# Limits of the Reasonable Inquiry: Artificial Intelligence and the Future of Rule 11

CELESTE GREAVES†

*Artificial intelligence ("AI") burst onto the legal scene and will soon revolutionize practice altogether. Despite enduring technical issues, like hallucinations and implicit bias, AI's capabilities and potential benefits have created massive impetus toward adoption. Recent cases like* Mata v. Avianca *highlight the growing need for adaptation in our existing frameworks to ensure the continued legitimacy of representations to the court. The "reasonable inquiry" standard of Rule 11 of the Federal Rules of Civil Procedure is woefully unequipped to meet this challenge for three primary reasons. First, traditional methods for determining whether a new technology has a high level of efficacy, and can therefore be relied upon, do not apply to AI. Second, when human interactions with AI create negative outcomes or mistakes, it disrupts conventional methods for assigning blame. And third, the "black box" inherent in generative models precludes the retrospective analysis and reasonable inquiry required to analyze the legitimacy of legal reasoning. Finally, even if a reasonable inquiry were possible, there is no legal consensus defining what constitutes "reasonable" in the context of AI. Ultimately, amending Rule 11 is necessary to better tailor the rule to new challenges faced by the judiciary in a rapidly evolving technological landscape.*

† J.D. Candidate 2026, University of California, College of the Law, San Francisco. My sincerest thanks to Professor Scott Dodson for his guidance and thoughtful insight throughout the development of this Note. To Skylar Schoemig, Editor-in-Chief of the *UC Law Journal* and my dear friend, thank you for believing in my work and for all your invaluable feedback. Thank you to Avani Chitre, Executive Production Editor, Madeleine Heldman, Executive Notes Editor, and the Notes and Production Teams for selecting and carefully refining my Note. Finally, to my fiancé, Dimitris, I am profoundly grateful for your devotion, love, and support, and to my parents for their unwavering encouragement and abiding belief in me.

## TABLE OF CONTENTS

## INTRODUCTION

Suppose an associate at a busy law firm, eager to incorporate the promise of emerging technology in her practice, decides to use artificial intelligence ("AI") to draft a motion. She uses AI to generate the argument. She feels confident that the tool is highly intelligent and will create reliable content because of the broad implementation in the legal field and publicity touting its abilities. Additionally, her firm and clients have both impressed upon her the importance of adopting the technology to increase her efficiency and minimize costs. The AI model quickly generates a persuasive motion. She reviews it, double-checks citations, and finds it to be in line with her expectations.

However, when opposing counsel challenges the motion and the court asks the associate to explain how she arrived at the argument presented, she is unable to provide an explanation. While the AI-generated motion appeared sound, the associate did not develop the reasoning herself, and the AI program did not provide a transparent explanation of its analytical process. Unlike traditional legal research, where the associate could have retraced her steps from primary source to conclusion, the "black box"[1] inherent in AI models prevented her from reconstructing the logic. She is left in the difficult position of either retroactively justifying the argument, requiring her to rewrite the motion and potentially rely on a completely different legal theory, or admit that because she relied on generative AI, she cannot independently verify the reasoning.

Threatened with Rule 11 sanctions, the young associate is between a rock and a hard place. Rule 11 assumes that attorneys can verify the accuracy and reliability of their submissions,[2] but traditional methods of verifying the accuracy of the technology do not work for AI. The young associate is unfamiliar with the relevant law and is unable to quickly verify the motion's reasoning. She will be held fully accountable for this failure.

This situation highlights a significant issue: practitioners' increased reliance on AI tools presents a new challenge in balancing the integrity of court filings with attorney accountability and expected due diligence.

Part I explains the historical and technological development of AI as well as key features of its inner workings that have profound effects on its appropriate use in legal practice. This Part also discusses the legal community's goals for AI and persistent technical issues that influence AI's adoption and use by legal professionals.

Part II provides background on the evolution of Rule 11 of the Federal Rules of Civil Procedure. It explains the rule's original rationale, the "reasonable inquiry" standard it establishes, and the role of judicial discretion, setting the stage for later analysis. Additionally, this Part examines how Rule 11 applies to

1. Cole Stryker & Mark Scapicchio, *What Is Generative AI?*, IBM, https://www.ibm.com/think/topics/generative-ai#257779831 (last visited Feb. 25, 2026).

2. FED. R. CIV. P. 11(b).

modern cases where attorneys have faced sanctions for failures to conduct a reasonable inquiry into AI-generated content.

Part III explores why the reasonable inquiry standard is inadequate for adjudicating AI-generated representations in court. First, AI outputs cannot be verified using traditional methods of assessing new technologies. Second, conventional approaches of allocating blame and responsibility in human-machine interactions can lead to unjust outcomes and, in the Rule 11 context, may place an undue burden on attorneys. Third, AI's technological limitations, particularly the black box problem,[3] create novel barriers to reasonable inquiry.

Part IV argues that even if a reasonable inquiry were feasible, the lack of consensus within the legal community on AI's role makes it difficult to establish a uniform standard of reasonableness. Additionally, insufficient education and the rapid evolution of AI technology make it challenging for attorneys and judges to stay informed on AI's capabilities and appropriate uses, further obscuring the standard for reasonableness.

Finally, Part V proposes new language to amend Rule 11 motivated by three key goals: (1) encouraging the use of AI and innovation in the legal field, (2) balancing attorney responsibility with the rapidly evolving legal landscape, and (3) tempering judicial discretion to ensure equitable and uniform application of the federal rules.

## I. AI AND ITS KEY FEATURES

AI is an emerging and transformative technology promising myriad revolutionary applications in the legal field.[4] However, AI's inner workings create opacity in its reasoning, a phenomenon known as the black box.[5] The black box obscures the model's decision-making process and allows for enduring technical issues like hallucinations and bias.[6] Despite these persistent issues, the legal community is rapidly adopting AI to harness its many benefits.[7]

---

3. *See* Stryker & Scapicchio, *supra* note 1.

4. *See CoCounsel Legal*, THOMSON REUTERS, https://legal.thomsonreuters.com/en/products/cocounsel-legal (last visited Feb. 25, 2026).

5. Stryker & Scapicchio, *supra* note 1*; see* Vikas Hassija, Vinay Chamola, Atmesh Mahapatra, Abhinandan Singal, Divyansh Goel, Kaizhu Huang, Simone Scardapane, Indro Spinelli, Mufti Mahmud & Amir Hussain, *Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence*, 16 COGNITIVE COMPUTATION 45, 46 (2024); Lou Blouin, *AI's Mysterious 'Black Box' Problem, Explained*, UNIV. MICH.-DEARBORN NEWS (Mar. 6, 2023), https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained.

6. *See* Matthew R. Caton, *Lawyers: Rely On 'Generative AI' at Your Peril*, 39 ME. BAR J. 48, 48 (2024); Bidyut Sarkar & Rudrendu Kumar Paul, *Ethical AI and Bias Mitigation Solution Strategy in AI Implementation*, 2 INT'L J. A.I. & MACH. LEARNING 88, 90 (2023); Hassija et al., *supra* note 5, at 47; Blouin, *supra* note 5.

7. *See* Press Release, Thomson Reuters, Thomson Reuters Launches CoCounsel Core, Legal GenAI Assistant, in Canada and Australia (Feb. 20, 2024), https://www.thomsonreuters.com/en/press-releases/2024/february/thomson-reuters-launches-cocounsel-core-legal-gen-ai-assistant-in-canada-and-australia ("[T]he first-ever legal AI assistant has been rolled out to more than 45 large law firms in the United States—including six of the Am Law 10—which employ in total more than 50,000 lawyers."); THOMSON REUTERS, *supra* note 4; Press Release, Thomson Reuters, Thomson Reuters Unveils Deeper Integration of CoCounsel 2.0 in Westlaw and Practical Law, Building on Its GenAI Assistant for Legal Professionals (Oct. 16, 2026),

### A. THE HISTORY AND DEVELOPMENT OF AI

AI has a dynamic history and development[8] that began long before most people became aware of the technology.[9] While the concept of AI was introduced as early as 1950 with Alan Turing's *Computing Machinery and Intelligence*,[10] the term "AI" was not coined until 1955 by John McCarthy.[11] By the late 1960s, AI theorist Marvin Minsky expressed that the purpose of the technology was to create machines capable of doing tasks that would require "intelligence if done by man."[12] AI development was relatively slow compared to similar technological advancements like the internet, and has ebbed and flowed, boomed and stagnated over the years.[13] The most recent explosion of development began in 2012 when Google developers, Jeff Dean and Andrew Ng, trained a neural network to recognize cats in unlabeled images.[14] Later in 2017, Facebook (now Meta) programmed chatbots to converse with each other, and over time, the models developed their own language.[15]

Since the origination of large language models ("LLMs"),[16] otherwise known as chatbots, the use cases for AI have rapidly expanded, causing it to become increasingly ubiquitous in everyday life.[17] Common, low-stakes applications, including virtual assistants,[18] targeted advertising,[19] gaming,[20] text

https://www.thomsonreuters.com/en/press-releases/2024/october/thomson-reuters-unveils-deeper-integration-of-cocounsel-2-0-in-westlaw-and-practical-law-building-on-its-genai-assistant-for-legal-professionals.html.

8. *See What Is the History of Artificial Intelligence (AI)?*, TABLEAU, https://www.tableau.com/data-insights/ai/history (last visited Feb. 26, 2026); *see, e.g.*, A. M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433, 433–36 (1950).

9. *See* Michelle Faverio & Alec Tyson, *What the Data Says About Americans' Views of Artificial Intelligence*, PEW RSCH. CTR. (Nov. 21, 2023), https://www.pewresearch.org/short-reads/2023/11/21/what-the-data-says-about-americans-views-of-artificial-intelligence (explaining that in 2023, only 30% of adults recognized all six examples of artificial intelligence and 42% had heard nothing about ChatGPT at all).

10. *See* Turing, *supra* note 8, at 433–34; Daniel Ben-Ari, Yael Frish, Adam Lazovski, Uriel Eldan & Dov Greenbaum, *"Danger, Will Robinson"? Artificial Intelligence in the Practice of Law: An Analysis and Proof of Concept Experiment*, RICH. J.L. & TECH., Mar. 15, 2017, at 1, 15.

