 10 business days.

(9) Refrain from initiating training of a covered model if there remains an unreasonable risk that an individual, or the covered model itself, may be able to use the hazardous capabilities of the covered model, or a derivative model based on it, to cause a critical harm.

(c) (1) Upon completion of the training of a covered model that is not the subject of a positive safety determination and is not a derivative model, the developer shall perform capability testing sufficient to determine whether the developer can make a positive safety determination with respect to the covered model pursuant to its safety and security protocol.

(2) Upon making a positive safety determination with respect to the covered model, a developer of the covered model shall submit to the Frontier Model Division a certification of compliance with the requirements of this section within 90 days and no more than 30 days after initiating the commercial, public, or widespread use of the covered model that includes both of the following:

(A) The basis for the developer’s positive safety determination.

(B) The specific methodology and results of the capability testing undertaken pursuant to this subdivision.

(d) Before initiating the commercial, public, or widespread use of a covered model that is not subject to a positive safety determination, a developer of the nonderivative version of the covered model shall do all of the following:

(1) Implement reasonable safeguards and requirements to do all of the following:

(A) Prevent an individual from being able to use the hazardous capabilities of the model, or a derivative model, to cause a critical harm.

(B) Prevent an individual from being able to use the model to create a derivative model that was used to cause a critical harm.

(C) Ensure, to the extent reasonably possible, that the covered model’s actions and any resulting critical harms can be accurately and reliably attributed to it and any user responsible for those actions.

(2) Provide reasonable requirements to developers of derivative models to prevent an individual from being able to use a derivative model to cause a critical harm.

(3) Refrain from initiating the commercial, public, or widespread use of a covered model if there remains an unreasonable risk that an individual may be able to use the hazardous capabilities of the model, or a derivative model based on it, to cause a critical harm.

(e) A developer of a covered model shall periodically reevaluate the procedures, policies, protections, capabilities, and safeguards implemented pursuant to this section in light of the growing capabilities of covered models and as is reasonably necessary to ensure that the covered model or its users cannot remove or bypass those procedures, policies, protections, capabilities, and safeguards.

(f) (1) A developer of a nonderivative covered model that is not the subject of a positive safety determination shall submit to the Frontier Model Division an annual certification of compliance with the requirements of this section signed by the chief technology officer, or a more senior corporate officer, in a format and on a date as prescribed by the Frontier Model Division.

(2) In a certification submitted pursuant to paragraph (1), a developer shall specify or provide, at a minimum, all of the following:

(A) The nature and magnitude of hazardous capabilities that the covered model possesses or may reasonably possess and the outcome of capability testing required by subdivision (c).

(B) An assessment of the risk that compliance with the safety and security protocol may be insufficient to prevent harms from the exercise of the covered model's hazardous capabilities.

(C) Other information useful to accomplishing the purposes of this subdivision, as determined by the Frontier Model Division.

(g) A developer shall report each artificial intelligence safety incident affecting a covered model to the Frontier Model Division in a manner prescribed by the Frontier Model Division. The notification shall be made in the most expedient time possible and without unreasonable delay and in no event later than 72 hours after learning that an artificial intelligence safety incident has occurred or learning facts sufficient to establish a reasonable belief that an artificial intelligence safety incident has occurred.

(h) (1) Reliance on an unreasonable positive safety determination does not relieve a developer of its obligations under this section.

(2) A positive safety determination is unreasonable if the developer does not take into account reasonably foreseeable risks of harm or weaknesses in capability testing that lead to an inaccurate determination.

(3) A risk of harm or weakness in capability testing is reasonably foreseeable, if, by the time that a developer releases a model, an applicable risk of harm or weakness in capability testing has already been identified by either of the following:

(A) Any other developer of a comparable or comparably powerful model through risk assessment, capability testing, or other means.

(B) By the United States Artificial Intelligence Safety Institute, the Frontier Model Division, or any independent standard-setting organization or capability-testing organization cited by either of those entities.

22604. A person that operates a computing cluster shall implement appropriate written policies and procedures to do all of the following when a customer utilizes compute resources that would be sufficient to train a covered model:

(a) Obtain a prospective customer's basic identifying information and business purpose for utilizing the computing cluster, including all of the following:

(1) The identity of that prospective customer.

(2) The means and source of payment, including any associated financial institution, credit card number, account number, customer identifier, transaction identifiers, or virtual currency wallet or wallet address identifier.

(3) The email address and telephonic contact information used to verify a prospective customer's identity.

(4) The Internet Protocol addresses used for access or administration and the date and time of each access or administrative action.

(b) Assess whether a prospective customer intends to utilize the computing cluster to deploy a covered model.

(c) Annually validate the information collected pursuant to subdivision (a) and conduct the assessment required pursuant to subdivision (b).

(d) Maintain for seven years and provide to the Frontier Model Division or the Attorney General, upon request, appropriate records of actions taken under this section, including policies and procedures put into effect.

(e) Implement the capability to promptly enact a full shutdown in the event of an emergency.

22605. (a) A developer of a covered model that provides commercial access to that covered model shall provide a transparent, uniform, publicly available price schedule for the purchase of access to that covered model at a given level of quality and quantity subject to the developer's terms of service and shall not engage in unlawful discrimination or noncompetitive activity in determining price or access.

(b) A person that operates a computing cluster shall provide a transparent, uniform, publicly available price schedule for the purchase of access to the computing cluster at a given level of quality and quantity subject to the developer's terms of service and shall not engage in unlawful discrimination or noncompetitive activity in determining price or access.

22606. (a) If the Attorney General has reasonable cause to believe that a person is violating this chapter, the Attorney General shall commence a civil action in a court of competent jurisdiction.

(b) In a civil action under this section, the court may award any of the following:

(1) (A) Preventive relief, including a permanent or temporary injunction, restraining order, or other order against the person responsible for a violation of this chapter, including deletion of the covered model and the weights utilized in that model.

(B) Relief pursuant to this paragraph shall be granted only in response to harm or an imminent risk or threat to public safety.

(2) Other relief as the court deems appropriate, including monetary damages to persons aggrieved and an order for the full shutdown of a covered model.

(3) A civil penalty in an amount not exceeding 10 percent of the cost, excluding labor cost, to develop the covered model for a first violation and in an amount not exceeding 30 percent of the cost, excluding labor cost, to develop the covered model for any subsequent violation.

(c) In the apportionment of penalties assessed pursuant to this section, defendants shall be jointly and severally liable.

(d) A court shall disregard corporate formalities and impose joint and several liability on affiliated entities for purposes of effectuating the intent of this section if the court concludes that both of the following are true:

(1) Steps were taken in the development of the corporate structure among affiliated entities to purposely and unreasonably limit or avoid liability.

(2) The corporate structure of the developer or affiliated entities would frustrate recovery of penalties or injunctive relief under this section.

22607. (a) Pursuant to subdivision (a) of Section 1102.5 of the Labor Code, a developer shall not prevent an employee from disclosing information to the Attorney General if the employee has reasonable cause to believe that the information indicates that the developer is out of compliance with the requirements of Section 22603.

(b) Pursuant to subdivision (b) of Section 1102.5 of the Labor Code, a developer shall not retaliate against an employee for disclosing information to the Attorney General if the employee has reasonable cause to believe that the information indicates that the developer is out of compliance with the requirements of Section 22603.

(c) The Attorney General may publicly release any complaint, or a summary of that complaint, pursuant to this section if the Attorney General concludes that doing so will serve the public interest.

(d) Employees shall seek relief for violations of this section pursuant to Sections 1102.61 and 1102.62 of the Labor Code.

(e) Pursuant to subdivision (a) of Section 1102.8 of the Labor Code, a developer shall provide clear notice to all employees working on covered models of their rights and responsibilities under this section.

SEC. 4. Section 11547.6 is added to the Government Code, to read:

11547.6. (a) As used in this section:

(1) "Hazardous capability" has the same meaning as defined in Section 22602 of the Business and Professions Code.

(2) "Positive safety determination" has the same meaning as defined in Section 22602 of the Business and Professions Code.

(b) The Frontier Model Division is hereby created within the Department of Technology.

(c) The Frontier Model Division shall do all of the following:

(1) Review annual certification reports received from developers pursuant to Section 22603 of the Business and Professions Code and publicly release summarized findings based on those reports.

(2) Advise the Attorney General on potential violations of this section or Chapter 22.6 (commencing with Section 22602) of Division 8 of the Business and Professions Code.

