Volume 8, No. 1 | January–February 2025

RAIL

The Journal of Robotics, Artificial Intelligence & Law

fastcase FULL COURT PRESS

RAIL
The Journal of Robotics, Artificial Intelligence & Law
Volume 8, No. 1 | January–February 2025

THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW (ISSN 2575-5633 (print) /ISSN 2575-5617 (online) at $495.00 annually is published six times per year by Full Court Press, a Fastcase, Inc., imprint. For customer support, please contact Fastcase, Inc., 729 15th Street, NW, Suite 500, Washington, D.C. 20005, 202.999.4777 (phone), or email customer service at support@fastcase.com.

Publishing Staff
Publisher: Leanne Battle
Production Editor: Sharon D. Ray
Cover Art Design: Juan Bustamante

Cite this publication as:

The Journal of Robotics, Artificial Intelligence & Law (Fastcase)

A Full Court Press, Fastcase, Inc., Publication

Editorial Office

729 15th Street, NW, Suite 500, Washington, D.C. 20005
https://www.fastcase.com/

POSTMASTER: Send address changes to THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW, 729 15th Street, NW, Suite 500, Washington, D.C. 20005.

Articles and Submissions

Direct editorial inquiries and send material for publication to:

Steven A. Meyerowitz, Editor-in-Chief, Meyerowitz Communications Inc., 26910 Grand Central Parkway, #18R, Floral Park, NY 11005, smeyerowitz@meyerowitzcommunications.com, 631.291.5541.

Material for publication is welcomed—articles, decisions, or other items of interest to attorneys and law firms, in-house counsel, corporate compliance officers, government agencies and their counsel, senior business executives, scientists, engineers, and anyone interested in the law governing artificial intelligence and robotics. This publication is designed to be accurate and authoritative, but neither the publisher nor the authors are rendering legal, accounting, or other professional services in this publication. If legal or other expert advice is desired, retain the services of an appropriate professional. The articles and columns reflect only the present considerations and views of the authors and do not necessarily reflect those of the firms or organizations with which they are affiliated, any of the former or present clients of the authors or their firms or organizations, or the editors or publisher.

QUESTIONS ABOUT THIS PUBLICATION?

For questions about the Editorial Content appearing in these volumes or reprint permission, please contact:

Leanne Battle, Publisher, Full Court Press at leanne.battle@vlex.com or at 202.999.4777

For questions or Sales and Customer Service:

Customer Service
Available 8 a.m.–8 p.m. Eastern Time
866.773.2782 (phone)
support@fastcase.com (email)

Sales
202.999.4777 (phone)
sales@fastcase.com (email)

ISSN 2575-5633 (print)
ISSN 2575-5617 (online)

General Counsel: Is Your Law Firm Using AI? Here's Why That Matters

Matthew H. Grady and Neha Krishna*

Given how ubiquitous artificial intelligence (AI) has become, everyone should have some understanding that AI may generate false information in response to requests. While enormous resources are being devoted to reducing or eliminating these hallucinations, such efforts overshadow other issues endemic in AI. You need to ask and understand how your company's confidential information is being protected when used with AI. Does your own use of AI jeopardize your rights? Do your partners use AI that can jeopardize your rights? Law firms are given access to some of the most sensitive company data. What protections does your firm have in place? If your law firm provides confidential information to an AI model that trains the model on your data, is it still confidential? Recent articles have focused on attorneys' misuse of AI-generated content in court, including fake case citations and fake authority, which has led to disciplinary action. It is important to understand that any use of AI may put company confidential information at risk. There is a need to understand what uses are being made, and what protections are in place to preserve attorney, firm, and client interests.

The advancement of artificial intelligence (AI) is posed to revolutionize industries across the board, and the legal industry as well. The advent of large language models has sparked the widespread use of generative AI. ChatGPT's release in November 2022 has inspired, what some may believe, an overzealous view of the capabilities of generative AI. Specifically, companies and even law firms have increasingly considered and/or adopted various AI tools in the past few years. A survey indicates that 64 percent of law firms use AI-enhanced technologies for legal research, and 47 percent for document review/analysis. Moreover, 77 percent of firms cite increasing efficiency as a primary benefit of the use of AI.[1] However, the integration of AI in legal practices has generated controversy among key stakeholders, including courts, clients, and lawyers. The large uptick in the use of generative AI in the legal industry must be balanced with a critical examination of its limitations, and an accurate assessment of its risks. Several questions

Robotics, Artificial Intelligence & Law / January–February 2025, Vol. 8, No. 1, pp. 19–27.

ISSN 2575-5633 (print) /ISSN 2575-5617 (online).

concerning the quality, legitimacy, and security of AI-generated work must be explored to determine the risk a lawyer, a firm, and its clients are exposed to in utilizing AI.