11. *Id.* at 19.

12. Edwina L. Rissland, Comment, *Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*, 99 YALE L.J. 1957, 1958 (1990) (quoting SEMANTIC INFORMATION PROCESSING v (Marvin Minsky ed., 1968)).

13. *See* TABLEAU, *supra* note 8.

14. *Id.*

15. *Id.*

16. *See* Cole Stryker, *What Are Large Language Models (LLMs)?*, IBM, https://www.ibm.com/think/topics/large-language-models (last visited Feb. 27, 2026).

17. *See* Kelly Raskovich, *AI Everywhere: Like Magic, but with Algorithms*, DELOITTE INSIGHTS (Dec. 11, 2024), https://www.deloitte.com/us/en/insights/topics/technology-management/tech-trends/2025/tech-trends-introduction-ai-is-everywhere.html.

18. *Assistants Migration Guide*, OPENAI, https://platform.openai.com/docs/assistants/migration (last visited Mar. 1, 2025).

19. *See generally* Mike Kaput, *AI in Advertising: Everything You Need To Know*, MKTG. A.I. INST. (Jan. 22, 2024), https://www.marketingaiinstitute.com/blog/ai-in-advertising (explaining AI integration in modern advertising).

20. *See, e.g.*, Harsh Panwar, The NPC AI of *The Last of Us*: A Case Study 2–3 (July 29, 2022) (unpublished manuscript), https://arxiv.org/pdf/2207.00682.

editing,[21] and more,[22] are already widely adopted.[23] The general public seems comfortable leaving low-stakes decisions to AI while retaining high-stakes decisions for humans.[24] However, that has not inhibited the integration of AI with consequential practice areas like medicine,[25] education,[26] criminal sentencing,[27] road safety,[28] and national security.[29]

### B. AI'S INNER WORKINGS AND THE BLACK BOX

AI systems typically emerge through machine learning, a broad category of AI training that involves pre-assigning weights to a model and feeding it massive amounts of data.[30] This data is often organized into datasets, which can be categorized and selected to tailor a model to be especially effective for a particular purpose.[31] By processing the data through its neural network, the model is able to learn and refine its outputs.[32] Neural networks can be understood to process information through layers of code in one direction, with each layer containing thousands or even millions of nodes.[33] Each node is assigned a weight, which it multiplies with every input.[34] If the multiplication meets a predetermined threshold, the node "fires" like a neuron and passes the

---

21. *See* Coco Poley, *How AI Text Editors Improve Content Creation*, TINY BLOG (Sep. 11, 2024), https://www.tiny.cloud/blog/how-ai-text-editors-improve-content-creation.

22. *See* Nikita Sachdeva, *20 Uses of Artificial Intelligence in Day-to-Day Life*, DAFFODIL: CURATED ENG'G INSIGHTS (Jan. 16, 2024, at 18:30 PT), https://insights.daffodilsw.com/blog/20-uses-of-artificial-intelligence-in-day-to-day-life.

23. *See* Brian Kennedy, Alec Tyson & Emily Saks, *Public Awareness of Artificial Intelligence in Everyday Activities*, PEW RSCH. CTR. (Feb. 15, 2023), https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities (explaining that a study showed 44% of people surveyed in 2023 did not think they regularly interacted with artificial intelligence).

24. *See Study Finds People Trust AI More for Low-Stakes Decisions*, AZO ROBOTICS (Feb. 18, 2025), https://www.azorobotics.com/News.aspx?newsID=15727.

25. *See* Daniel A. Hashimoto, Guy Rosman, Daniela Rus & Ozanan R. Meireles, *Artificial Intelligence in Surgery: Promises and Perils*, 268 ANN. SURG. 70, 73 (2018) ("Clinician-machine interaction has already been demonstrated to augment decision-making").

26. *See* Carrie Spector, *How Technology Is Reinventing Education*, STAN. REP. (Feb. 14, 2024), https://news.stanford.edu/stories/2024/02/technology-in-education.

27. For a discussion about artificial intelligence being used to predict recidivism, see Allyson Brunette, *Humanizing Justice: The Transformational Impact of AI in Courts, From Filing to Sentencing*, THOMSON REUTERS (Oct. 25, 2024), https://www.thomsonreuters.com/en-us/posts/ai-in-courts/humanizing-justice.

28. *See* Dimitrios I. Tselentis, Eleonora Papadimitriou & Pieter van Gelder, *The Usefulness of Artificial Intelligence for Safety Assessment of Different Transport Modes*, ACCIDENT ANALYSIS & PREVENTION, Mar. 28, 2023, at 1.

29. *See Artificial Intelligence Security Center*, NAT'L SEC. AGENCY, https://www.nsa.gov/AISC (last visited Feb. 28, 2026).

30. Dave Bergmann, *What Is Machine Learning?*, IBM, https://www.ibm.com/topics/machine-learning (last visited Feb. 28, 2026).

31. *See Finding the Best Training Data for Your AI Model*, KEYLABS (Jan. 29, 2024), https://keylabs.ai/blog/finding-the-best-training-data-for-your-ai-model/.

32. Bergmann, *supra* note 30.

33. *See What Is Deep Learning?*, IBM, https://www.ibm.com/think/topics/deep-learning (last visited Feb. 28, 2026); FangFang Lee, *What Is a Neural Network*, IBM, https://www.ibm.com/think/topics/neural-networks (last visited Feb. 28, 2026).

34. Lee, *supra* note 33.

value forward.[35] By the time the model generates an output based on the "firing" of these "neurons," developers and users can no longer trace how it arrived at its conclusion.[36] Developers have identified this phenomenon as the black box, which obscures our ability to determine the reasoning of models that use this type of processing.[37] The most common models that operate in this fashion are LLMs like ChatGPT, Llama-4, etc.[38]

New AI models are being trained using reinforcement learning rather than traditional machine learning, which may lessen the opacity of the black box.[39] Reinforcement learning is a more evolved method of model training where, instead of the model's neurons firing based on predetermined weights, the model adapts its weights in real time to optimize for "rewards."[40] Models trained using reinforcement learning are beginning to be able to "think aloud" or create a chain of thought, which may help in making the black box more translucent.[41] However, thousands—if not millions—of nodes are involved in the neural network's decision-making process.[42] So, even if some chains of thought are produced and viewable by the developer or user, significant unknowns remain both with respect to the unrevealed chains of thought and within the meta-reasoning of the model.[43] Meta-reasoning involving the synthesis of numerous logical chains of thought would still be shrouded in mystery.[44]

Like a child, AI learns by recognizing patterns and categorizing experiences until it can effectively respond to new situations.[45] However, just as a child may struggle to explain why they know something, AI systems also lack transparency in how they reach conclusions.[46] AI's reasoning is embedded in complex layers of learned associations, rather than clear, traceable logic. This phenomenon poses substantial issues when models are used to produce novel

---

35. *See id.*

36. *See id.*; Stryker & Scapicchio, *supra* note 1.

37. Stryker & Scapicchio, *supra* note 1; *see* Hassija et al., *supra* note 5, at 47.

38. Matthew Kosinski, *What Is Black Box Artificial Intelligence (AI)?*, https://www.ibm.com/think/topics/black-box-ai (last visited Mar. 1, 2026).

39. *See* Sampsa Samila, *Chain of Thought Reasoning, the New LLM Breakthrough*, IESE BUS. SCH.: A.I. & FUTURE MGMT. BLOG NETWORK (Oct. 22, 2024), https://blog.iese.edu/artificial-intelligence-management/2024/chain-of-thought-reasoning-the-new-llm-breakthrough.

40. *What Is an AI Model?*, IBM, https://www.ibm.com/think/topics/ai-model (last visited Mar. 1, 2026).

41. ChatGPT-4 C-LARA-Instance & Manny Rayner, Reinforcement Learning for Chain of Thought Reasoning: A Case Study Using Tic-Tac-Toe 3 (July 2024) (unpublished manuscript), https://www.researchgate.net/publication/382496965_Reinforcement_Learning_for_Chain_of_Thought_Reasoning_A_Case_Study_Using_Tic-Tac-Toe.

42. *See supra* note 33.

43. Lance Eliot, *Speculation Brims That OpenAI O1 Leverages Multiple Chain-Of-Thought and Meta-Reasoning*, FORBES (Sep. 19, 2024, at 04:09 ET), https://www.forbes.com/sites/lanceeliot/2024/09/19/speculation-brims-that-openai-o1-leverages-multiple-chain-of-thought-and-meta-reasoning.

44. *Id.*

45. Blouin, *supra* note 5.

46. *Id.*

legal arguments, as discussed later,[47] because the deeper reasoning and presumptions on which the model builds its argument are unknown to the user.

### C. MODERN CAPABILITIES AND LEGAL APPLICATIONS OF LLMS

Non-specific LLMs, like ChatGPT-4, are trained on diverse datasets encompassing thousands of subjects, areas of study, and various corners of the internet.[48] These models perform shockingly well on standard legal tasks, such as articulating the basis for a claim, identifying required elements for a claim, providing legal strategies, writing cross-examinations, drafting arguments based on case and statutory law, and determining a likely judgment based on its own arguments for each side of the dispute.[49] OpenAI launched its "o1" model in September 2024, which already significantly outperforms OpenAI's previous model, ChatGPT-4, in numerous areas.[50] The o1 model particularly excels in tasks requiring intricate reasoning and knowledge integration across various fields.[51] However, experts testing the software have found that the program still makes errors on simple problems and struggles with highly specialized concepts.[52]

Applied to the legal field, the o1 model is not limited to research. It is touted as being able to "'think' through" problems in the same way a human would, identify nuances in judicial reasoning, and offer suggestions that human researchers may overlook.[53] Some models have been trained on legal data, leading to the development of legal models like CoCounsel.[54] These models are developed and deployed[55] to assist attorneys in tasks such as deposition preparation, document review, research, and identifying key information, allowing lawyers more time to "focus on higher-value" human work like strategy and client relationships.[56] CoCounsel has even been incorporated as an embedded software in preeminent legal research databases like Westlaw and Westlaw's Practical Law program.[57] Integration of AI with traditional legal

47. *See infra* Subpart.III.C.