(3) (A) Issue guidance, standards, and best practices sufficient to prevent unreasonable risks from covered models with hazardous capabilities including, but not limited to, more specific requirements on the duties required under Section 22603 of the Business and Professions Code.

(B) Establish an accreditation process and relevant accreditation standards under which third parties may be accredited for a three-year period, which may be extended through an appropriate process, to certify adherence by developers to the best practices and standards adopted pursuant to subparagraph (A).

(4) Publish anonymized artificial intelligence safety incident reports received from developers pursuant to Section 22603 of the Business and Professions Code.

(5) Establish confidential fora that are structured and facilitated in a manner that allows developers to share best risk management practices for models with hazardous capabilities in a manner consistent with state and federal antitrust laws.

(6) (A) Issue guidance describing the categories of artificial intelligence safety events that are likely to constitute a state of emergency within the meaning of subdivision (b) of Section 8558 and responsive actions that could be ordered by the Governor after a duly proclaimed state of emergency.

(B) The guidance issued pursuant to subparagraph (A) shall not limit, modify, or restrict the authority of the Governor in any way.

(7) Appoint and consult with an advisory committee that shall advise the Governor on when it may be necessary to proclaim a state of emergency relating to artificial intelligence and advise the Governor on what responses may be appropriate in that event.

(8) Appoint and consult with an advisory committee for open-source artificial intelligence that shall do all of the following:

(A) Issue guidelines for model evaluation for use by developers of open-source artificial intelligence models that do not have hazardous capabilities.

(B) Advise the Frontier Model Division on the creation and feasibility of incentives, including tax credits, that could be provided to developers of open-source artificial intelligence models that are not covered models.

(C) Advise the Frontier Model Division on future policies and legislation impacting open-source artificial intelligence development.

(9) Provide technical assistance and advice to the Legislature, upon request, with respect to artificial intelligence-related legislation.

(10) Monitor relevant developments relating to the safety risks associated with the development of artificial intelligence models and the functioning of markets for artificial intelligence models.

(11) Levy fees, including an assessed fee for the submission of a certification, in an amount sufficient to cover the reasonable costs of administering this section that do not exceed the reasonable costs of administering this section.

(12) (A) Develop and submit to the Judicial Council proposed model jury instructions for actions brought by individuals injured by a hazardous capability of a covered model.

(B) In developing the model jury instructions required by subparagraph (A), the Frontier Model Division shall consider all of the following factors:

(i) The level of rigor and detail of the safety and security protocol that the developer faithfully implemented while it trained, stored, and released a covered model.

(ii) Whether and to what extent the developer’s safety and security protocol was inferior, comparable, or superior, in its level of rigor and detail, to the mandatory safety policies of comparable developers.

(iii) The extent and quality of the developer’s safety and security protocol’s prescribed safeguards, capability testing, and other precautionary measures with respect to the relevant hazardous capability and related hazardous capabilities.

(iv) Whether and to what extent the developer and its agents complied with the developer’s safety and security protocol, and to the full degree, that doing so might plausibly have avoided causing a particular harm.

(v) Whether and to what extent the developer carefully and rigorously investigated, documented, and accurately measured, insofar as reasonably possible given the state of the art, relevant risks that its model might pose.

(d) There is hereby created in the General Fund the Frontier Model Division Programs Fund.

(1) All fees received by the Frontier Model Division pursuant to this section shall be deposited into the fund.

(2) All moneys in the account shall be available, only upon appropriation by the Legislature, for purposes of carrying out the provisions of this section.

SEC. 5. Section 11547.7 is added to the Government Code, to read:

11547.7. (a) The Department of Technology shall commission consultants, pursuant to subdivision (b), to create a public cloud computing cluster, to be known as CalCompute, with the primary focus of conducting research into the safe and secure deployment of large-scale artificial intelligence models and fostering equitable innovation that includes, but is not limited to, all of the following:

(1) A fully owned and hosted cloud platform.

(2) Necessary human expertise to operate and maintain the platform.

(3) Necessary human expertise to support, train, and facilitate use of CalCompute.

(b) The consultants shall include, but not be limited to, representatives of national laboratories, public universities, and any relevant professional associations or private sector stakeholders.

(c) To meet the objective of establishing CalCompute, the Department of Technology shall require consultants commissioned to work on this process to evaluate and incorporate all of the following considerations into its plan:

(1) An analysis of the public, private, and nonprofit cloud platform infrastructure ecosystem, including, but not limited to, dominant cloud providers, the relative compute power of each provider, the estimated cost of supporting platforms as well as pricing models, and recommendations on the scope of CalCompute.

(2) The process to establish affiliate and other partnership relationships to establish and maintain an advanced computing infrastructure.

(3) A framework to determine the parameters for use of CalCompute, including, but not limited to, a process for deciding which projects will be supported by CalCompute and what resources and services will be provided to projects.

(4) A process for evaluating appropriate uses of the public cloud resources and their potential downstream impact, including mitigating downstream harms in deployment.

(5) An evaluation of the landscape of existing computing capability, resources, data, and human expertise in California for the purposes of responding quickly to a security, health, or natural disaster emergency.

(6) An analysis of the state's investment in the training and development of the technology workforce, including through degree programs at the University of California, the California State University, and the California Community Colleges.

(7) A process for evaluating the potential impact of CalCompute on retaining technology professionals in the public workforce.

(d) The Department of Technology shall submit, pursuant to Section 9795, an annual report to the Legislature from the commissioned consultants to ensure progress in meeting the objectives listed above.

(e) The Department of Technology may receive private donations, grants, and local funds, in addition to allocated funding in the annual budget, to effectuate this section.

(f) This section shall become operative only upon an appropriation in a budget act for the purposes of this section.

SEC. 6. The provisions of this act are severable. If any provision of this act or its application is held invalid, that invalidity shall not affect other provisions or applications that can be given effect without the invalid provision or application.

SEC. 7. This act shall be liberally construed to effectuate its purposes.

SEC. 8. The duties and obligations imposed by this act are cumulative with any other duties or obligations imposed under other law and shall not be construed to relieve any party from any duties or obligations imposed under other law.

SEC. 9. No reimbursement is required by this act pursuant to Section 6 of Article XIII B of the California Constitution because the only costs that may be incurred by a local agency or school district will be incurred because this act creates a new crime or infraction, eliminates a crime or infraction, or changes the penalty for a crime or infraction, within the meaning of Section 17556 of the Government Code, or changes the definition of a crime within the meaning of Section 6 of Article XIII B of the California Constitution.

The New York Times https://www.nytimes.com/2023/12/29/nyregion/michael-cohen-ai-fake-cases.html

Michael Cohen Used Artificial Intelligence in Feeding Lawyer Bogus Cases

Donald Trump's former fixer had sought an early end to court supervision after his 2018 campaign finance conviction. He enlisted the help of Google Bard.

By Benjamin Weiser and Jonah E. Bromwich
Dec. 29, 2023

Michael D. Cohen, the onetime fixer for former President Donald J. Trump, mistakenly gave his lawyer bogus legal citations concocted by the artificial intelligence program Google Bard, he said in court papers unsealed on Friday.

The fictitious citations were used by the lawyer in a motion submitted to a federal judge, Jesse M. Furman. Mr. Cohen, who pleaded guilty in 2018 to campaign finance violations and served time in prison, had asked the judge for an early end to the court's supervision of his case now that he is out of prison and has complied with the conditions of his release.

The ensuing chain of misunderstandings and mistakes ended with Mr. Cohen asking the judge to exercise "discretion and mercy."

In a sworn declaration made public on Friday, Mr. Cohen explained that he had not kept up with "emerging trends (and related risks) in legal technology and did not realize that Google Bard was a generative text service that, like ChatGPT, could show citations and descriptions that looked real but actually were not."

He also said he had not realized that the lawyer filing the motion on his behalf, David M. Schwartz, "would drop the cases into his submission wholesale without even confirming that they existed."

The episode could have implications for a Manhattan criminal case against Mr. Trump in which Mr. Cohen is expected to be the star witness. The former president's lawyers have long attacked Mr. Cohen as a serial fabulist; now, they say they have a brand-new example.

The ill-starred filing was at least the second this year by lawyers in Manhattan federal court in which lawyers cited bogus decisions generated by artificial intelligence. The legal profession, like others, is struggling to account for a novel technology meant to mimic the human brain.

Artificial intelligence programs like Bard and ChatGPT generate realistic responses by hazarding guesses about which fragments of text should follow other sequences. Such programs draw on billions of examples of text ingested from across the internet. Although they can synthesize vast amounts of information and present it persuasively, there are still bugs to be worked out.

