

CONFIDENTIALITY OF AI CONVERSATIONS: PROTECTING SELF-REPRESENTED LITIGANTS WHO USE CHATGPT FOR LEGAL ADVICE

ANOO DINESH VYAS[†]

ABSTRACT

When a layperson uses ChatGPT to obtain feedback on a legal matter, attorney-client privilege may not apply, as ChatGPT is not a lawyer, much less a human. Further, while lawyers are entitled to protection for their opinion work-product, it is not clear whether self-represented litigants are entitled to the same protection. Additionally, the broader duty of confidentiality binds only attorneys, not AI systems like ChatGPT.

The public increasingly uses AI tools such as ChatGPT. If a layperson employs ChatGPT for legal advice, particularly in a civil matter, such communications may be discoverable and potentially admissible. This presents an access to justice issue because a self-represented litigant who seeks to understand the scope of their legal rights may not realize that their AI communications can be used against them. Alternatively, they may not be able to afford an attorney, and thus decide to take the risk of communicating with AI anyway.

This Article argues that self-represented litigants should enjoy protection for opinion work-product, and further, AI responses to self-represented litigants should also be permitted to count as opinion work-product. In addition, this Article proposes a discovery management protocol so courts may handle AI communications in a practicable manner. The work-product solution may be implemented more easily than other options, such as extending attorney-client privilege, though that also may be advisable in the interests of justice. Finally, as a backstop, judges could also consider excluding such AI communications from admission under Rule 403.

[†] Assistant Professor of Law, University of Massachusetts School of Law. The author would like to thank Benjamin Anchill, Margaret Drew, Justine Dunlap, Akhi Johnson, Sean Lyness, Eric Mitnick, Gary Spring, Anna-Marie Tabor and Evan Zoldan for reviewing drafts of this piece. Thank you to Rebecca Valentine and Jessica Almeida for their helpful research assistance. Thank you to the wonderful team at Duke Law & Technology Review. All errors are my own. Finally, thank you to my family. Kai, Shantha, and A.J. — this is for you.

TABLE OF CONTENTS

Introduction	187
I. The Public is Turning to AI for Legal Advice	191
II. Existing Confidentiality Doctrines are Inadequate	195
A. Attorney-Client Privilege	195
B. Work-Product Doctrine	199
C. The Duty of Confidentiality	202
D. Cases Involving AI Communications of Parties	203
E. Federal Rules of Evidence.....	206
F. Stored Communications Act	207
III. The Simpler Fix: Focus on the Work-Product Doctrine	208
A. Interpreting the Current Rules as Permitting Opinion Work- Product for Self-Represented Litigants Who Use AI Tools	208
B. Amending the Federal Rules of Civil Procedure.....	209
C. Discovery Management Protocol	211
D. Exclusion Under Rule 403 as Backstop	212
IV. Attorney-Client Privilege is More Complicated	212
A. Benefits of Expanding Attorney-Client Privilege	213
B. The Harms of Expanding Attorney-Client Privilege.....	217
C. Duty of Confidentiality	219
V. Exploring An AI-Specific Rule	220
A. Prismatic Framing: A Separate “Horse Law” vs. Existing Bodies of Law That Happen to Apply to Horses	220
B. AI Specific Rules for Confidentiality.....	222
VI. Philosophical & Sociological Justifications for AI Confidentiality.....	225
A. The Veil of Ignorance.....	225
B. The Performed Self.....	228
VII. Conclusion	230

INTRODUCTION

Given its ease of use and low cost, it is no surprise that the public is increasingly turning to large language models such as ChatGPT, Gemini, and Claude for assistance with legal issues.¹ ChatGPT and other AI systems have well-documented limitations, such as making up cases

¹ See Steve Carroll, *New Survey Identifies How Consumers Would Be Willing to Use Generative AI to Address Legal Needs*, LEXIS NEXIS (May 4, 2023), <https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/post/s/new-survey-identifies-how-consumers-would-be-willing-to-use-generative-ai-to-address-legal-needs>.

that do not exist or providing incorrect information.² Nonetheless, surveys indicate that a prevalent use case of ChatGPT by the general public is to obtain legal advice.³