48. *See* LEE TIEDRICH, KARINE PERSET & SARA FIALHO ESPOSITO, OECD, ARTIFICIAL INTELLIGENCE PAPERS NO.33, INTELLECTUAL PROPERTY ISSUES IN ARTIFICIAL INTELLIGENCE TRAINED ON SCRAPED DATA 14–15, 19 (2025).

49. Kwan Yuen Iu & Vanessa Man-Yi Wong, ChatGPT by OpenAI: The End of Litigation Lawyers? 15–17 (Feb. 2, 2023) (unpublished manuscript), https://www.ssrn.com/abstract=4339839 (generated by GPT-3.5).

50. *See OpenAI o1 System Card*, OPENAI (Dec. 5, 2024), https://openai.com/index/openai-o1-system-card.

51. Tianyang Zhong et al., Evaluation of OpenAI o1: Opportunities and Challenges of AGI 94 (Nov. 24, 2025) (unpublished manuscript), https://arxiv.org/abs/2409.18486.

52. *Id.* at 23.

53. Kevin P. Lee, LINKEDIN, *The Transformative Impact of the OpenAI o1 Model on Legal Practice* (Sep. 16, 2024), https://www.linkedin.com/pulse/transformative-impact-openai-o1-model-legal-practice-kevin-p-lee--aclxe.

54. THOMSON REUTERS, *supra* note 4.

55. *See* Thomson Reuters Launches CoCounsel Core, Legal GenAI Assistant in Canada and Australia, *supra* note 7.

56. THOMSON REUTERS, *supra* note 4.

57. Thomson Reuters Unveils Deeper Integration of CoCounsel 2.0 in Westlaw and Practical Law, Building on Its GenAI Assistant for Legal Professionals, *supra* note 7.

databases enables high level capabilities like identifying mischaracterizations of the law because, as discussed previously,[58] AI models provided with large swaths of high quality, highly specific data become precise and effective in that field.[59] For example, integration with traditional legal databases increases the high quality case law the model can access which allows it to gain the ability to interpret nuanced judicial holdings more effectively.[60] While the future scope of AI's role remains unclear, a recent survey from August 2025 showed that of the legal professionals who responded, eighty percent believed AI would have a highly transformative impact on their work in the next five years, and half of the responding firms said exploring and implementing AI was their highest priority.[61]

### D. Enduring Technical Issues: Hallucinations and Bias

Alongside the leaps and bounds AI has made toward the ultimate goal of a human-like general intelligence, significant limitations persist. One of the more notorious issues is a phenomenon called "hallucination," where an AI model outputs information that is inaccurate, or even entirely fabricated.[62] This technical issue is already pervading the legal landscape, illustrated by a string of cases where judges issued sanctions against attorneys who submitted hallucinated cases to the court.[63]

The general recommendation to mitigate the harms of AI hallucinations is to double-check all AI-generated outputs.[64] While this may seem intuitive and workable in the context of legal research, double-checking is impractical when applied to legal reasoning. Unlike verifying citations, which can be easily checked against traditional databases, attorneys cannot easily verify the legitimacy of a novel argument, particularly when it occurs within the black box. As discussed later, this creates a tension between an attorney's obligations under Rule 11 and the increasingly ubiquitous use of AI in legal practice.[65]

Additionally, AI models are trained using datasets derived from human activity over time, so the data reflects historical biases.[66] In turn, the historical biases of our world are transcribed into the model's reasoning.[67] When models adopt these biases, they may apply faulty reasoning based on bias that can

58. *See supra* Subpart.I.B.

59. *See* KEYLABS, *supra* note 31.

60. Thomson Reuters Unveils Deeper Integration of CoCounsel 2.0 in Westlaw and Practical Law, Building on Its GenAI Assistant for Legal Professionals, *supra* note 7.

61. Marjorie Richter, *How AI Is Transforming the Legal Profession*, THOMSON REUTERS (Aug. 18, 2025), https://legal.thomsonreuters.com/blog/how-ai-is-transforming-the-legal-profession.

62. Caton, *supra* note 6.

63. *See* discussion *infra* Subpart.II.B; Mata v. Avianca, Inc., 678 F. Supp. 3d 443, 449 (S.D.N.Y 2023).

64. Caton, *supra* note 6, at 50.

65. *See infra* Subpart.III.C.

66. Sarkar & Paul, *supra* note 6, at 89–90.

67. *Id.* at 90.

perpetuate negative social stereotypes and have disparate impacts.[68] In the legal context, these biases can raise constitutional issues if the AI system relies on unconstitutional presumptions prohibited by the Equal Protection Clause.[69] To mitigate these harms, developers can alter the training data by removing common patterns of bias, but these strategies are not usually practicable due to the size of datasets.[70] Moreover, anti-bias interventions can distort the model in unexpected ways, making it less reliable or stripping it of context so that it no longer creates outputs that comport with reality.[71] As explored later, bias and potentially problematic presumptions pose risks to attorneys using the model because the black box prohibits attorneys from determining when bias may be at play.[72]

The enduring technical issues of hallucination and bias underscore some of the core challenges attorneys face when adopting AI in their practice. This tension is further explored later in the discussion regarding decision-making in the black box and its intersection with an attorney's obligations under Rule 11.[73]

## II. THE HISTORY OF RULE 11 AND MODERN-DAY APPLICATIONS TO AI-GENERATED CONTENT

Rule 11 was originally intended to deter frivolous lawsuits by ensuring that attorneys file with the court in good faith.[74] Over the years, through several amendments, the Rule has evolved to include a "reasonable inquiry" requirement, mandating attorneys conduct a reasonable inquiry beyond the four corners of their filings to determine the plausibility and validity of the assertions

68. *Id.* at 89. For an example of how bias in AI models can have disparate impacts, consider when Amazon scrapped its AI hiring tool after discovering it discriminated against female candidates. The model was trained on past resumes, most of which came from men, due to the technology industry's male dominance. *See* Jeffrey Dastin, *Insight – Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women*, REUTERS (Oct. 10, 2018, at 17:50 PT), https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG. As a result, the AI learned to favor male candidates, penalizing resumes that included words like "women's," such as "women's chess club," and downgrading candidates who attended all-women's colleges. *Id.*

69. *See infra* Subpart.III.A.

70. *See* Sarkar & Paul, *supra* note 6.

71. For example, ImageNet attempted to remove offensive or inappropriate images from AI training datasets; however, this caused concern that the model would fail to be representative of the real world. *See* Will Knight, *AI Is Biased. Here's How Scientists Are Trying to Fix It*, WIRED (Dec. 19, 2019, at 12:34 PT), https://www.wired.com/story/ai-biased-how-scientists-trying-fix. Additionally, in 2015, Google's image recognition misclassified Black individuals as gorillas, so Google blocked the model from identifying primates. This solved the immediate issue but reduced the overall accuracy of the model's representations of reality. *See* James Thorn, *Bias in Artificial Intelligence*, TOWARDS DATA SCI. (July 14, 2020), https://towardsdatascience.com/bias-in-artificial-intelligence-a3239ce316c9.

72. *See infra* Subpart.III.C.

73. *Id.*

74. Robert L. Carter, *The History and Purposes of Rule 11*, 54 FORDHAM L. REV. 4, 4 (1985).

therein.[75] Judicial discretion plays a key role in determining whether that standard has been met and whether to issue sanctions as a result.[76]

In the context of AI-generated content, there is already confusion about how this standard applies. So far, its application has been limited to circumstances like hallucinated cases,[77] while broader applications, such as its impact on AI-generated legal reasoning, remain unexplored.

### A. The History and Rationale of Rule 11 and the Reasonable Inquiry Standard

Rule 11 is at the core of the legal profession's interest in preserving the legitimacy, decorum, and efficacy of our judicial system. The original goal of the rule was to prevent frivolous lawsuits by requiring an attorney's signature, ensuring the filing was in "good faith," but this did not involve any responsibility of inquiry on behalf of the attorney.[78] Additionally, the rule included a "pure heart, empty head" defense that allowed litigants to avoid sanctions when they acted in good faith but failed to conduct sufficient research.[79]

The Rule was amended in 1983 to include, among other things, the reasonable inquiry standard requiring attorneys to look beyond the four corners of their filing to determine the validity and truth of the assertions.[80] The duty to investigate gave the rule teeth and replaced the prior requirement for only good faith belief.[81] At that time, what constituted reasonable inquiry under the circumstances was based on the time allocated for investigation, whether only their client had the information, and whether the claims were legally plausible.[82] Another notable change was the removal of judicial discretion. The amendment established that once a failure to reasonably inquire was established, sanctions were mandatory, not optional.[83]

Later, the 1993 amendment to Rule 11 re-implemented judicial discretion to determine whether to impose sanctions when an attorney failed to conduct a reasonable inquiry.[84] And, while the amendment introduced other changes like establishing a safe harbor provision, a deterrence-based approach to sanctions,

75. *Id.* at 4–5, 7.

76. Kelly D. Wheaton, *The Amended Rule 11 Sanctions: New and Improved, or Just New?*, ARMY LAW., Sep. 1994, at 23, 29.

77. *See, e.g.*, Mata v. Avianca, Inc., 678 F. Supp. 3d 443, 449 (S.D.N.Y. 2023); Park v. Kim, 91 F.4th 610, 615 (2d Cir. 2024).

78. Carter, *supra* note 74, at 4, 6.

79. Jessica R. Gunder, *Rule 11 Is No Match for Generative AI*, 27 STAN. TECH. L. REV. 308, 323 (2024) (quoting Marguerite L. Butler, *Rule 11—Sanctions and a Lawyer's Failure to Conduct Competent Legal Research*, 29 CAP. U.L. REV. 681, 688 (2001)).

80. Carter, *supra* note 74, at 4–5.

81. *Id.* at 4, 6–7.

82. Wheaton, *supra* note 76, at 25.

83. *Id.* at 24.