The three citations in Mr. Cohen's case appear to be hallucinations created by the Bard chatbot, taking bits and pieces of actual cases and combining them with robotic imagination. Mr. Schwartz then wove them into the motion he submitted to Judge Furman.

Mr. Cohen, in his declaration, said he understood Bard to be "a supercharged search engine," which he had used previously to find accurate information online.

it sent to your inbox.

Mr. Schwartz, in his own declaration, acknowledged using the citations and said he had not independently reviewed the cases because Mr. Cohen indicated that another lawyer, E. Danya Perry, was providing suggestions for the motion.

“I sincerely apologize to the court for not checking these cases personally before submitting them to the court,” Mr. Schwartz wrote.

Barry Kamins, a lawyer for Mr. Schwartz, declined to comment on Friday.

Ms. Perry has said she began representing Mr. Cohen only after Mr. Schwartz filed the motion. She wrote to Judge Furman on Dec. 8 that after reading the already-filed document, she could not verify the case law being cited.

In a statement at the time, she said that “consistent with my ethical obligation of candor to the court, I advised Judge Furman of this issue.”

She said in a letter made public on Friday that Mr. Cohen, a former lawyer who has been disbarred, “did not know that the cases he identified were not real and, unlike his attorney, had no obligation to confirm as much.”

“It must be emphasized that Mr. Cohen did not engage in any misconduct,” Ms. Perry wrote. She said on Friday that Mr. Cohen had no comment, and that he had consented to the unsealing of the court papers after the judge raised the question of whether they contained information protected by the attorney-client privilege.

The imbroglio surfaced when Judge Furman said in an order on Dec. 12 that he could not find any of the three decisions. He ordered Mr. Schwartz to provide copies or “a thorough explanation of how the motion came to cite cases that do not exist and what role, if any, Mr. Cohen played.”

The matter could have significant implications, given Mr. Cohen’s pivotal role in a case brought by the Manhattan district attorney that is scheduled for trial on March 25.

The district attorney, Alvin L. Bragg, charged Mr. Trump with orchestrating a hush money scheme that centered on a payment Mr. Cohen made during the 2016 election to a pornographic film star, Stormy Daniels. Mr. Trump has pleaded not guilty to 34 felony charges.

Seeking to rebut Mr. Trump’s lawyers’ claims that Mr. Cohen is untrustworthy, his defenders have said that Mr. Cohen lied on Mr. Trump’s behalf but has told the truth since splitting with the former president in 2018 and pleading guilty to the federal charges.

On Friday, Mr. Trump’s lawyers immediately seized on the Google Bard revelation. Susan R. Necheles, a lawyer representing Mr. Trump in the coming Manhattan trial, said it was “typical Michael Cohen.”

“He’s an admitted perjurer and has pled guilty to multiple felonies and this is just an additional indication of his lack of character and ongoing criminality,” Ms. Necheles said.

Ms. Perry, the lawyer now representing Mr. Cohen on the motion, said that Mr. Cohen’s willingness to have the filings unsealed showed he had nothing to hide.

“He relied on his lawyer, as he had every right to do,” she said. “Unfortunately, his lawyer appears to have made an honest mistake in not verifying the citations in the brief he drafted and filed.”

A spokeswoman for Mr. Bragg declined to comment on Friday.

Prosecutors may argue that Mr. Cohen’s actions were not intended to defraud the court, but rather, by his own admission, were a product of a woeful misunderstanding of new technology.

The issue of lawyers relying on chatbots exploded into public view earlier this year after another federal judge in Manhattan, P. Kevin Castel, fined two lawyers $5,000 after they admitted filing a legal brief filled with nonexistent cases and citations, all generated by ChatGPT.

Such cases appear to be rippling through the nation's courts, said Eugene Volokh, a law professor at U.C.L.A. who has written about artificial intelligence and the law.

Professor Volokh said he had counted a dozen cases in which lawyers or litigants representing themselves were believed to have used chatbots for legal research that ended up in court filings. "I strongly suspect that this is just the tip of the iceberg," he said.

Stephen Gillers, a legal ethics professor at New York University School of Law, said: "People should understand that generative A.I. is not the bad guy here. It holds much promise."

"But lawyers cannot treat A.I. as their co-counsel and just parrot what it says," he added.

The nonexistent cases cited in Mr. Schwartz's motion — United States v. Figueroa-Flores, United States v. Ortiz and United States v. Amato — came with corresponding summaries and notations that they had been affirmed by the U.S. Court of Appeals for the Second Circuit.

Judge Furman noted in his Dec. 12 order that the Figueroa-Flores citation actually referred to a page from a decision that was issued by a different federal appeals court and "has nothing to do with supervised release."

The Amato case named in the motion, the judge said, actually concerned a decision of the Board of Veterans' Appeals, an administrative tribunal.

And the citation to the Ortiz case, Judge Furman wrote, appeared "to correspond to nothing at all."

William K. Rashbaum contributed reporting.

Benjamin Weiser is a reporter covering the Manhattan federal courts. He has long covered criminal justice, both as a beat and investigative reporter. Before joining The Times in 1997, he worked at The Washington Post. More about Benjamin Weiser

Jonah E. Bromwich covers criminal justice in New York, with a focus on the Manhattan district attorney's office, state criminal courts in Manhattan and New York City's jails. More about Jonah E. Bromwich

A version of this article appears in print on , Section A, Page 1 of the New York edition with the headline: Cohen Used Artificial Intelligence to Give Lawyer Bogus Cases

United States Court of Appeals
For the Second Circuit

August Term 2023

Submitted: December 15, 2023
Decided: January 30, 2024

No. 22-2057

MINHYE PARK,

Plaintiff-Appellant,

v.

DAVID DENNIS KIM,

Defendant-Appellee,

Appeal from the United States District Court
for the Eastern District of New York
No. 20CV02636, Pamela K. Chen, *Judge*.

Before: PARKER, NATHAN, and MERRIAM, *Circuit Judges*.

Plaintiff-Appellant Minhye Park appeals from an August 25, 2022, judgment of the United States District Court for the Eastern District of New York (Chen, J.) dismissing her action against Defendant-Appellee David Dennis Kim, pursuant to Rules 37 and 41(b) of the Federal Rules of Civil Procedure, for her persistent and knowing violation of court orders. The record demonstrates Park's sustained and willful intransigence in spite of repeated warnings that failure to comply would result in the dismissal of the action. Accordingly, we AFFIRM the judgment of dismissal.

We separately address the fact that Park's counsel, Attorney Jae S. Lee, has admitted to citing a non-existent state court decision in her reply brief to this Court. Counsel reports that she relied on a generative artificial intelligence tool, ChatGPT, to identify precedent that might support her arguments, and did not read or otherwise confirm the validity of the (non-existent) decision she cited. Because this conduct falls well below the basic obligations of counsel, we refer Attorney Lee to the Court's Grievance Panel, and further ORDER Attorney Lee to furnish a copy of this decision to her client.

JAE S. LEE, JSL Law Offices P.C., Uniondale, NY, *for Plaintiff-Appellant.*

ALEJANDRA R. GIL, Heidell, Pittoni, Murphy & Bach, LLP, White Plains, NY, *for Defendant-Appellee.*

PER CURIAM:

Plaintiff-Appellant Minhye Park appeals from the August 25, 2022, judgment of the United States District Court for the Eastern District of New York (Chen, J.) dismissing her action against Defendant-Appellee David Dennis Kim, pursuant to Rules 37 and 41(b) of the Federal Rules of Civil Procedure. We assume the parties' familiarity with the underlying facts, procedural history, and issues on appeal, to which we refer only as necessary to explain our decision to affirm the District Court's judgment.

We separately address the conduct of Park's counsel, Attorney Jae S. Lee. Lee's reply brief in this case includes a citation to a non-existent case, which she

admits she generated using the artificial intelligence tool ChatGPT. Because citation in a brief to a non-existent case suggests conduct that falls below the basic obligations of counsel, we refer Attorney Lee to the Court's Grievance Panel, and further direct Attorney Lee to furnish a copy of this decision to her client, Plaintiff-Appellant Park.

STANDARD OF REVIEW

"We review a district court's imposition of sanctions for abuse of discretion." Wolters Kluwer Fin. Servs., Inc. v. Scivantage, 564 F.3d 110, 113 (2d Cir. 2009); see also Agiwal v. Mid Island Mortg. Corp., 555 F.3d 298, 302 (2d Cir. 2009) (dismissal pursuant to Rule 37); Baptiste v. Sommers, 768 F.3d 212, 216 (2d Cir. 2014) (dismissal pursuant to Rule 41(b)).