AI Hallucinations

A primary problem regarding generative AI is its generation of false information in response to requests. This issue is more commonly regarded as an "AI hallucination" where a model generates false, misleading, or illogical information but presents it as if it were a fact.[2] This issue is amplified in large language models (LLMs) that do not understand the underlying semantics of generated words. Instead, these models use probabilistic methods to anticipate the next word to be generated in a sentence. Hallucinations can occur for a variety of reasons, including biased or low-quality training and, more generally, due to current limitations of generative models. In the past two years, several disconcerting instances of AI hallucinations have emerged during court cases, further confirming the limitations of AI.

For example, in the U.S. District Court for the Eastern District of New York, in *Park v. Kim*, a lawyer cited a nonexistent state decision in a reply brief.[3]

Similarly, in the Missouri Eastern District Court of Appeals, only two out of 24 citations (found using AI) within a brief were genuine, and the lawyer was sanctioned $10,000 for gross negligence.[4]

AI hallucinations have reached discussion in the chambers of the U.S. Supreme Court, with Chief Justice John Roberts warning lawyers in his 2023 Year-End Report on the Federal Judiciary of hallucinations and that "any use of AI requires caution and humility."[5] Additionally, as of May 2024, more than 25 federal judges issued standing orders requiring attorneys to disclose the use of AI.[6] The prevalence of AI hallucinations in the courtroom has generated significant legal scrutiny regarding the lack of rigorous checks on the use of AI.

Mitigation Strategies

While hallucinations remain an issue for all users, there has been an effort to uncover strategies designed to mitigate the likelihood of hallucinations in final products.

First, an MIT article encourages users to diversify their sources by cross-checking AI output with other reliable/expert sources or AI-generated content from other models.[7]

Second, several models, including ChatGPT Playground, allow users to experiment with certain hyperparameters, including the "temperature" of the LLM. The temperature of an LLM determines the degree of importance the model places on randomness and creativity over predictability. For example, a higher temperature denotes a diminished emphasis on the probability the next word occurs and a greater emphasis on creative/random outputs.

Finally, using AI platforms catered for legal research can reduce the likelihood of hallucinations. A study found that general purpose LLMs hallucinated 58 to 82 percent of the time on legal queries. AI platforms trained for the unique needs of legal work can produce more quality results due to more refined, higher-quality training data. However, it is important not to overstate the improvements in legal-focused models. According to recent articles, Lexis+ AI produced incorrect information more than 17 percent of the time, and Westlaw's AI-Assisted Research hallucinated more than 34 percent of the time. [8] While implementing the above measures may reduce the likelihood of hallucinations, a need for a "human in the loop" of automation is not eliminated. For in-house counsel, it becomes imperative to know when employed lawyers are using AI to ensure the appropriate safeguards are in place.

Significant Issues

And while hallucinations have become the headliner regarding the shortcomings of AI, there remain significant issues that are frequently overlooked.

First, bias in generative AI (and in AI more broadly) has been a critical and yet neglected topic. A study done on Stable Diffusion—a popular deep learning, text-to-image model—revealed that the model generated images of people with darker skin tones 70 percent of the time for the key phrase "fast food worker" even though 70 percent of fast food workers in the United States are white. There were multiple similar instances in which the model amplified gender and racial stereotypes.[9] Another study found concerning racial patterns of toxicity when assigning ChatGPT to take on the persona of certain individuals.[10] Generative AI can

often overstate certain issues and provide misleading and harmful information.

Second, law firms must have policies in place to ensure lawyers pay close attention to data privacy concerns with the use of generative AI. It behooves all in-house counsel to ensure that their firm and partners have considered and addressed this issue. For example, OpenAI may share users' personal information with unspecified third parties. Moreover, users are unable to request that OpenAI delete specific prompts from storage. According to the Congressional Research Service, proposed legislation is in place for notice and disclosure of changes in privacy policies, opt out of tracking user inputs, and deletion and minimization requirements.[11] However, as of today, there are several privacy hazards regarding the use of AI in law firms.

There are many ways that a law firm or any partner can protect confidential information while continuing to use AI. One recommendation is contract based and includes licensing agreements with an AI provider/platform having strict confidentiality provisions to prevent the platform from knowingly uploading confidential information to be retained in AI models or accessed by unauthorized persons. Any such measures should be thoroughly vetted. Other options include technical architectures where confidential information is used or supplied to internal-only AI models. Sandboxed or secured architectures may be sufficient to protect rights. "Internal" in this context covers the systems and hardware that are wholly under the control of the personnel using them.