84. *Id.* at 26, 29.

and limits on monetary sanctions, the reasonable inquiry standard remained unchanged.[85]

When thinking about AI-generated attorney work-product, it is important to distinguish between the standards for filings challenged by opposing counsel via motion versus challenges raised sua sponte by the judge. In the former, when the opposing party files a motion for sanctions, the judge may issue sanctions if counsel failed to carry out a reasonable inquiry, despite not having acted in bad faith or with any contempt.[86] A claim by a litigant is objectively unreasonable if "it is patently clear that [the] claim has absolutely no chance of success under the existing precedents, and where no reasonable argument can be advanced to extend, modify, or reverse the law as it stands."[87] In the latter, when the judge raises a challenge sua sponte, it is widely accepted that there must be extra care and sanctions should only be assigned when counsel's conduct is "akin to [ ] contempt," especially since the safe harbor provision is unavailable.[88] The bad faith requirement for sua sponte Rule 11 challenges is a higher standard aimed not just at negligence but at more egregious attorney conduct.[89]

This distinction is pertinent because sua sponte challenges are subject to the bad faith akin to contempt standard rather than the lower reasonable inquiry standard.[90] As a result, not all sanctions rulings examine whether a reasonable inquiry was conducted.[91] This is especially relevant in the next Part, as all the existing case law on judicial discretion for imposing sanctions for AI-generated content involves sua sponte challenges.[92] This leaves the application of the reasonable inquiry standard to any type of AI-generated content uncertain and untested.

## B. MODERN DAY APPLICATIONS OF RULE 11 TO THE USE OF AI BY ATTORNEYS

The first application of Rule 11 to an attorney's failure to conduct a reasonable inquiry into AI-generated court filings occurred in 2023 in *Mata v. Avianca*.[93] In that case, an attorney filed a brief containing AI-hallucinated cases with the court without verifying their legitimacy.[94] The attorney explained that

---

85. *Id.* at 28–30.

86. Gunder, *supra* note 79, at 329.

87. Adams v. N.Y. State Educ. Dep't., No. 08 Civ. 5996, 2010 WL 4970011, at *7 (S.D.N.Y. Dec. 8, 2010) (quoting Catcove Corp. v. Heaney, 685 F. Supp. 2d 328, 337 (E.D.N.Y. 2010) (alteration in original)).

88. *In re* Pennie & Edmonds LLP, 323 F.3d 86, 90 (2d Cir. 2003); *see also* Hunter v. Earthgrains Co. Bakery, 281 F.3d 144, 151 (4th Cir. 2002); Barber v. Miller, 146 F.3d 707, 711 (9th Cir. 1998); *In re* Bees, 562 F.3d 284, 287 (4th Cir. 2009).

89. *In re Pennie & Edmonds*, 323 F.3d at 89–90.

90. *See* Gunder, *supra* note 79, at 329; *see also In re Pennie & Edmonds*, 323 F.3d at 90; *Hunter*, 281 F.3d at 151; *Barber*, 146 F.3d at 711.

91. *See*, *e.g.*, *In re Pennie & Edmonds*, 323 F.3d at 90; *Hunter*, 281 F.3d at 151; *Barber*, 146 F.3d at 711.

92. *See*, *e.g.*, Mata v. Avianca, Inc., 678 F. Supp. 3d 443, 462 (S.D.N.Y. 2023); United States v. Cohen, 724 F. Supp. 3d 251, 258 (S.D.N.Y. 2024); Park v. Kim, 91 F.4th 610, 613 (2d Cir. 2024).

93. 678 F. Supp. 3d at 448.

94. *See id.* at 451.

he had mistakenly believed that LLMs are like a "super search engine," rather than a generative technology.[95] The judge scheduled a show-cause hearing and ultimately imposed sanctions, not for the attorney's failure to verify the cases, but for the attorney's bad faith conduct, including lying to the court in an attempt to avoid the hearing and claiming that the cases were not fabricated.[96] Because the challenge was raised sua sponte, the court did not need to carry out a reasonable inquiry analysis to determine whether sanctions applied because the judge found the higher standard of bad faith to be met.[97]

Shortly thereafter, the situation reoccurred in *United States v. Cohen*, where counsel for Michael Cohen submitted a letter to the court that contained three fictitious, hallucinated cases.[98] The judge threatened sanctions sua sponte, and again, counsel admitted ignorance of the technology he was using, stating that he did not realize Google Bard was a generative text service, and that he thought it was a "super-charged search engine."[99] Despite the error, the court cited the requirement for bad faith and found that the attorney had not met that standard, so sanctions were not issued.[100] The issue was addressed once again without the judge having to apply the reasonable inquiry standard due to the challenge being raised sua sponte.

Later, the Second Circuit clarified the standard inquiry regarding AI-generated case law in *Park v. Kim*.[101] Similar to *Mata* and *Cohen*, an attorney submitted cases hallucinated by ChatGPT to the court.[102] The district judge raised a challenge sua sponte and imposed sanctions.[103] On appeal, the Second Circuit clarified that failing to check research was in violation of the adversarial system and constituted an attempt to get the court to rely on fake opinions.[104] The court reasoned that failing to verify the legitimacy of case law reveals an inability to determine whether the legal argument made is legally tenable.[105] Presumably, the Second Circuit's holding extends the application of Rule 11 beyond AI-generated legal research—to legal reasoning produced by the model. However, the court did not elaborate on how the facts could be assessed under the reasonable inquiry standard and does not conduct further analysis regarding what would be reasonable.[106]

95. *Id.* at 456.
96. *Id.* at 452, 457–58.
97. *Id*. at 462–63.
98. No. 18-CR-602, 2023 WL 8635521, at *1 (S.D.N.Y. Dec. 12, 2023).
99. Letter from E. Danya Perry to Hon. Jesse M. Furman, at 3, United States v. Cohen, 724 F. Supp. 3d 251 (S.D.N.Y. 2023) (No. 18-CR-602).
100. *Cohen*, 724 F. Supp. 3d at 258 (citing Muhammad v. Walmart Stores E., L.P., 732 F.3d 104, 108 (2d Cir. 2013)).
101. 91 F.4th 610, 615 (2d Cir. 2024); Caton, *supra* note 6.
102. *See* Caton, *supra* note 6, at 48–49.
103. *Park*, 91 F.4th at 612.
104. *Id.* at 615.
105. *Id*.
106. *See id.* at 611, 614–15.

Sanctions have been issued on other occasions, but only insofar as LLMs provide hallucinated cases that the attorneys adopt as legitimate research and provide the to court.[107] There has yet to be a case where AI-generated legal reasoning has been submitted to the court, and the judge must conduct a reasonable inquiry analysis into the underlying reasoning of the legal argument, not based solely on the existence of case law.

## III. RULE 11'S REASONABLE INQUIRY STANDARD IS INCOMPATIBLE WITH ATTORNEY'S USE OF AI FOR LEGAL REASONING

The reasonable inquiry standard is incompatible with the emerging and ubiquitous adoption of AI in legal practice[108] for three key reasons.

First, unlike previous technologies that underwent an adoption phase involving parallel verification,[109] AI cannot be validated in the same way without frustrating its purpose. As a result, what constitutes reasonable reliance or use of AI remains unclear and unprecedented. Second, traditional methods of assigning blame for harm caused by human-machine interactions[110] may create disparate impacts for attorneys using AI. The inclination to hold attorneys solely responsible for AI failures may not always be reasonable or equitable. Third, the black box obscures AI's reasoning, causing major issues for interpretability.[111] This makes it difficult to assess whether enduring technical issues, like bias, are influencing legal arguments in impermissible ways.[112] Even developers have only limited interpretability of AI decision-making.[113] Thus, attorneys cannot reliably scrutinize AI outputs without significantly diminishing the technology's usefulness.

---

107. *See id.* at 615; Mata v. Avianca, Inc., 678 F. Supp. 3d 443, 466 (S.D.N.Y. 2023); *see also, e.g.*, Couvrette v. Wisnovsky, No. 1:21-CV-00157, 2024 WL 5126286, *15–17 (D. Or. Dec. 16 2026); Daniel Gentry v. Calvin Thompson, No. CV 25-1260, 2026 WL 787563, *2–3 (E.D. La. Mar. 20, 2026). For a catalog of cases where attorneys have faced Rule 11 sanctions and other penalties for AI related errors, see *AI Hallucination Cases*, DAMIEN CHARLOTIN, https://www.damiencharlotin.com/hallucinations/?q=&sort_by=date&states=USA&period_idx=0&nature_categories=fabricated (last visited Mar. 26, 2026).

108. *See*, *e.g.*, *AI Adoption in Legal Sector Soars, Study Finds*, LEGAL.IO (Oct. 10, 2024), https://www.legal.io/articles/5543376/AI-Adoption-in-Legal-Sector-Soars-Study-Finds; Andrew Arruda, *An Ethical Obligation to Use Artificial Intelligence? An Examination of the Use of Artificial Intelligence in Law and the Model Rules of Professional Responsibility*, 40 AM. J. TRIAL ADVOC. 443, 458 (2017); *ABA Task Force on Law and Artificial Intelligence Releases Survey on AI and Legal Education*, A.B.A. (June 24, 2024), https://www.americanbar.org/news/abanews/aba-news-archives/2024/06/aba-task-force-law-and-ai-survey.

109. *See* Martin E. Hsia, *From Typewriters to Artificial Intelligence: Issues in Everyday Legal Practice as Technology Evolved Over the Last 40 Years*, HAW. BAR J., Sep. 2024, at 10.

110. *See* Madeleine Clare Elish, *Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction*, 5 ENGAGING SCI., TECH., & SOC'Y 40, 41–42 (2019).

111. *See* Stryker & Scapicchio, *supra* note 1; Hassija et al., *supra* note 5, at 45–46; H. James Wilson & Paul R. Daugherty, *Collaborative Intelligence: Humans and AI Are Joining Forces*, HARV. BUS. REV., July–Aug. 2018, at 115, 117; Ananya, *AI's Understanding and Reasoning Skills Can't Be Assessed by Current Tests*, SCI. NEWS (July 10, 2024, at 09:00 PT), https://www.sciencenews.org/article/ai-understanding-reasoning-skill-assess.

112. *See* Wilson & Daugherty, *supra* note 111, at 118.