RULE 37 AND RULE 41(b)

Rule 37 provides: "If a party . . . fails to obey an order to provide or permit discovery . . . the court where the action is pending may issue further just orders . . . [including] dismissing the action or proceeding in whole or in part." Fed. R. Civ. P. 37(b)(2)(A)(v). Dismissal under Rule 37 is appropriate "only when a court finds willfulness, bad faith, or any fault" by the non-compliant litigant. Bobal v. Rensselaer Polytechnic Inst., 916 F.2d 759, 764 (2d Cir. 1990) (citation and

quotation marks omitted). "Whether a litigant was at fault or acted willfully or in bad faith are questions of fact, and we review the District Court's determinations for clear error." Agiwal, 555 F.3d at 302.

> Several factors may be useful in evaluating a district court's exercise of discretion to dismiss an action under Rule 37. These include: (1) the willfulness of the non-compliant party or the reason for noncompliance; (2) the efficacy of lesser sanctions; (3) the duration of the period of noncompliance, and (4) whether the non-compliant party had been warned of the consequences of . . . noncompliance.

Id. (citation and quotation marks omitted).

Rule 41(b) authorizes a district court to dismiss an action "[i]f the plaintiff fails to prosecute or to comply with [the] rules or a court order." Fed. R. Civ. P. 41(b). We consider five factors in reviewing a Rule 41(b) dismissal:

> (1) the duration of the plaintiff's failure to comply with the court order, (2) whether plaintiff was on notice that failure to comply would result in dismissal, (3) whether the defendants are likely to be prejudiced by further delay in the proceedings, (4) a balancing of the court's interest in managing its docket with the plaintiff's interest in receiving a fair chance to be heard, and (5) whether the judge has adequately considered a sanction less drastic than dismissal.

Lucas v. Miles, 84 F.3d 532, 535 (2d Cir. 1996).

DISCUSSION

Over the course of the litigation before the District Court, Park continually and willfully failed to respond to and comply with the District Court's discovery

orders. Magistrate Judge Bloom issued numerous discovery orders and provided ample warning to Park and her counsel that timely compliance was expected. For example, at a telephonic conference on August 11, 2021, Judge Bloom granted Kim's motion to compel, warning Attorney Lee: "[Y]our client can be subject to sanctions, which could be as severe as dismissal of the case, if she fails to comply." J. App'x at 81-82.[1] On November 29, 2021, Judge Bloom again warned of the consequences of non-compliance, in no uncertain terms: "Plaintiff shall have one final opportunity to comply with the Court's discovery Order . . . . **This is a Court Order and plaintiff must comply.** This is plaintiff's last chance." J. App'x at 7 (emphasis in original). In that same order, Judge Bloom set a briefing schedule for filing a motion to dismiss for non-compliance, should such a motion be necessary. Finally, having still not received the ordered discovery more than seven months after Judge Bloom's August 2021 order, Kim moved to dismiss based on Park's failure to comply with court orders and discovery obligations. See J. App'x at 292-93.

[1] The written order issued after that conference was also very clear: "**This is a Court Order and plaintiff must comply.** Plaintiff is warned that if she fails to comply with the Court's Order to produce discovery, she may be subject to sanctions, which could include dismissal of this action." J. App'x at 5 (emphasis in original).

In her report and recommendation, Judge Bloom carefully considered all of the requirements of Rule 37 and Rule 41(b), including the availability of lesser sanctions, and concluded that dismissal was appropriate. Judge Chen, the presiding District Judge, reviewed Park's objections to the report and recommendation in detail, overruled them, adopted the report and recommendation, and issued an order of dismissal on August 24, 2022.

On appeal, Park reiterates her complaints about Kim's alleged discovery abuses, as well as her conclusory assertion that she in fact complied with the relevant discovery orders. As Judge Bloom and Judge Chen found, these arguments are meritless, lack foundation in the record, and completely ignore the actual orders issued by Judge Bloom. Accordingly, we conclude that Park's noncompliance amounted to "sustained and willful intransigence in the face of repeated and explicit warnings from the court that the refusal to comply with court orders . . . would result in the dismissal of [the] action." Valentine v. Museum of Mod. Art, 29 F.3d 47, 50 (2d Cir. 1994). As such, we affirm the judgment of the District Court.

PLAINTIFF'S IMPROPER BRIEFING BEFORE THIS COURT

We must also address a separate matter concerning the conduct of Park's

counsel, Attorney Lee. Park's reply brief in this appeal was initially due May 26, 2023. After seeking and receiving two extensions of time, Attorney Lee filed a defective reply brief on July 25, 2023, more than a week after the extended due date. On August 1, 2023, this Court notified Attorney Lee that the late-filed brief was defective, and set a deadline of August 9, 2023, by which to cure the defect and resubmit the brief. Attorney Lee did not file a compliant brief, and on August 14, 2023, this Court ordered the defective reply brief stricken from the docket. Attorney Lee finally filed the reply brief on September 9, 2023.[2]

The reply brief cited only two court decisions. We were unable to locate the one cited as "Matter of Bourguignon v. Coordinated Behavioral Health Servs., Inc., 114 A.D.3d 947 (3d Dep't 2014)." Appellant's Reply Br. at 6. Accordingly, on November 20, 2023, we ordered Park to submit a copy of that decision to the Court by November 27, 2023. On November 29, 2023, Attorney Lee filed a Response with the Court explaining that she was "unable to furnish a copy of the decision." Response to November 20, 2023, Order of the Court, at 1,

[2] Attorney Lee filed the reply brief together with a motion to reconsider the Court's prior order striking the non-compliant brief. The Court later granted that motion to reconsider and accepted the September 9, 2023, version of the reply brief.

Park v. Kim, No. 22-2057-cv (2d Cir. Nov. 29, 2023), ECF No. 172 (hereinafter, "Response"). Although Attorney Lee did not expressly indicate as much in her Response, the reason she could not provide a copy of the case is that it does not exist – and indeed, Attorney Lee refers to the case at one point as "this non-existent case." Id. at 2.

Attorney Lee's Response states:

> I encountered difficulties in locating a relevant case to establish a minimum wage for an injured worker lacking prior year income records for compensation determination . . . . Believing that applying the minimum wage to in injured worker in such circumstances under workers' compensation law was uncontroversial, I invested considerable time searching for a case to support this position but was unsuccessful.
> . . .
> Consequently, I utilized the ChatGPT service, to which I am a subscribed and paying member, for assistance in case identification. ChatGPT was previously provided reliable information, such as locating sources for finding an antic furniture key. The case mentioned above was suggested by ChatGPT, I wish to clarify that I did not cite any specific reasoning or decision from this case.

Id. at 1-2 (sic).

All counsel that appear before this Court are bound to exercise professional judgment and responsibility, and to comply with the Federal Rules of Civil Procedure. Among other obligations, Rule 11 provides that by presenting a submission to the court, an attorney "certifies that to the best of the person's

knowledge, information, and belief, formed after an inquiry reasonable under the circumstances . . . 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." Fed. R. Civ. P. 11(b)(2); see also N.Y. R. Pro. Conduct 3.3(a) (McKinney 2023) ("A lawyer shall not knowingly: (1) make a false statement of . . . law to a tribunal."). "Rule 11 imposes a duty on attorneys to certify that they have conducted a reasonable inquiry and have determined that any papers filed with the court are well grounded in fact, [and] legally tenable." Cooter & Gell v. Hartmarx Corp., 496 U.S. 384, 393 (1990). "Under Rule 11, a court may sanction an attorney for, among other things, misrepresenting facts or making frivolous legal arguments." Muhammad v. Walmart Stores E., L.P., 732 F.3d 104, 108 (2d Cir. 2013) (per curiam).

At the very least, the duties imposed by Rule 11 require that attorneys read, and thereby confirm the existence and validity of, the legal authorities on which they rely. Indeed, we can think of no other way to ensure that the arguments made based on those authorities are "warranted by existing law," Fed. R. Civ. P. 11(b)(2), or otherwise "legally tenable." Cooter & Gell, 496 U.S. at

393. As a District Judge of this Circuit recently held when presented with non-existent precedent generated by ChatGPT: "A fake opinion is not 'existing law' and citation to a fake opinion does not provide a non-frivolous ground for extending, modifying, or reversing existing law, or for establishing new law. An attempt to persuade a court or oppose an adversary by relying on fake opinions is an abuse of the adversary system." Mata v. Avianca, Inc., No. 22CV01461(PKC), 2023 WL 4114965, at *12 (S.D.N.Y. June 22, 2023).