Even with internal-only architectures, it is important to keep in mind cybersecurity, especially with the use of cloud-based resources. Ostensibly "internal" models or systems may in fact be hosted in the cloud, making the security of the cloud provider determinative of whether the model is being used on "internal" resources or is even secure at all. The same cloud resources can expose confidential information if not properly secured against attack. In-house counsel should be aware or request information on partners' architecture and agreements. Review any agreements covering use or licensing of AI tools, remain aware of cloud resource issues, and identify your law firms' and any partners' analysis of how the agreement or architecture protects all confidential information.

Furthermore, public models such as ChatGPT use information for reinforcement of the model. User inputs are used to further improve the model for other users with similar requests. The State

Bar of California has specifically cautioned attorneys that they should avoid inputting confidential client information into an AI product that uses inputted information to train its model.[12] Moreover, the New York State Bar Association also issued a warning in safeguarding client confidentiality with the use of generative AI.[13]

To mitigate risks, law firms must consider using platforms that offer private, enterprise-level models with security protocols in place to protect stored information. Moreover, implementing strict internal policies and trainings for the use of AI for employees can further enhance client confidentiality through discussion and transparency. Thus, it becomes necessary for in-house counsel to ensure these measures are in place—or potentially, require that any outside partner or law firm's work not utilize AI tools at all.

Privacy Concerns

It is important to note that any use of AI can produce risk when used with confidential information. For example, there are several data privacy concerns that law firms and any in-house counsel should specifically consider prior to permitting employees to use generative AI.

First, companies should consider the security risks associated with the storage of their inputs and results. Often, companies either do not allow for the deletion of inputs or may unlawfully retain these inputs. For example, with Alexa—a cloud-based voice assistant—Amazon repeatedly assured users that they could delete voice recordings. However, it was later discovered that the company unlawfully retained the data to improve its algorithm, and Amazon was charged by the Department of Justice for violating the Children's Online Privacy Protection Act.[14]

Second, companies should consider the risks associated with a change of privacy policies to allow third-party access or unauthorized internal access. Ring, a manufacturer of home security cameras and alarm systems, failed to restrict employees' and contractors' access to private customers' videos. Furthermore, it used these videos to train algorithms without the consent of consumers. The Federal Trade Commission (FTC) filed the complaint and final order in the U.S. District Court for the District of Columbia.[15] Both internal and third-party access present serious threats to client confidentiality.

Finally, like other digital platforms, companies must consider the associated cybersecurity risks. In another instance at Ring, a bad actor used a breached consumer's account to gain access to stored videos and live video streams of approximately 55,000 customers. These concerns should be thoroughly investigated and addressed through rigid policies and procedures for a law firm. Only when reasonable precautions are in place should counsel permit AI use on their projects.

Due to the above-described concerns, law firms should take necessary measures to protect their employees and clients from breaches of privacy. In addition to the State Bar of California's recommendation, law firms should avoid inputting confidential information into AI platforms without rigorous security measures in place for internal and external access.

Additionally, it is important to keep up to date with the privacy/security policies in place for the AI platforms in use within the firm. Discrete privacy policy changes that are not in line with client confidentiality measures can temporarily expose a company to bad actors. Finally, it is more generally advised that law firms should avoid using open-source and continuous-learning models like ChatGPT. However, if employees are using ChatGPT, there are options to improve privacy measures firmwide.

First, users can opt out of their inputs being used for training purposes.

Second, an Enterprise or API account (paid service options) can improve the security of the stored prompts in ChatGPT. ChatGPT's services for businesses, such as Team, Enterprise, or API, do not use content to train the model. In light of these concerns, companies should adopt stringent measures to safeguard the firm's data privacy and client confidentiality. In-house counsel should ensure that any AI platforms used within law firms meet the strict standards outlined above.

The use of generative AI within the workplace presents both great improvements in efficiency and significant challenges. All practitioners should be cautious of the impact of AI hallucinations. This has played out in the courts in a series of decisions highlighting the likelihood of misleading and fictional outputs, as well as in new specific requirements and rules governing the use of AI. These AI hallucinations and new rules have emphasized the necessity for robust validation mechanisms for AI outputs. In addition to hallucinations, there are several other issues including bias, transparency,

data privacy, and data confidentiality. It is crucial to consider the ethical and legal implications of AI-driven decisions with respect to its shortcomings and the impact those shortcomings can have on your company's rights and obligations. There are several recorded instances of violation of data privacy laws with AI that highlight the urgent need for well-formulated policy and guardrails designed to protect from liabilities arising from the use of AI.

In conclusion, while AI holds a transformative potential for any industry, a balanced approach—including well-defined policy, education, and technical implementation—is needed to ensure confidentiality and accuracy.