113. Hassija et al., *supra* note 5.

### A. AI DEFIES TRADITIONAL MEHODS OF VERIFICATION FOR EMERGING TECHNOLOGIES

The development of technology for the purpose of diminishing or changing the types of work that require human intelligence is not a new phenomenon.[114] From calculators to rudimentary computers, all technologies go through an adoption phase where they are used and their outputs checked for accuracy.[115] Generally, this involves a human predicting the correct output for a given problem and then comparing that prediction to the output created by the new technology.[116] This practice of "parallel problem-solving"[117] allows the user to check the accuracy of the technology. If the outputs do not match, the program can be modified to increase accuracy.[118] Parallel problem solving is ideal for testing technologies and inquiries with quantifiable or objective answers, such as solving equations.[119] However, legal reasoning differs in three key ways and prevents parallel problem-solving from reliably verifying AI-generated outputs in the legal context.

First, parallel problem solving undermines the purpose of using AI because it requires the user to already know the expected solution. For attorneys to assess AI-generated legal analysis, they must independently conduct all necessary research and reasoning. Since AI is primarily used to save time, increase efficiency, and allow attorneys to focus their time on meaningful, client-oriented tasks,[120] requiring full legal analysis before verification negates its benefits. Even if experienced lawyers can predict likely outcomes and check AI-generated results, novice attorneys often lack the necessary background knowledge needed to efficiently verify the outputs.

Second, legal reasoning often does not yield a single, quantifiable solution. Even if an attorney conducts independent legal analysis to verify an AI-generated output, the model's response may differ substantially. The model might follow the same legal route and reach a different conclusion or take an entirely different path and arrive at the same conclusion. Given that multiple legitimate legal theories can apply to a single issue, parallel problem-solving

114. Rissland, *supra* note 12, at 1959; *see What Is Artificial General Intelligence (AGI)?*, IBM, https://www.ibm.com/think/topics/artificial-general-intelligence (last visited Mar. 6, 2026); *Artificial General Intelligence, the Holy Grail of AI*, ENTEFY (June 14, 2022), https://www.entefy.com/blog/artificial-general-intelligence-the-holy-grail-of-ai.

115. Hsia, *supra* note 109, at 11.

116. *Id*. at 10.

117. "Parallel problem-solving" is a term I will use to describe the practice of solving problems alongside a machine to prove or disprove its accuracy. This term does not exist in the literature to which I cite but, rather, will be used for efficiency in describing the practice.

118. Hsia, *supra* note 109, at 11.

119. *Id.*

120. *See 7 Ways Artificial Intelligence Can Benefit Your Law Firm*, A.B.A. (Sep. 2017), https://www.americanbar.org/news/abanews/publications/youraba/2017/september-2017/7-ways-artificial-intelligence-can-benefit-your-law-firm.

may be ineffective, as variations in reasoning and conclusions do not necessarily indicate errors.

Third, AI models can introduce bias into legal reasoning without the user recognizing it because the conclusion may appear legitimate. Legal analysis hinges on the reasoning behind a conclusion rather than the conclusion itself. If an AI model inherits biases from its training data and embeds them in its reasoning, attorneys may be unable to detect or correct them, as the decision-making process occurs within a black box.[121] This creates a risk that AI-generated conclusions rest on unconstitutional or flawed presumptions without the user's knowledge. Simply verifying that an answer seems plausible does not ensure the model reached it through sound legal reasoning. The potential for pervasive bias in AI models is a unique challenge of the technology that complicates their use in the legal field.

Because AI outputs cannot be reliably verified without undermining the very efficiency and time-saving benefits that make the technology appealing, expecting attorneys to conduct a reasonable inquiry into AI-generated content becomes impractical. Legal professionals cannot meaningfully check AI outputs through parallel problem solving, as doing so requires independent legal analysis that negates AI's utility. Moreover, the absence of a single correct answer in legal reasoning, combined with the opacity of AI decision-making, prevents attorneys from detecting when bias or flawed logic has shaped the model's conclusions. Since a reasonable inquiry demands that attorneys critically assess the legitimacy of their sources and reasoning, the inability to trace or validate AI-generated outputs makes such an inquiry inherently less feasible.

### B. CONVENTIONAL METHODS OF ATTRIBUTING BLAME IN HUMAN-MACHINE INTERACTIONS MAY CREATE UNJUST OUTCOMES

Utilizing AI to substitute human intelligence has been a primary goal of the technology from its inception.[122] But what if an attorney's reliance on AI is a double-edged sword? If attorneys delegate legal reasoning to AI, and yet maintain complete liability for assertions made to the court under Rule 11, doesn't that risk attorneys absorbing liability for errors they couldn't reasonably be expected to predict? Attorneys take a risk in melding their practice with automated systems, especially in the context of Rule 11, because of the "moral crumple zone."[123]

In *Moral Crumple Zones: Cautionary Tales in Human-Robot Interactions*, anthropologist Madeleine Clare Elish explored the concept of blame allocation in relationships between humans and automated systems.[124] Elish posits that responsibility for failures by automated systems is frequently misattributed to

121. *See infra* Subpart.III.C.
122. Rissland, *supra* note 12, at 1980; *see* Bergmann & Stryker, *supra* note 114; ENTEFY, *supra* note 114.
123. Elish, *supra* note 110, at 41.
124. *Id.*

humans who have very little control over the situation.[125] Human operators often absorb responsibility for failures by automated systems in (1) the aftermath of a highly publicized event or accident, when (2) the party "should have known" that a failure had occurred or that the technology was malfunctioning.[126] The disproportionate absorption of liability by humans when automated systems fail is called the moral crumple zone.[127]

Elish uses the crash of Air France Flight 447 to illustrate the misallocation of blame implicated in human-machine interactions.[128] There, an Airbus crashed into the Atlantic Ocean, tragically killing everyone on board.[129] A malfunction of the Pitot tubes caused the airplane's autopilot to show false readings, and, although recovering the situation may have been possible, the false readings caused panic and confusion that prevented the pilots from stabilizing the airplane.[130] News reports primarily blamed the pilots for failing to react appropriately and promptly, and for losing "cognitive control of the situation."[131]

But there was more to the story. The Pitot tubes had a known tendency to freeze and had been replaced on other airbuses.[132] Also, aviation experts claimed that the autopilot system could "ensure that the plane flies safely at all times, even though the pilot may make an error."[133] At the time, professionals thought the autopilot was infallible and even "smart enough to protect the airplane and the people aboard it from any dumb moves by the pilot."[134] Elish explains that while courts consistently upheld pilots' ultimate liability, increased automation reduced pilots' control over the aircraft; this results in pilots remaining fully liable despite having diminished control.[135] The pilots' absorption of liability despite having limited control placed them in the moral crumple zone.[136]

The adoption of AI in legal practice risks placing attorneys in the moral crumple zone, where increased automation reduces their control, yet full liability under Rule 11 remains with them. As mentioned above, Elish discusses two factors for when the moral crumple zone is likely to occur. Both are present in the temporal adoption of AI by legal practitioners. First, following *Mata*, the legal community was in an uproar of shock and confusion that such an error had occurred.[137] The sanctions assigned and broader publicity of the case likely

---

125. *Id.*
126. *See id.* at 42.
127. *Id.*
128. *Id.* at 46.
129. *Id.*
130. *Id.* at 47–48.
131. *Id.* at 48 (citation omitted).
132. *Id.*
133. *Id.* at 49 (citation omitted).
134. *Id.* (citation omitted).
135. *Id.*
136. *Id.* at 50.
137. *See* Christopher F. Lyon, *Fake Cases, Real Consequences: Misuse of ChatGPT Leads to Sanctions*, 28 NYSBA NYLITIGATOR 8, 8 (2023); William A. Ryan, Allen Garrett & Brad Sears, *Practical Lessons from the Attorney AI Misstep in* Mata v. Avianca, ASS'N OF CORP. COUNS. (Aug. 8, 2023), https://www.acc.com/resource-

satisfy the first prong of Elish's theory because the event was novel and highly publicized.[138]

Second, considering whether the party "should have known" when there was a failure of the technology is more ambiguous. All the offending attorneys in *Mata*, *Cohen*, and *Park* showed a substantial lack of awareness regarding AI, its uses, and its capabilities.[139] Two of the transgressing attorneys specifically explained that they were under the impression that generative AI services were supercharged search engines.[140] Substantial lack of awareness in the legal community regarding how AI works and its limitations makes it less likely that attorneys "should have known" when the system failed. However, the fact that sanctions have been issued under Rule 11 for bad faith[141] shows that, regardless of an attorney's understanding, there is an expectation within the legal community that attorneys know when their use of AI results in an error. Like how the pilots of Air France Flight 447 were expected to know when the autopilot malfunctioned despite receiving false readings, attorneys are also likely to be held to the standard of knowing when AI malfunctions. This reality places attorneys squarely in the moral crumple zone when it comes to liability under Rule 11 challenges.

### C. THE BLACK BOX PRECLUDES REASONABLE INQUIRY

With AI beginning to conduct legal reasoning, interpretation, and strategy, it is unclear how attorneys can conduct a reasonable inquiry because of the opacity of the black box.[142] Some scholars have proposed that legal experts could serve as "explainers" of model outputs if the outputs are unfair, illegal, or

library/practical-lessons-attorney-ai-missteps-mata-v-avianca; Owen R. Wolfe, Eddy Salcedo & Jamie Anderson, *Update on the ChatGPT Case: Counsel Who Submitted Fake Cases Are Sanctioned*, SEYFARTH SHAW LLP (June 26, 2023), https://www.seyfarth.com/news-insights/update-on-the-chatgpt-case-counsel-who-submitted-fake-cases-are-sanctioned.html. For evidence that this was not only nationally, but internationally publicized, see Mary Or, *Lawyer Blames ChatGPT for 'Bogus Citations,'* NZ LAW. (May 31, 2023), https://www.thelawyermag.com/nz/news/general/lawyer-blames-chatgpt-for-bogus-citations/447559.