Attorney Lee states that "it is important to recognize that ChatGPT represents a significant technological advancement," and argues that "[i]t would be prudent for the court to advise legal professionals to exercise caution when utilizing this new technology." Response at 2. Indeed, several courts have recently proposed or enacted local rules or orders specifically addressing the use of artificial intelligence tools before the court.[3] But such a rule is not necessary to

[3] See, e.g., Notice of Proposed Amendment to 5th Cir. R. 32.3, U.S. Ct. of Appeals for the Fifth Cir., https://www.ca5.uscourts.gov/docs/default-source/default-document-library/public-comment-local-rule-32-3-and-form-6 [https://perma.cc/TD4F-WLV2] (Proposed addition to local rule: "[C]ounsel and unrepresented filers must further certify that no generative artificial intelligence program was used in drafting the document presented for filing, or to the extent such a program was used, all generated text, including all citations and legal analysis, has been reviewed for accuracy and approved by a human."); E.D. Tex. Loc. R. AT-3(m) ("If the lawyer, in the exercise of his or her professional legal

inform a licensed attorney, who is a member of the bar of this Court, that she must ensure that her submissions to the Court are accurate.

Attorney Lee's submission of a brief relying on non-existent authority reveals that she failed to determine that the argument she made was "legally tenable." Cooter & Gell, 496 U.S. at 393. The brief presents a false statement of law to this Court, and it appears that Attorney Lee made no inquiry, much less the reasonable inquiry required by Rule 11 and long-standing precedent, into the validity of the arguments she presented. We therefore **REFER** Attorney Lee to the Court's Grievance Panel pursuant to Local Rule 46.2 for further investigation, and for consideration of a referral to the Committee on Admissions and

---

judgment, believes that the client is best served by the use of technology (e.g., ChatGPT, Google Bard, Bing AI Chat, or generative artificial intelligence services), then the lawyer is cautioned that certain technologies may produce factually or legally inaccurate content and should never replace the lawyer's most important asset – the exercise of independent legal judgment. If a lawyer chooses to employ technology in representing a client, the lawyer continues to be bound by the requirements of Federal Rule of Civil Procedure 11, Local Rule AT-3, and all other applicable standards of practice and must review and verify any computer-generated content to ensure that it complies with all such standards."); Self-Represented Litigants (SRL), U.S. Dist. Ct. for the E. Dist. of Mo., https://www.moed.uscourts.gov/self-represented-litigants-srl [https://perma.cc/Y7QG-VVEF] ("No portion of any pleading, written motion, or other paper may be drafted by any form of generative artificial intelligence. By presenting to the Court . . . a pleading, written motion, or other paper, self-represented parties and attorneys acknowledge they will be held responsible for its contents. See Fed. R. Civ. P. 11(b).").

Grievances. See 2d Cir. R. 46.2.

We further **ORDER** Attorney Lee to provide a copy of this ruling to Plaintiff-Appellant Park – translated into Korean if necessary to permit Park to understand it – within twenty-one days, and to file a certification on the docket in this case attesting that she has done so.

RECENT ETHICS DEVELOPMENTS IN BANKRUPTCY AND LITIGATION

February 28, 2024

PANELISTS

- Carolynn Beck – Eisner, LLP
- Jeffrey S. Kwong – Levene, Neale, Bender, Yoo & Golubchik L.L.P.
- Alston L. Lew – Murphy, Pearson, Bradley & Feeney, P.C.
- Mike Paek – U.S. Bankruptcy Court, Southern District of New York

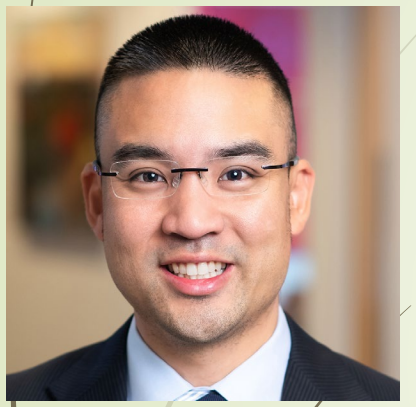
Carolynn Beck

Carolynn operates a national practice representing companies in commercial litigation, employment litigation and employment counseling. She represents companies in a wide variety of industries including technology and life sciences.

Carolynn's many litigation successes include her representation of a female reproductive health company in a breach of contract and tortious interference action against a household-name genetic testing company, which settled confidentially before trial.

A leader in multiple bar organizations, Carolynn previously served on the Board of Governors of the National Asian Pacific American Bar Association (NAPABA), the voice of more than 60,000 members of the Asian Pacific American legal community. The wife of a retired Navy SEAL, she co-chairs the NAPABA Military & Veteran's Network, which spearheaded a NAPABA resolution supporting federal and state legislation to ease bar waiver rules for military spouses.

Jeffrey S. Kwong



Mr. Kwong is a Partner at Levene, Neale, Bender, Yoo & Golubchik L.L.P. ("LNBYG"), and represents Chapter 11 debtors, unsecured creditor committees, secured and unsecured creditors, and parties in bankruptcy litigation and appeals from a variety of industries, including hotels and hospitality, lending and banking, commercial real estate, restaurants, retail, and healthcare. Mr. Kwong's prior Chapter 11 debtor engagements include Cornerstone Apparel, Inc., Anna's Linens, Tala Jewelers, Inc., and Green Fleet Systems, LLC. Further, he has represented commercial landlords in some of the largest retail bankruptcy cases filed across the country in recent years.

Prior to joining LNBYG, Mr. Kwong served for two years as law clerk to the Honorable Deborah J. Saltzman, United States Bankruptcy Judge for the Central District of California. Mr. Kwong obtained his J.D. in 2012 from the University of California, Berkeley, School of Law, where he served as an editor for the Berkeley Journal of International Law and a Senior Articles Editor for the Asian American Law Journal. He received his undergraduate degree, summa cum laude, from the University of California, San Diego.

Alston L. Lew

Mr. Lew is a Director at Murphy, Pearson, Bradley & Feeney P.C.'s San Francisco office. He represents businesses and individuals in all phases of civil and business litigation. Mr. Lew counsels clients in a variety of matters related to business and real estate law, professional liability, and insurance defense. A portion of his practice is dedicated to the defense of attorneys against legal malpractice claims either in litigation or before the California State Bar Court. Mr. Lew also defends other licensed professionals such as CPAs, real estate brokers, insurance brokers and doctors in professional liability cases. As part of his real estate practice, Mr. Lew defends landlords and homeowner associations involved in litigation relating to management/ownership of real property. Prior to joining Murphy, Pearson, Bradley & Feeney, Mr. Lew worked as a research attorney and law clerk for the Honorable Judge Stuart Hing in the Superior Court of California, County of Alameda. He also was an in- house litigation attorney for a nationally recognized insurance carrier.

Mr. Lew received his law degree from Golden Gate University School of Law with an emphasis on intellectual property and his undergraduate degree in Political Science from the University of San Francisco. Mr. Lew is an active member of AABA Bay Area, NAPABA's Insurance Committee and the Mass Torts and Class Action Committee, and the Asian American Insurance Network.

Mike Paek

Kyu (Mike) Paek is the Chief Deputy Clerk for the Southern District of New York Bankruptcy Court. In that role, Mike manages and oversees all aspects of court operations for one of the busiest bankruptcy courts in the country. Mike is dedicated to fostering a first-class venue for distressed companies to reorganize and for individuals to achieve a fresh start. His leadership has been instrumental in guiding the court through the countless obstacles created by the Covid-19 pandemic.

Prior to his current role, Mike served as career law clerk to the Honorable Stuart M. Bernstein, assisting with numerous high-profile "mega" chapter 11 cases, chapter 15 cross-border cases, brokerage liquidations under the Securities Investor Protection Act, and multi-million-dollar adversary proceedings. He previously worked in the Business Reorganization Group of Schulte Roth & Zabel LLP, where he represented debtors, creditors, DIP lenders and investors.

DISCLAIMER

- The opinions expressed during the panel discussion are those of the panelists and not necessarily the entities that they work for.