Key Takeaways

- Be cautious—the quality, legitimacy, and security of AI-generated work pose risks in any context.
- Outside partners and AI use warrants the same level of supervision as your internal controls.
- Keep in mind that even courts and regulatory bodies advise caution given the pervasiveness of AI and potential pitfalls.
- In-house counsel should ensure that if any partners holding confidential information are using AI, that policies, guidelines, and technical controls are in place and sufficient to mitigate risk.
- Be vigilant as confidentiality may be at risk even when policies and guidelines are in place.
- In-house counsel should review how partners and law firms are protecting confidential information, including technical architectures, agreements, and security provisions.
- Identify, review, and understand any licensing agreements around AI. Confirm your partners' understanding of any confidentiality and use terms.
- Even with external security measures in place, also ensure rigorous internal security measures, including access control, to protect against external threats and internal actors.
- There are specific architectures and model configurations that partners should use to prevent unauthorized access of confidential information. In-house counsel should verify that any tool or platforms provide sufficient safeguards and are configured properly.

Notes

* Matthew H. Grady (matthew.grady@wolfgreenfield.com) is a shareholder in Wolf, Greenfield & Sacks, P.C. Neha Krishna is a technology specialist intern at the firm.

1. The Business Case for AI-Enabled Legal Technology, Thomson Reuters (2021), https://legalsolutions.thomsonreuters.co.uk/content/dam/ewp-m/documents/legal-uk/en/pdf/reports/the-business-case-for-ai-enabled-legal-technology.pdf.

2. Key Legal Issues with Generative AI for Legal Professionals (2024, May 20), https://legal.thomsonreuters.com/blog/generative-ai-for-legal-professionals-top-use-cases/.

3. Park v. Kim, no. 22-2057 (2d Cir. 2024), Justia Law (2024), https://law.justia.com/cases/federal/appellate-courts/ca2/22-2057/22-2057-2024-01-30.html.

4. In the Missouri Court of Appeals Eastern District (2024), https://www.courts.mo.gov/file.jsp?id=205455.

5. J. Roberts, 2023 Year-End Report on the Federal Judiciary (n.d.), https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf.

6. Tracking Federal Judge Orders on Artificial Intelligence, Pulse, Law360 (2024), https://www.law360.com/pulse/ai-tracker.

7. When AI Gets It Wrong: Addressing AI Hallucinations and Bias, MIT Sloan Teaching & Learning Technologies (2024, May 7), https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/.

8. AI on Trial: Legal Models Hallucinate in 1 Out of 6 (or More) Benchmarking Queries, Stanford HAI (n.d.), https://hai.stanford.edu/news/ai-trial-legal-models-hallucinate-1-out-6-or-more-benchmarking-queries.

9. L. Nicoletti & D. Bass, Humans Are Biased. Generative AI Is Even Worse (2023, June 9) https://www.bloomberg.com/graphics/2023-generative-ai-bias/.

10. A. Deshpande et al., Toxicity in ChatGPT: Analyzing Persona-Assigned Language Models, ArXiv (2023).

11. Generative Artificial Intelligence and Data Privacy: A Primer (2023, May 23), https://crsreports.congress.gov/product/pdf/R/R47569.

12. D. Coy, Using AI in Legal Work: COPRAC's Tips on Confidentiality and Competence, The Bar Association of San Francisco (2024, May 24), https://www.sfbar.org/blog/using-ai-in-legal-work-copracs-tips-on-confidentiality-and-competence/.

13. D. Alexander, Protecting Client Confidentiality Remains Paramount at the Outset of the AI Era, New York State Bar Association (2024, Jan. 18), https://nysba.org/protecting-client-confidentiality-remains-paramount-at-the-outset-of-the-ai-era/.

14. FTC and DOJ Charge Amazon with Violating Children's Privacy Law by Keeping Kids' Alexa Voice Recordings Forever and Undermining Parents' Deletion Requests (2023, May 31), https://www.ftc.gov/news-events/news/

press-releases/2023/05/ftc-doj-charge-amazon-violating-childrens-privacy-law-keeping-kids-alexa-voice-recordings-forever.

15. H. Liu & Staff in the Office of Technology, FTC Says Ring Employees Illegally Surveilled Customers, Failed to Stop Hackers from Taking Control of Users' Cameras, Federal Trade Commission (2023, May 31), https://www.ftc.gov/news-events/news/press-releases/2023/05/ftc-says-ring-employees-illegally-surveilled-customers-failed-stop-hackers-taking-control-users#:~:text=The%20Federal%20Trade%20Commission%20charged%20home%20security%20camera,take%20control%20of%20consumers%E2%80%99%20accounts%2C%20cameras%2C%20and%20videos.