138. *See* Benjamin Weiser, *Here's What Happens When Your Lawyer Uses ChatGPT*, N.Y. TIMES (May 28, 2023), https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html; Megan Cerullo, *A Lawyer Used ChatGPT to Prepare a Court Filing. It Went Horrible Awry*, CBS NEWS (May 29, 2023, at 17:40 ET), https://www.cbsnews.com/news/lawyer-chatgpt-court-filing-avianca; Ramishah Maruf, *Lawyer Apologizes for Fake Court Citations From ChatGPT*, CNN (May 28, 2023, at 15:28 ET), https://www.cnn.com/2023/05/27/business/chat-gpt-avianca-mata-lawyers/index.html; John Naughton, *A Lawyer Got ChatGPT to Do His Research, but He Isn't AI's Biggest Fool*, GUARDIAN (June 3, 2023, at 11:00 ET), https://www.theguardian.com/commentisfree/2023/jun/03/lawyer-chatgpt-research-avianca-statement-ai-risk-openai-deepmind.

139. *See* Mata v. Avianca, Inc., 678 F. Supp. 3d 443, 456 (S.D.N.Y. 2023); Park v. Kim, 91 F.4th 610, 614–15 (2d Cir. 2024); Letter from E. Danya Perry to Hon. Jesse M. Furman, *supra* note 99.

140. *See Mata*, 678 F. Supp. 3d at 456; Letter from E. Danya Perry to Hon. Jesse M. Thurman, *supra* note 99.

141. *See Mata*, 678 F. Supp. 3d at 464; *see also Park*, 91 F.4th at 615–16 (referring attorney to a disciplinary committee rather than imposing sanctions).

142. *See* Lee, *supra* note 53; *see also* Harry Surden, *Artificial Intelligence and Law: An Overview*, 35 GA. STATE U.L. REV. 1305, 1336 (2019) (discussing the lack of transparency around AI systems' decision-making process); Stryker & Scapicchio, *supra* note 1 ("Many generative AI models are 'black box' models, meaning it can be challenging or impossible to understand their decision-making processes.").

wrong.[143] However, without information about the model's weights and training data, the path the model took to reach any given conclusion will still be obscured.[144] As discussed above, even if legal experts are able to verify the accuracy of a model's output, there is still be no way of knowing whether the model arrived at that conclusion through an appropriate, legal manner. For example, this issue was illustrated when a researcher tested an AI model's ability to do multiplication and despite outputting the correct value, all the model's sub-steps were incorrect.[145]

As a result of this opacity, a practitioner who used AI to craft an argument could advocate for a legal outcome that is based on unconstitutional or discriminatory presumptions. For example, in an employment discrimination case, the employer may allege that the AI model used to screen applications lawfully rejected the plaintiff for neutral and job-related criteria. However, because the AI's decision-making process occurs inside a black box, where its internal weighing of factors and reasoning are not visible to the employer or the attorney, the attorney lacks the normal ability to assess whether the AI's logical pathway was legitimate.[146] The AI model may have arrived at its decision by relying on historical biases or presumptions derived from otherwise neutral elements of the plaintiff's application like zip code, name characteristics, personal interests, or other proxies for protected classes. However, without access to the model's internal "reasoning" or training data, the attorney is incapable of determining whether that reasoning complies with anti-discrimination laws or if it raises equal protection concerns. As a result, the attorney is precluded from conducting a meaningful pre-filing inquiry as required by the reasonable inquiry standard of Rule 11 because the factual foundations of a claim would be unchallengeable and inaccessible within the black box.

A parallel example exists through the lens of consumer lending and mortgage applications, but this time from the perspective of the attorney using AI to evaluate the case. Suppose an attorney uses AI to assess whether a bank's denial of a mortgage application is lawful. The model was trained on data instilled with historical biases and determines that the plaintiff was duly denied based on objective creditworthiness factors.[147] However, due to the black box,

143. *See* Wilson & Daugherty, *supra* note 111, at 117–18.

144. *See id.* at 117; Ananya, *supra* note 111.

145. Ananya, *supra* note 111.

146. For a similar case, see *Mobley v. Workday, Inc.*, 740 F. Supp. 3d 796, 801 (2024), where an anti-discrimination claim was brought against Workday for their hiring algorithm.

147. Examples of this type of bias in AI model's determining creditworthiness are, unfortunately, pretty common. *See, e.g.*, James Kleimann, *Mortgage Underwriting Algorithm at Heart of Wells Fargo's Racial Disparity Lawsuit*, HOUSINGWIRE (June 13, 2024, at 12:10 PT), https://www.housingwire.com/articles/mortgage-underwriting-algorithm-at-heart-of-wells-fargos-racial-disparity-lawsuit; Dan Armstrong, *AI Exhibits Racial Bias in Mortgage Underwriting Decisions*, LEHIGH UNIV. NEWS (Aug. 20, 2024), https://news.lehigh.edu/ai-exhibits-racial-bias-in-mortgage-underwriting-decisions; Press Release, U.S. Att'y's Off. Dist. of Colo., Justice Department Sues Rocket Mortgage, Appraisal Management Company and Appraiser

the attorney would be unable to verify whether the model reached its conclusion through sound legal logic or bias. This could likely lead to the attorney failing to conduct a reasonable inquiry as required by Rule 11 and facing sanctions as a result.

Despite being an impressive tool for legal reasoning, the obscurity of the decision-making process means that attorneys who use the technology have no meaningful way of conducting a reasonable inquiry. Without replicating the work done by the AI model independently to ensure its legitimacy, the attorney will always be unsure of the AI output to some degree, which renders the technology essentially useless.

## IV. EVEN IF A REASONABLE INQUIRY WERE POSSIBLE, APPLICATION OF THE STANDARD WOULD BE INCONSISTENT

Even if a reasonable inquiry were possible, the application of the standard would be inconsistent and inequitable. There is a wide disparity in the judicial and wider legal community when it comes to AI, what type of use is appropriate, and what type of inquiry satisfies the standard.

### A. JUDICIAL UNDERSTANDINGS AND A PATCHWORK OF STANDING ORDERS

As attorneys' reliance on AI becomes more ubiquitous, judges will be faced with the task of determining whether an attorney has conducted a reasonable inquiry of legal arguments that were generated or heavily influenced by AI. The Second Circuit and the Southern District of New York agree that failing to check AI-generated case law is evidence of bad faith.[148] Additionally, an attorney may reasonably be expected to verify the existence of cases because doing so would not frustrate the purpose of using AI in their legal practice.[149] But, judges' approaches to the regulation of the use of AI in their courtrooms via standing orders vary widely across jurisdictions.

Following *Mata*, judges across various jurisdictions issued a swath of standing orders, all requiring different standards of care and disclosure, ranging from a complete ban on the use of AI tools to free use subject to disclosure.[150] However, banning the use of generative AI in lieu of figuring out how to

---

for Race Discrimination in Mortgage Refinance Application (Oct. 21, 2024), https://www.justice.gov/usao-co/pr/justice-department-sues-rocket-mortgage-appraisal-management-company-and-appraiser-race.

148. *See, e.g.*, Mata v. Avianca, Inc., 678 F. Supp. 3d 443, 460 (S.D.N.Y. 2023); *see also* Park v. Kim, 91 F.4th 610, 614–15 (2d Cir. 2024) (holding that the attorney "made *no* inquiry, much less the reasonable inquiry required by Rule 11").

149. *See supra* Subpart.III.A.

150. Shweta Watwe, *Judges Reflect on GenAI Use One Year After ChatGPT's Debut*, BLOOMBERG L. (Nov. 28, 2023, at 02:00 PT), https://news.bloomberglaw.com/litigation/judges-reflect-on-genai-use-one-year-after-chatgpts-debut; *see also* Mike Langan, *Column: Second Circuit Indicates That Rule 11 Covers Relying on Generative AI to Conduct Legal Analysis*, N. DIST. OF N.Y. FED. CT. BAR ASS'N (Apr. 10, 2024), https://ndnyfcba.org/column-second-circuit-indicates-that-rule-11-covers-relying-on-generative-ai-to-conduct-legal-analysis (explaining that courts are amending their local rules to address AI usage); Gunder, *supra* note 79, at 322 (discussing the shortcomings of regulating attorney conduct under Rule 11 and local rules).

adjudicate its use is shortsighted and ignores the massive potential advantages that AI offers the legal community.[151]

First, just three days after sanctions were issued in *Mata*, Judge Brantley Starr of the U.S. District Court of the Northern District of Texas issued the first standing order.[152] It required attorneys and pro se litigants to certify that "no portion of any filing [was] drafted by generative artificial intelligence" or that if there was the use of generative AI, it was checked "by a human being."[153] Other courts followed suit; for example, one week later, the U.S. District Court for the Eastern District of Pennsylvania issued a standing order requiring both disclosure of AI use and certification that all citations to the law or the record be verified for accuracy.[154]

Further, a Magistrate Judge of the U.S. District Court for the Northern District of Illinois specified a standing order providing in part that "the Court will continue to presume that the Rule 11 certification is a representation by filers, as living, breathing, thinking human beings, that they themselves have read and analyzed all cited authorities . . . ."[155] Ultimately, judges' expectations range from abstinence, to disclosure, to complete human verification to satisfy the existing reasonable inquiry standard.[156] As a result, application of the reasonable inquiry standard will likely be inconsistent or even disparate. Markedly, references to the use of legal reasoning or novel legal arguments are absent in the standing orders issues thus far, which leaves a big question mark as to what judicial expectations are for such uses of AI.

---

151. Wilson & Daugherty, *supra* note 111.

152. Wilson & Daugherty, *supra* note 111; Joseph Anderson, *AI and the Legal Puzzle: Filling Gaps, but Missing Pieces*, 75 MERCER L. REV. 1521, 1544 (2024) (citation omitted).

153. Paul W. Grimm, Maura R. Grossman & Daniel G. Brown, *Is Disclosure and Certification of the Use of Generative AI Really Necessary?*, JUDICATURE, 2023, at 69, 70.

154. *Id.*

155. *Id.* at 70–71.