(1) CASE STUDY: JUDGE DAVID JONES AND JACKSON WALKER

- Case Study: Judge David Jones and Jackson Walker
- Attorney Disclosure Obligations, and Relevant Rules
- Judge Disclosure Obligations, and Relevant Rules
- ABA Model/ State Bar Ethics Rules
- Implications for Attorneys Practicing Bankruptcy or Litigation Generally

CASE STUDY: Judge Jones and Jackson walker

- 1. **Judge Jones**: Bankruptcy Judge David R. Jones was appointed to the bench in 2011. Judge Jones was a prominent corporate bankruptcy judge with the largest* commercial chapter 11 docket in the country (such as JCPenney, Neiman Marcus, Denbury Resources).**
- 2. **Ms. Freeman**: Elizabeth Freeman was a former law clerk to Judge Jones, and a partner at Jackson Walker LLP ("JW") from at least 2017 until December 2022. In December 2022, she founded the Law Office of Liz Freeman ("Freeman Law Office").
- 3. **JW**: JW is a prominent regional firm in Texas and frequently served as local counsel for corporate debtors who filed chapter 11 petitions in Bankr. S.D. Tex. where Judge Jones served.
- 4. **Incident**: Judge Jones was in an intimate relationship*** with Freeman and cohabitated since approximately 2017. JW regularly appeared before Judge Jones. Prior to the relationship becoming "public," the relationship was not disclosed in the employment applications filed by JW with the Court, and Judge Jones did not recuse himself from JW/ Freeman Law Office cases. Judge Jones approved at least $13 million of fees to JW (Freeman worked on some of those cases)

***In 2023, of the 54 large Chapter 11 cases filed, 25 landed in SDTX, where only two judges, including Jones, oversaw large restructurings. https://www.ft.com/content/574f0940-d82e-4e4a-98bd-271058cce434 ("Houston brought a service orientation to the bankruptcy court system. They woke up Delaware and SDNY [Southern District of New York].").**
****In 2016, Judge Jones creates the complex case pool that no longer randomly assigns "mega-cases" to all SDTX Bankruptcy Judges.**
*****According to the NY Post, "a survivorship agreement attached to the deed of the house was obtained by The Journal, which reported that it lists both Jones and Freeman as owners, blatantly stating that the two own the property jointly. If one of the two dies, the other inherits the property[.]" https://nypost.com/2023/10/16/bankruptcy-judge-resigns-after-relationship-with-attorney-revealed/**

RELEVANT RULES REGARDING BANKRUPTCY EMPLOYMENT

- **Bankruptcy Code Section 327(a)**
  - (a) Except as otherwise provided in this section, the trustee, with the court's approval, may employ one or more attorneys, accountants, appraisers, auctioneers, or other professional persons, that do not hold or represent an interest adverse to the estate, and that are disinterested persons, to represent or assist the trustee in carrying out the trustee's duties under this title.
- **Bankruptcy Code Section 327(e)**
  - (e) The trustee, with the court's approval, may employ, for a specified special purpose, other than to represent the trustee in conducting the case, an attorney that has represented the debtor, if in the best interest of the estate, and if such attorney does not represent or hold any interest adverse to the debtor or to the estate with respect to the matter on which such attorney is to be employed.
- **Federal Rule of Bankruptcy Procedure 2014**
  - (a) APPLICATION FOR AND ORDER OF EMPLOYMENT. . . . The application shall state the specific facts showing the necessity for the employment, the name of the person to be employed, the reasons for the selection, the professional services to be rendered, any proposed arrangement for compensation, and, to the best of the applicant's knowledge, all of the person's connections with the debtor, creditors, any other party in interest, their respective attorneys and accountants, the United States trustee, or any person employed in the office of the United States trustee. The application shall be accompanied by a verified statement of the person to be employed setting forth the person's connections with the debtor, creditors, any other party in interest, their respective attorneys and accountants, the United States trustee, or any person employed in the office of the United States trustee.

RELEVANT RULES REGARDING BANKRUPTCY COMPENSATION

- **Bankruptcy Code Section 328**
  - (a) The trustee, or a committee appointed . . . with the court's approval, may employ or authorize the employment of a professional person under section 327 or 1103 of this title, as the case may be, on any reasonable terms and conditions of employment, including on a retainer, on an hourly basis, on a fixed or percentage fee basis, or on a contingent fee basis. Notwithstanding such terms and conditions, the court may allow compensation different from the compensation provided under such terms and conditions after the conclusion of such employment, if such terms and conditions prove to have been improvident in light of developments not capable of being anticipated at the time of the fixing of such terms and conditions.
  - (c) . . . the court may deny allowance of compensation for services and reimbursement of expenses of a professional person employed . . . if, at any time during such professional person's employment . . . such professional person is not a disinterested person, or represents or holds an interest adverse to the interest of the estate with respect to the matter on which such professional person is employed.
- **Bankruptcy Code Section 330(a)**
  - (2) The court may, on its own motion or on the motion of the United States Trustee, the United States Trustee for the District or Region, the trustee for the estate, or any other party in interest, award compensation that is less than the amount of compensation that is requested.
  - (5) The court shall reduce the amount of compensation awarded under this section by the amount of any interim compensation awarded under section 331, and, if the amount of such interim compensation exceeds the amount of compensation awarded under this section, may order the return of the excess to the estate.

RELEVANT RULES REGARDING JUDGE'S ABILITY TO PRESIDE OVER, OR APPROVE COMPENSATION IN BANKRUPTCY CASES

- **Bankruptcy Rule 5002**
  - (a) APPROVAL OF APPOINTMENT OF RELATIVES PROHIBITED . . . The employment of an individual as attorney, accountant, appraiser, auctioneer, or other professional person pursuant to §§327, 1103, or 1114 may be approved by the court if the individual is a relative* of the United States trustee in the region in which the case is pending, ***unless the court finds that the relationship with the United States trustee renders the employment improper under the circumstances of the case***."
  - (b) JUDICIAL DETERMINATION THAT APPROVAL OF APPOINTMENT OR EMPLOYMENT IS IMPROPER. A bankruptcy judge may not. . . approve the employment of a person as an attorney, accountant, appraiser, auctioneer, or other professional person . . . ***if that person is or has been so connected with such judge or the United States trustee as to render the appointment or employment improper***.

* Bankruptcy Code Seciton101(34) defines relative to be an "individual related by affinity or consanguinity within the third degree as determined by the common law, or individual in a step or adoptive relationship within such third degree."

RELEVANT RULES REGARDING JUDGE'S ABILITY TO PRESIDE OVER, OR APPROVE COMPENSATION IN BANKRUPTCY CASES

- **Bankruptcy Rule 5004(b)**
  - (a) [APPLICABILITY OF 28 U.S.C. § 455]
    - 28 U.S.C. § 455(a) ("Any . . . Judge . . .shall disqualify himself in any proceeding in which his ***impartiality might reasonably be questioned.***").
    - 28 U.S.C. § 455(b) ("He shall also disqualify himself in the following circumstances . . . (4) He knows that he . . . or his spouse . . . has a ***financial interest in the subject matter in controversy*** . . . or any **other interest that could be substantially affected by the outcome of the proceeding** . . . (5) He or his spouse, or a person within the third degree of relationship to either of them, or the spouse of such a person . . . Is acting as a lawyer in the proceeding [or] . . . Is known by the judge to have an interest that could be substantially affected by the outcome of the proceeding[.]").
  - (b) DISQUALIFICATION OF JUDGE FROM ALLOWING COMPENSATION. A bankruptcy judge shall be disqualified from allowing compensation to a person who is a relative of the bankruptcy judge or with whom the judge is so connected as to render it improper for the judge to authorize such compensation.

AFTERMATH: JUDGE JONES

- 5th Circuit Complaint: Chief Judge Priscilla Richman filed a judicial misconduct complaint against Judge Jones on Oct. 13, 2023.
  - Code of Conduct for U.S. Judges, Canon 2: "A judge should avoid impropriety and the appearance of impropriety in all activities." *See also* 28 U.S.C. § 455.
  - Canon 2B: "A judge should now allow family, . . . or other relationship to influence judicial conduct or judgment. A judge should neither lend the prestige of the judicial office to advance the private interests of the judge or others nor convey or permit others to ***convey the impression that they are in a special position to influence the judge***."
  - Canon 3B(3): "A just should exercise the power of appointment fairly and only on the basis of merit, avoiding unnecessary appointments, nepotism, and favoritism."
  - Canon 3C(1): "A judge shall disqualify himself or herself in a proceeding in which the judge's impartiality might reasonably be questioned . . . ."
  - Canon 3C(1)(c): The judge should be disqualified if "the judge knows that . . . the judge's spouse . . . has a financial interest in the subject matter in controversy or in a party to the proceeding, or any other interest that could be affected substantially by the outcome of the proceeding."
  - Canon 3C(1)(d): The judge should be disqualified if "the judge or the judge's spouse, or a person related to either within the third degree of relationship, or the spouse of such a person is . . . acting as a lawyer in the proceeding [or] known by the judge to have an interest that could be substantially affected by the outcome of the proceeding[.]"
- Resignation: Ethics probe dropped by the Circuit following Judge Jones' resignation on November 15, 2023.
- Judge Jones' Assertion of Judicial Immunity, And Request for DOJ to Defend him in Civil Suit.