156. Judge Michael Newman from the Southern District of Ohio issued a standing order to take effect in December of 2023 banning the use of AI in the preparation of any filing submitted to the court, excluding Westlaw and Lexis. *See* Michael J. Newman, *Standing Order Governing Civil Cases*, S. DIST. OF OHIO, https://www.ohsd.uscourts.gov/sites/ohsd/files//MJN%20Standing%20Civil%20Order%208.27.25.pdf (last visited Mar. 8, 2026). Taking a different approach, many federal judges including Judge Araceli Martinex-Olguin from the Northern District of California, Magistrate Judge Jeffrey Cole from the Northern District of Illinois, Judge Evelyn Padin from the District of New Jersey and many more, issued standing orders requiring some combination of disclosure and certification for materials created by or heavily influenced by AI. *See Standing Order for Civil Cases Before District Judge Araceli Martínez Olguín*, N. DIST. OF CAL. (Nov. 14, 2025), https://cand.uscourts.gov/sites/default/files/standing-orders/AMO-CivilStandingOrder-10-20-2025.pdf; Jeffrey Cole, *The Use of "Artificial Intelligence" in the Preparation of Documents Filed Before This Court*, N. DIST. OF ILL. https://www.ilnd.uscourts.gov/_assets/_documents/_forms/_judges/Cole/Artificial%20Intelligence%20 standing%20order.pdf (last visited Mar. 8, 2026); *Judge Evelyn Padin's General Pretrial and Trial Procedures*, DIST. CT. OF N.J., https://www.njd.uscourts.gov/sites/njd/files/EPProcedures.pdf (last visited Feb. 20, 2025).

### B. SCHOLARLY ATTITUDES ON AI ADOPTION AND ETHICAL OBLIGATIONS

With scholarly perspectives ranging from adopting an innovation-forward approach[157] to tempering the use of AI with a human touch in all circumstances,[158] scholars are torn on how to approach the issue of AI integration with attorney work-product. The lack of consistent baseline knowledge about AI within the judicial community[159] is mirrored among practicing attorneys, significantly shaping scholarly perspectives on the issue.[160]

Harry Surden, a professor and AI expert at the University of Colorado School of Law, argued that existing AI case law, where attorneys submitted materials directly created by AI, indicates a lack of AI literacy amongst the general legal community.[161] He also argued that it was key for attorneys to put in the time to understand the strengths and weaknesses of the technology.[162] While he acknowledges the advantages of saving time and money by using AI in processes like discovery, contract review, and template creation, he advises that AI use must be ultimately reviewed by a human.[163] This "human in the loop" approach may be appropriate to adjudicate AI-generated content in its current form, but it does not account for more complex applications of AI, like generating novel legal arguments. Surden may respond to this point by arguing that areas of legal practice involving human cognition, like legal reasoning, "will likely be difficult to replace given the current state of AI technology."[164] While this may be true, AI technology is developing at an exponential rate, so it's unlikely to be true for long.[165]

Similarly, Dr. Ilia Kolochenko, a partner in the cyber security practice at Plat Law LLP, pointed out that in the wake of the hallucination cases, the core of the issue regarding AI and representations to the court lies not in the instances of hallucinations but rather in "insufficiently trained or inexperienced lawyers

---

157. *See* Arruda, *supra* note 108, at 457–58. For some discussion about how an innovation-forward approach is important to provide tools to help close the access to justice gap, see Tomas Weber, *Artificial Intelligence and the Law: Legal Scholars on the Potential for Innovation and Upheaval*, STAN. L. SCH. (Dec. 5, 2023), https://law.stanford.edu/stanford-lawyer/articles/artificial-intelligence-and-the-law.

158. *See, e.g.*, Rebecca Crootof, Margot E. Kaminski & W. Nicholson Price II, *Humans in the Loop*, 76 VAND. L. REV. 429, 473–74 (2023); Surden, *supra* note 142, at 1330; I. Glenn Cohen, Boris Babic, Sara Gerke, Qiong Xia, Theodoros Evgeniou & Klaus Wertenbroch, *How AI Can Learn From the Law: Putting Humans in the Loop Only on Appeal*, NPJ DIG. MED., Aug. 25, 2023, at 1, 1.

159. *See supra* Subpart.IV.A.

160. Gunder, *supra* note 79, at 322; *see also* Solomon Klappholz, *Legal Professionals Face Huge Risks When Using AI at Work*, ITPRO (Feb. 27, 2025), https://www.itpro.com/technology/artificial-intelligence/legal-professionals-face-huge-risks-when-using-ai-at-work (stating that the real problem may be the widespread use of generative AI by lawyers who do not understand its well-understood limitations, rather than AI itself).

161. *Id.*

162. Surden, *supra* note 142, at 1326.

163. *Id.* at 1330–31.

164. *Id.* 1330–32.

165. The exponential growth of AI has exceeded that of traditional computing, even exceeding the projections of Moore's Law for exponential growth. *See* Don Cowan, *Exponential Growth*, SCI. OF MACH. LEARNING & AI, https://www.ml-science.com/exponential-growth (last visited Mar. 8, 2025).

. . . who may all have a knowledge gap . . . ."[166] Furthermore, even if attorneys are checking the outputs of AI for accuracy before submitting them to the court, "some lawyers simply cannot spot legal mistakes related to both procedure and substance of the case."[167] This perspective is shared by Professor Surden, who also argues attorneys' failures to spot mistakes and issues in AI-generated content is ultimately the issue, rather than the use of the technology.[168] However, characterizing the problem as attorneys failing to identify mistakes in AI-generated content not only places attorneys squarely in the moral crumple zone but also does not account for the increasing ubiquitousness of AI implementation,[169] the massive industry push to adopt early,[170] or the disparate impact on new attorneys with less experience spotting mistakes.

On the other hand, some scholars are less forgiving. Andrew Perlman, the Dean of Suffolk University Law School, has explained that he views a failure to verify AI-generated content as a blatant showing of incompetence.[171] And further, firms like Morgan & Morgan have committed to firing any attorney who files an illegitimate, AI-generated case with the court.[172] Since the Rule 11 reasonable inquiry standard is objective and does not account for individual attorney knowledge, a lack of general knowledge and education regarding AI is especially problematic.[173]

Alternatively, some scholars argue that there is an ethical obligation to adopt AI into legal practice.[174] Basing this assertion on the Model Rules of Professional Conduct's requirement for lawyers to "provide competent representation,"[175] the fact that twenty-six states have adopted an ethical duty of technology competence, and the fiscal implications for transitioning substantial attorney hours to AI hours, some scholars view the use of AI as an obligation.[176]

Paradoxically, attorneys are stuck between a rock and a hard place. While there is a massive lack of general knowledge and education regarding AI and its capabilities, there has been a rush by some lawyers and pro se litigants to promptly adopt the technology regardless.[177] Adoption of AI in the legal sector has dramatically increased from 19% to 79% between 2023 and 2024; 70% of

166. Grace Robbie, *AI Literacy (Not Hallucinations) the Big Worry, Experts Warn*, LAWS. WEEKLY (Feb. 24, 2025), https://www.lawyersweekly.com.au/biglaw/41562-ai-literacy-not-hallucinations-the-big-worry-experts-warn.

167. Klappholz, *supra* note 160.

168. *Id.*

169. *See supra* Subpart.III.B; LEGAL.IO, *supra* note 108; A.B.A., *supra* note 108.

170. *See, e.g.*, Arruba, *supra* note 108; LEGAL.IO, *supra* note 108.

171. Robbie, *supra* note 166.

172. Sara Merken, *AI 'Hallucinations' in Court Papers Spell Trouble for Lawyers*, REUTERS (Feb. 18, 2025, at 12:55 PT), https://www.reuters.com/technology/artificial-intelligence/ai-hallucinations-court-papers-spell-trouble-lawyers-2025-02-18.

173. Gunder, *supra* note 79, at 336, 347.

174. Arruda, *supra* note 108, at 455–57.

175. MODEL RULES OF PRO. CONDUCT r. 1.1 (A.B.A. 2023).

176. Arruda, *supra* note 108, at 457–58.

177. Caton, *supra* note 62, at 50.

clients surveyed were either positive or neutral to the use of AI by firms.[178] With offerings of increased productivity, less reliance on outside counsel, and nearly three-fourths of a firm's hourly billable tasks potentially automatable, there is massive pressure on attorneys to adopt this technology.[179] Unfortunately, only 43% of attorneys surveyed reported having even a basic grasp of generative AI, and only 42% said their firms offered educational opportunities.[180] Other attorneys indicated that there were incomplete training policies and an overall lack of alignment within the firms regarding how and where AI should be used.[181]

In June 2024, the American Bar Association ("ABA") released the results of a survey of law school administrators and faculty regarding the integration of AI into the curriculum.[182] The survey found that 55% of responding schools offered an AI specific course, and 85% of schools were contemplating changing their curriculum in response to the increased prevalence and AI tools.[183] This shows substantial steps are being taken to better prepare future lawyers for integrating their work product with that of generative AI. Unfortunately, that same effort has not been shown in the workforce to educate already licensed attorneys.[184] The California State Bar has issued a technology rule requiring attorneys to keep up to date with technological advances, but has left it up to the individual and their comfort level with the technology to determine whether they should use it in their practice.[185] Therefore, the onus for comprehending how AI works and its current capabilities, discerning whether its use is appropriate, and applying it in practice falls largely on the attorney's shoulders.[186] The lack of higher regulation of use is a mistake that will ultimately lead attorneys to make errors and inappropriately rely on this technology.

178. LEGAL.IO, *supra* note 108.

179. *Id.*

180. *Id.*

181. *Id.*

182. A.B.A., *supra* note 108.

183. *Id.*

184. *See, e.g., MCLE Requirements*, STATE BAR OF CAL., https://www.calbar.ca.gov/legal-professionals/maintaining-compliance/mcle/mcle-requirements (last visited Mar. 8, 2026) (requiring only one hour of technology training).

185. *California Lawyers Association Task Force on Artificial Intelligence*, CAL. LAWS. ASS'N, https://calawyers.org/california-lawyers-association/california-lawyers-association-task-force-on-artificial-intelligence (last visited Mar. 8, 2026).

186. *See also ABA Issues First Ethics Guidance on a Lawyer's Use of AI Tools*, A.B.A., (July 29, 2024), https://www.americanbar.org/news/abanews/aba-news-archives/2024/07/aba-issues-first-ethics-guidance-ai-tools (discussing the attorneys' responsibility to "fully consider their applicable ethical considerations" before using AI and emphasizing that "lawyers must be vigilant"); *see generally* Matic, *AI in Law: Opportunities, Challenges, and Ethical Considerations for Lawyers*, AUTOGPT (Apr. 16, 2024), https://autogpt.net/ai-in-law-opportunities-challenges-and-ethical-considerations-for-lawyers (discussing attorney's responsibility to embrace and stay ahead of AI and its development while understanding and avoiding its limitations).

## V. THE SOLUTION: AMEND RULE 11

For the aforementioned reasons, Rule 11 should be adapted to set realistic expectations and guidelines for attorneys as they integrate AI into their practice. The proposed language advances three main goals: encouraging AI adoption and innovation in the legal field, ensuring a level of transparency and explainability in AI-generated content, and tempering judicial discretion by providing a uniform framework for assessing a reasonable inquiry. These AI-specific updates will help the legal profession remain at the forefront of technological advancements while enhancing the quality of client representation.

### A. PROPOSED LANGUAGE AND GOALS

In 1983, Rule 11 was amended to better support the legitimacy and integrity of the court.[187] The increasing prevalence of AI is another opportunity to redefine the judicial system's expectations for attorneys' submissions to the court.[188] An amendment must balance embracing AI as a massively advantageous development in the legal field while bolstering our rules to better protect litigants and the integrity of our judicial system.

In its current form, Rule 11 is inadequate for adjudicating filings by attorneys who increasingly incorporate AI-generated content into their work-product. The reasonable inquiry standard is incompatible with AI because, unlike previous technologies, AI cannot be tested and vetted through parallel problem solving.[189] The black box further precludes meaningful inquiry into the AI model's reasoning.[190] As a result, traditional methods of assigning fault may place undue burden on attorneys who use the technology.[191] Even if a reasonable inquiry were possible, its application would be inconsistent and inequitable due to the lack of consensus on appropriate AI use within the judicial, professional, and scholarly communities.[192] To address these discrepancies, an amendment to Rule 11 is necessary.

The goals of the proposed amendment to Rule 11 are threefold. First and foremost, for the legal profession to remain relevant and helpful to litigants, it must embrace and encourage the use of AI. Prohibiting the use of AI in court filing is averse to this aim and should not be promoted. Second, the public's trust in the legal system will be maintained by requiring a certain amount of transparency or explainability regarding materials attorneys submit to the court. Third, conversely to the 1983 amendment, judicial discretion should be limited in the application of Rule 11 to AI-generated or influenced court filings. It is

187. *See supra* Subpart.II.A; Carter, *supra* note 74.

188. *See generally* Carter, *supra* note 74 (discussing how prior amendments to Rule 11 clarified the extent to which an attorney had an affirmative duty to investigate the prima facie validity of the claims asserted and, thus, increased the expectations on attorneys).

189. *See supra* Subpart.III.A.

190. *See supra* Subpart.III.C.

191. *See supra* Subpart.III.B.

192. *See supra* Part.IV.

unreasonable to expect judges to fully understand and adjudicate this cutting-edge and ever-changing technology without guidance. Allowing too much judicial discretion opens attorneys to the threat of the moral crumple zone. This Note suggests the following amendment to Rule 11.[193]

Rule 11(b) as Amended (modifications are bolded)

> (b) Representations to the Court. By presenting to the court a pleading, written motion, or other paper—whether by signing, filing, submitting, or later advocating it—an attorney or unrepresented party certifies that to the best of the person's knowledge, information, and belief, formed after an inquiry reasonable under the circumstances:
>
> (1) it is not being presented for any improper purpose, such as to harass, cause unnecessary delay, or needlessly increase the cost of litigation;
>
> (2) 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;
>
> (3) the factual contentions have evidentiary support or, if specifically so identified, will likely have evidentiary support after a reasonable opportunity for further investigation or discovery; **~~and~~**
>
> (4) the denials of factual contentions are warranted on the evidence or, if specifically so identified, are reasonably based on belief or a lack of information**; and**
>
> **(5) if an attorney uses an AI-based tool to assist, prepare, or create any filing, they must take reasonable steps to verify the accuracy, reliability, and legality of the AI-generated content, however:**
>
> > **(A) Liability for errors or misrepresentations arising from AI-generated content shall be assessed in proportion to the attorney's ability to reasonably verify or interpret the AI output.**
> >
> > > **(i) Sanctions shall not be issued for inaccuracies or biases inherent in AI that cannot be readily detected through inquiry.**
> >
> > **(B) Courts may not prohibit the use if AI in court filings. However, attorneys must be prepared to provide a reasonable explanation for AI-generate legal arguments for factual assertions.**
> >
> > **(C) Judicial discretion in sanctioning AI-related errors must be guided by uniform standards that account for AI's evolving nature.**

The proposed language advances all three stated goals. First, 5(B) ensures that AI won't cause sanctions outright which encourages its use in the legal field and limits deterrence caused by fear of sanctions. This allows the legal community to reap the benefits of using AI in practice and support the goal of embracing and encouraging the adoption of AI in the legal profession.

Second, 5(A) addresses tensions and ambiguity regarding attorney liability for AI-related errors and the threat of the moral crumple zone. By proportionally

193. FED. R. CIV. P. 11(b).

allocating accountability based on an attorney's ability to reasonably verify AI-generated output, it mitigates the dilemma of attorneys bearing ultimate responsibility despite having limited control. This approach promotes fair review and transparency while preventing undue penalties on attorneys.

Third, 5(C) limits judicial discretion by requiring judges to carry out an analysis using uniform standards to determine whether sanctions are appropriate in the given situation.[194] Tasking judges with assessing whether attorneys took reasonable steps based on uniform standards, rather than their own interpretations of the complexities of AI, ensures that judicial enforcement remains fair and uniform while preventing undue burdens on judges or attorneys.

## B. Anticipated Benefits of the Proposed Amendment

The new language of 5(B) prohibits all federal courts from banning the use of AI in attorney filings, fostering judicial engagement with the technology and encouraging a more informed approach to adjudicating AI-related legal issues. By requiring judges to interact with AI-generated content, rather than excluding it outright, the amendment promotes a legal landscape that keeps pace with technological advancements. Furthermore, allowing AI in all federal courts enhances litigation efficiency and effectiveness by ensuring all litigants have access to modern tools that improve legal research, drafting, and argumentation.

The provision also balances accountability and practicability by requiring attorneys to provide a reasonable explanation for AI-generated legal arguments, without demanding an exact reproduction of the model's reasoning. Since AI operates as a black box, expecting attorneys to fully decipher its decision-making would be unrealistic. Instead, experienced litigants can approximate a legally sound rationale, allowing greater flexibility in leveraging AI while the technology continues to evolve.

This approach benefits both judges and litigants. Judges receive higher-quality work product on average, as AI can assist in refining legal arguments and reducing errors. Litigants, in turn, gain the ability to incorporate AI into their filings without the undue burden of proving the inner workings of the model. By preventing outright bans while maintaining reasonable safeguards, 5(B) ensures AI integration in legal practice is both practical and equitable.

The new language of 5(A) introduces a crucial safeguard by stating that liability for errors or misrepresentations involving AI-generated content "shall be assessed in proportion to the attorney's ability to reasonably verify or interpret the AI output." This provision directly addresses concerns about attorneys being unfairly held responsible for AI's inherent flaws, such as black box reasoning and biases stemming from training data. By reducing liability in cases where errors arise from these systemic AI limitations, 5(A) allows attorneys to utilize AI with less fear of undue repercussions.

194. *See supra* Subpart.V.B.

This proportional approach mitigates the risk of attorneys being placed in the moral crumple zone, where human operators bear full responsibility for failures despite having limited control. Under a traditional liability structure, attorneys could be unfairly blamed for AI's unpredictable or opaque outputs, but by introducing proportional accountability, 5(A) disrupts this dynamic. Attorneys are shielded from the rigid expectation that they must foresee and correct AI-generated errors that are beyond their capacity to detect.

Further reinforcing this protection, 5(A)(i) explicitly prohibits sanctions for inaccuracies or biases inherent in AI that are not readily detectable through reasonable inquiry. By incorporating these measures, the amendment fosters responsible AI integration in legal practice while safeguarding attorneys from disproportionate liability.

The new language of 5(C) establishes a uniform federal standard for assigning sanctions to errors involving AI, promoting consistency and fairness across jurisdictions. By implementing a single national framework, 5(C) eliminates the need for judges to issue numerous and frequently changing standing orders to keep pace with AI's rapid evolution. Instead, it provides a clear, standardized expectation for both attorneys and judges, reducing uncertainty and inconsistencies.

A unified standard also ensures that AI's role in legal practice can be assessed and adapted over time by professionals knowledgeable in both AI and the law. This approach allows for thoughtful, informed adjustments as AI evolves, rather than ad hoc, reactionary rulemaking at the court level. Judges benefit from the removal of the administrative burden of crafting AI-related policies, freeing them to focus on substantive legal matters. Meanwhile, litigants gain confidence in a predictable and consistent regulatory environment, knowing that AI-related sanctions will be applied equitably regardless of the jurisdiction in which they practice. By centralizing AI-related sanctioning criteria, 5(C) fosters a more stable and adaptable legal system, one that accommodates technological advancements while maintaining fairness and procedural clarity.

## CONCLUSION

As AI integrates with legal practice, it is imperative that the profession adapts rather than resists. The current framework of Rule 11 fails to account for the reality that attorneys using AI do not have the same ability to conduct a reasonable inquiry as they do with traditional legal research and reasoning. The inability to efficiently and effectively verify AI-generated outputs, the black box, the risk to particularly novice attorneys who are less likely to identify AI errors, and the wide variety of understandings and attitudes in the judicial and scholarly communities all indicate that an update to Rule 11 is necessary.

The proposed amendment to Rule 11 offers a partial solution that balances innovation, accountability, and fairness. However, this proposal is only the beginning of a much larger conversation. Future scholarship must explore how AI literacy can be incorporated into continued legal education, how courts can standardize AI disclosure requirements, and whether AI itself should be held liable for errors it produces. Failing to adapt to these technological advances will leave attorneys subject to an outdated and ill-fitting regulatory system. Instead, the law should adapt alongside the technology that is reshaping it.

***