AFTERMATH: JW AND FREEMAN LAW OFFICES

- Creditors and the government "watchdog" in bankruptcy cases, the Office of the United States Trustee, have filed motions for reconsideration and/or disgorgement of fees previously awarded to JW and Freeman Law Office, claiming that the orders approving compensation and/or employment are "tainted."
- Potential Consequences
  - At least $13 million of approved fees and expenses to JW are at risk of being disgorged or "clawed back"
  - Rejection of future fees
  - Disqualification
  - Criminal prosecution?
    - In a 1998 case, bankruptcy partner John Gellene of what's now Milbank LLP was accused of filing a false Rule 2014 disclosure declaration by federal prosecutors. After he was convicted of "knowingly and fraudulently" making false declarations under oath in two Rule 2014 bankruptcy applications, he was sentenced to 15 months in prison, and his firm had to disgorge $1.8 million in fees.*
- "Screening" Sufficient?

* https://www.americanbar.org/groups/business_law/resources/podcast/john-gellene-a-bankruptcy-lawyer-goes-to-jail/; James Nani, Bankruptcy Rules for Disclosing Relationships: Explained (Bloomberg Law News, November 2023).

TAKEAWAYS

- Takeaways: "**Disclose, Disclose, Disclose**." Its always better to err on being safe than to face the consequences later.
  - **Don't try to determine what is material**. Judge Jones previously stated that he believed he was not required to disclose the connection because Freeman did not appear in his courtroom, and that there was no economic benefit to him from her legal work and the two are not married.
- Failure to disclose may also risk disgorgement by co-counsel.
  - In Jones, an amended complaint filed by Van Deelen argued that "Judge Jones and Freeman plainly deceived the public and interested parties in bankruptcies by failing to disclose their relationship," the complaint said. "But they did not deceive Jackson Walker or Kirkland & Ellis. Both firms knew of the relationship and used it to profit."*
- Bankruptcy professionals may want to perform more stringent conflicts checks, and make more potent disclosures in employment applications. For example, disclosing that certain law clerks worked for certain judges in a district.
- Disclosure issues don't only impact bankruptcy practitioners. For litigators, you may serve, or one days serve as as "special litigation counsel" to the Debtor or to the Bankruptcy Estate.
  - *See e.g.*, *In re Paris*, 568 B.R. 810 (Bankr. C.D. Cal. 2017)

* https://news.bloomberglaw.com/bankruptcy-law/kirkland-jackson-walker-named-in-bankruptcy-judge-ethics-suit

Relevant Model ABA/ State Ethics Rules

- **ABA Model Rule 3.3 - Candor Toward the Tribunal**
  - "(a) A lawyer shall not knowingly: (1) make a false statement of fact or law to a tribunal or fail to correct a false statement of material fact or law previously made to the tribunal by the lawyer . . . ."
  - "(b)A lawyer who represents a client in an adjudicative proceeding and who knows that a person intends to engage, is engaging or has engaged in criminal or fraudulent conduct related to the proceeding shall take reasonable remedial measures, including, if necessary, disclosure to the tribunal."
  - "(c) The duties stated in paragraphs (a) and (b) continue to the conclusion of the proceeding, and apply even if compliance requires disclosure of information otherwise protected by Rule 1.6."
- See also California, New York, and Texas rules regarding duty of candor towards the Court.

(2) ETHICS AND ARTIFICIAL INTELLIGENCE

- Case Study: *Mata v. Avianca, Inc.*, No. 22-CV-1461 (PKC), 2023 U.S. Dist. LEXIS 108263 (S.D.N.Y. June 22, 2023).
- Ethical Duties
  - Duty of Competence and Diligence
  - Duty of Confidentiality (ABA Model Rule 1.6)
  - Duty to Supervise (ABA Model Rules 5.1 and 5.3)
  - Duty of Candor to Tribunal (ABA Model Rule 3.1, 3.3)
    - See California Guidelines*
- Local Rules and Judges Requiring Attorneys To Attest as to whether they have used generative artificial intelligence in court filings.
  - Hon. Brantley Starr, "Mandatory Certification Regarding Generative Artificial Intelligence [Standing Order]," (N.D. Tex.).
  - Hon. Gabriel A. Fuentes, "Standing Order For Civil Cases Before Magistrate Judge Fuentes," [N.D. Ill.]

*https://www.calbar.ca.gov/Portals/0/documents/ethics/Generative-AI-Practical-Guidance.pdf

Mata v. Avianca, Inc., No. 22-CV-1461 (PKC), 2023 U.S. Dist. LEXIS 108263 (S.D.N.Y. June 22, 2023).

ChatGPT

Ethical Duties

Local Rules Regarding Disclosure of AI Usage

- Hon. Brantley Starr, "Mandatory Certification Regarding Generative Artificial Intelligence [Standing Order]," (N.D. Tex.).
- Hon. Gabriel A. Fuentes, "Standing Order For Civil Cases Before Magistrate Judge Fuentes," [N.D. Ill.]

Select Artificial Intelligence Cases

- "*Michael Cohen Used Artificial Intelligence in Feeding Lawyer Bogus Cases*" available at https://www.nytimes.com/2023/12/29/nyregion/michael-cohen-ai-fake-cases.html (describing issue with Mr Cohen's usage of Google Bard).

Use of AI in E-Discovery

- *Moore v. Publicis Groupe*, 287 F.R.D. 182, 193 (S.D.N.Y. 2012), *adopted sub nom. Moore v. Publicis Groupe SA*, No. 11 CIV. 1279 ALC AJP, 2012 WL 1446534 (S.D.N.Y. Apr. 26, 2012)
  - ("What the Bar should take away from this Opinion is that computer-assisted review is an available tool and should be seriously considered for use in large-data-volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review. Counsel no longer have to worry about being the "first" or "guinea pig" for judicial acceptance of computer-assisted review. As with keywords or any other technological solution to ediscovery, counsel must design an appropriate process, including use of available technology, with appropriate quality control testing, to review and produce relevant ESI while adhering to Rule 1 and Rule 26(b)(2)(C) proportionality. Computer-assisted review now can be considered judicially-approved for use in appropriate cases.").
- *Entrata, Inc. v. Yardi Sys., Inc.*, No. 2:15-CV-00102, 2018 WL 5470454, at *7 (D. Utah Oct. 29, 2018
  - ("This case is unique because the parties never reached any agreement. Nevertheless, the court is persuaded that because it is "black letter law" that courts will permit a producing party to utilize TAR [Technology-Assisted Review for expedited review of discovery documents], Entrata was not required to seek approval from the Magistrate Court to use TAR where there was never an agreement to utilize a different search methodology. The court agrees with the Magistrate Judge that if Yardi had concerns about Entrata's use of TAR, it should have sought intervention long before the last day of fact discovery. "

State Bar of California Approves Guidelines Concerning Ethics In Using Generative AI In Legal Practice

- On November 16, 2023, the State Bar of California approved guidelines based on existing rules to assist lawyers with their ethical obligations in the use of generative AI. The State Bar called it: "Practical Guidance for the Use of Generative Artificial Intelligence in the Practice of Law"
- The California State Bar emphasizes existing applicable State Bar rules based on the Rules of Professional Conduct.
- Highlighted Rule Clarifications Include:
  - Rule 1.1 Duty of Competence
  - Rule 1.4 Duty to Provide Copies of significant documents to expressly permit provision by "electronic or other means."
  - Rule 1.16(e) – Lawyer's Duty to Release all client materials when terminating a representation to expressly include release of client materials that are in "electronic format."
  - Rule 4.4 – Requires a lawyer who receives inadvertently produced materials that are subject to the attorney-client privilege or confidential to notify the sender.
  - Rule 7.2(a) – clarifies advertising rules to provide a lawyer may advertise through "electronic means of communication, including public media."
  - Rule 7.5 – clarifies the scope of the rule governing a lawyer's professional designation to include logos and URLs.
  - https://www.calbar.ca.gov/Attorneys/Conduct-Discipline/Ethics/Ethics-Technology-Resources

Applicable California Rules of Professional Conduct Governing the Use of AI

- Rule 1.1 – Duty of Competence – a lawyer shall not intentionally, recklessly, repeatedly, or with gross negligence, fail to provide legal services with competence.
  - In the case of the Chat GPT case in New York, it was evident that the attorneys representing the plaintiff failed to closely supervise, review or ensure the accuracy of the legal citations submitted to the court as part of their brief that was generated by AI. Essentially, the attorneys submitted falsified and non-existent legal authority to the court thereby damaging their client's position.
  - Attorneys as part of this rule have an obligation to keep up to date on changes in the law and its practice, which includes the benefits and risks associated with relevant technology.

- Rule 1.4 Client Communication
  - Under Rule 1.4(a)(2): A lawyer shall reasonably consult with the client about the means by which to accomplish the client's objectives in the representation.
  - Given the present unreliability of AI, this rule potentially places a requirement on an attorney to disclose to the client the use of AI during the course of the representation either in the submission of a brief to the court; storage of data; use of the client's data via third party vendors such as Google, Bloomberg, etc.; issues concerning client confidentiality as a result of using AI generally.
  - From a budgetary standpoint, attorneys may need to go over the pros and cons on the use of AI with clients and also communicate potential conflicts that can arise in the use of AI.

- Rule 1.6 Safeguarding Client Information
  - Rule 1.6(a): "A lawyer shall not reveal information protected from disclosure by Business and Professions Code section 6068(e)(1) unless the client gives informed consent, subject to certain exceptions.
  - Source of consideration for attorneys that use AI is that they may be inadvertently disclosing client confidential information to a third-party vendor when inputting search terms into an AI system such as ChatGPT. Chat GPT stores all content inputted into the model, analyzes the information and maintains the information for future use. Inputted search terms may include sensitive client information, legal strategies and privileged communications.
  - Law Firms need to safeguard client information from improper disclosure by their attorneys if AI is used as part of legal research and also disclose the use of AI to their clients.
  - Attorneys must be well-trained on first and second-level review of AI-assisted document productions, which fall under the umbrella of "reasonable steps" attorneys must take to avoid inadvertent disclosure. (*Regents of Univ. of Cal. v. Superior Court*, 165 Cal.App.4th 672, 681-82 (2008).)

- Rule 3.1 and 3.3 – Meritorious Claims and Candor To Tribunal
- Rule 3.1(a)(2)P: A lawyer shall not present a claim or defense in litigation that is not warranted under existing law, unless it can be supported by a good faith argument for an extension, modification, or reversal of existing law.
- Rule 3.3(a)(1)-(2): A lawyer shall not knowingly make a false statement of fact or law to a tribunal or fail to correct a false statement of material fact or law previously made to the tribunal by the lawyer…or knowingly misquote to a tribunal the language of a book, statute, decision or other authority.

- Other potential applicable rules:
- 5.3 Supervision of non-lawyers – Rule 5.3(b) requires a lawyer to have direct supervision over a nonlawyer during a represented matter, employee or otherwise, to make reasonable efforts to ensure the person's conduct is compatible with the professional obligations of a lawyer.
- Rule 8.4.1 prohibits unlawful discrimination in the course of representing a client and the use of AI potentially can cause a lawyer to adversely affect a lawyer's representation of a client based on inherit biases in an AI's existing algorithms.

Applicable ABA Model Rules

- **ABA Resolution 112 & Model Rules**
- ABA Resolution 112 specifically calls out the expectation that lawyers understand the risks of AI, including the risks of bias and harm to the legal system. Indeed, the ABA **model rules of professional conduct** can already be seen as holding lawyers accountable for understanding AI tools.
- **Rule 1.1 Competence and Comment [8]** requires a lawyer to competently represent each client, "keep[ing] abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology."
- **Rule 1.4 Communication** requires reasonable consultation with clients about methods and choices for accomplishing the client's objectives. That should include accurately communicating the availability, effectiveness, risk, and overall impact on costs of relevant AI systems, including obtaining competencies for effective operation.

- **Rule 1.5 Fees** charges lawyers with responsibility not to demand unreasonable fees, necessitating consideration of AI's potential for accelerating work (such as by high recall, high precision document review).
- **Rule 1.6 Confidentiality** requires understanding the operation and security of AI systems to avoid unintended access to client information—e.g., by unsecure systems, "smart" assistants that transmit to the vendor, use of AI trained on a different client's data, open-source licenses that require sharing.
- **Rules 1.7, 1.9 Conflict of Interest** may implicate reuse of AI systems trained on client data, or taking a trained technology to a new firm.
- **Rule 2.1 Advisor** requires exercise of independent professional judgment, implicating lawyer understanding of the design, training, and operation of AI systems on whose outcomes the lawyer relies.
- **Rules 3.3 (Candor Toward the Tribunal), 3.4 (Fairness to Opposing Party and Counsel) and 4.1 (Truthfulness in Statements to Others)** require sound information on the effectiveness (as operated) of AI systems that provide information on which a lawyer relies in making factual assertions—e.g., a prosecutor's sentencing memo containing algorithmic assessment of recidivism risk, a rule 26(g) representation of reasonable completion of production, assertions on reliability of evidence emanating from AI systems.
- **Rules 5.1, 5.3 Supervision** impose on senior lawyers responsibility to supervise subordinates and third parties to ensure compliance with the rules of ethics; this in turn generates need for significant information on the trustworthiness of outputs from AI systems in order to supervise the uses of and representations about those outputs.

- **Rule 8.4(d) Misconduct** requires avoidance of conduct prejudicial to the administration of justice. The responsible lawyer must not only know the effectiveness of AI systems whose output is offered to or used by the court, but also envision both the impact on participants in the justice system and the ability of the justice system to police for bias, prejudice, and other unintended consequences.
- **Rule 8.4(g) Misconduct** prohibits professional conduct that a lawyer reasonably should know is discriminatory. Accordingly, a lawyer must understand potential biases of AI systems they employ—insight that the system designer or distributor may not have gathered. For example, hiring algorithms trained on the data of a non-diverse firm's current partners will likely select resumes reflecting similar schools, hobbies, zip codes, and accordingly candidates. An AI system that recommends prospective trial teams based on successes from prior non-diverse teams may well replicate their characteristics.

Potentially upcoming legislation in California concerning the use of AI that could serve as a guideline for other States

- September 6, 2023 – California's Governor Gavin Newsome signed an Executive Order – Order N-12-23 governing the use of generative AI by California agencies. Under the order, California government agencies have until January 2024 to issue guidelines for the procurement, uses and training required for use of generative AI.
  - "Whereas the State of California is committed to accuracy, reliability, and ethical outcomes when adopting AI technology, engaging and supporting historically vulnerable and marginalized communities, and serving its residents, workers, and businesses in a transparent, engaged and equitable way…"
  - "Whereas the unprecedented speed of innovation and deployment of GenAI technologies necessitates measured guardrails to protect against potential risks or malicious uses, including, but not limited to, bioterrorism, cyberattacks, disinformation, deception, and discrimination or bias.."
  - By January 2024, the Government Operations Agency and several other agencies will issue general guidelines for public sector use of GenAI and rely on guidelines issued by the White House.
  - No later than March 2024, the California Cybersecurity Integration Center and California State Threat Assessment Center and various other agencies will perform a joint risk analysis of potential threats to California's infrastructure.

New York City adopts one of the first laws in the country based on AI

- New York City adopted a local rule – Local Rule 144, codified in the NYC Administrative Code 20-870, *et. seq.*, which prohibits employers from using an automated employment decision tool (AEDT) to screen a candidate or employee for an employment decision unless the tool has been subject to a bias audit conducted within one year prior, and the employer complies with other notice requirements. The law went into effect July 5, 2023.
  - There are questions concerning enforceability of the law.
  - Questions about what type of access an auditor would get to a company's information in order to determine how the audited company's hiring practices work.
  - The law requires employers relying on AI to inform candidates of the use of AI as part of their hiring practices.

Federal Government Guidelines on AI use

- The Biden Administration's Recommendations concerning the use of AI by the federal government:
  - Executive Order 13960 of December 3, 2020 – "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government, 85 Fed Reg. 236 at 78939 (December 8, 2020). Federal agencies must "design, develop, acquire, and use AI in a manner that fosters public trust and confidence while protecting privacy, civil rights, civil liberties, and American values.
  - The Federal Government issued what it calls a "Blueprint For An AI Bill of Rights" https://www.whitehouse.gov/ostp/ai-bill-of-rights/#discrimination
  - Federal Government has created a website specifically indicating the Biden Administration's position on the use of AI: www.ai.gov.