Laypersons may not realize that their statements to ChatGPT, or ChatGPT's statements to them, could be discoverable and admissible.⁴ Further, even if they comprehend that their conversations could be used

² See Damien Charlotin, *AI Hallucination Cases*, <https://www.damiencharlotin.com/hallucinations/> (last accessed March 1, 2026) (identifying at least 979 situations where AI hallucinated cases were utilized in a court proceeding, with 683 of those occurring in the United States); Adi Robertson, *I tested ChatGPT's deep research with the most misunderstood law on the internet. It got the facts right but the story wrong*, THE VERGE (Feb. 7, 2025), <https://www.theverge.com/openai/607587/chatgpt-deep-research-hands-on-section-230>; Gaby Del Valle, *Why do lawyers keep using ChatGPT?*, THE VERGE, (June 1, 2025), <https://www.theverge.com/policy/677373/lawyers-chatgpt-hallucinations-ai> (observing one situation where “the [court] filing included ‘significant misrepresentations and misquotations of supposedly pertinent case law and history[.]’”). Additionally, hallucinations are a product of AI LLMs and their usage in legal practice is growing. Brandon M. Fierro, *Short Circuit Court: AI Hallucinations in Legal Filings and How To Avoid Making Headlines*, COLE SCHOTZ (Aug. 4, 2025), <https://www.coleschotz.com/news-and-publications/short-circuit-court-ai-hallucinations-in-legal-filings-and-how-to-avoid-making-headlines/> (“The growth of the [AI hallucination] problem is literally exponential.”).

³ Carroll, *supra* note 1. It is not just that ChatGPT is changing how laypersons interact with the law, but generative AI is recalibrating how lawyers themselves interface with it. See 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/> (noting that “[o]ne of the most compelling benefits of AI is time savings. AI-powered tools can accelerate repetitive yet necessary tasks that often consume significant portions of a lawyer’s day. For example, drafting standard contracts — a process traditionally involving tedious cutting, pasting, deleting, and editing — can now be handled more quickly and accurately with AI assistance. These tools can also support legal professionals in conducting faster research, reviewing precedents, and summarizing source material across platforms.”). See also PURDUE GLOBAL LAW SCHOOL, *The Potential Benefits and Risks of ChatGPT in Legal Practice* (Apr. 2, 2024), <https://www.purduegloballawschool.edu/blog/news/risks-and-benefits-chat-gpt-legal-practice>.

⁴ See ChatGPT CEO warns: *Your private chats could end up in court*, THE NAIROBIAN, 2025 WLNR 18778659 (July 28, 2025) (“OpenAI CEO Sam Altman has shared a startling reality which has created a buzz: conversations with ChatGPT, even deeply personal ones, are not protected and could be used as legal evidence in court. ‘Young people especially talk to ChatGPT about everything from relationship problems to emotional struggles,’ Altman said in an interview with podcaster Theo Von. ‘They use it like a therapist or life coach. But unlike a therapist or lawyer, ChatGPT doesn’t come with confidentiality protections.’”).

against them, they may decide to take the risk of communicating with ChatGPT for legal advice because they cannot afford a lawyer.

At least three doctrines protect the confidentiality of legal communications: (1) attorney-client privilege, (2) work-product doctrine, and (3) the duty of confidentiality.⁵ Attorney-client privilege may be inapplicable because ChatGPT is not a lawyer, nor a lawyer's representative.⁶ Moreover, there is tension over whether the Federal Rules of Civil Procedure protects opinion work-product (i.e., mental impressions, conclusions, opinions, and legal theories) of both self-represented litigants and AI-generated responses.⁷ Finally, the duty of confidentiality binds attorneys.⁸

Therefore, self-represented litigants who use AI to assist them with their case are exposed to the possibility that their AI communications could be discovered by the other side.⁹ Because many individuals lack the

⁵ Douglas R. Richmond, *The Attorney-Client Privilege and Associated Confidentiality Concerns in the Post-Enron Era*, 110 PENN ST. L. REV. 381, 382 (2005) ("Confidentiality is central to the practice of law. Indeed, confidentiality is a substantial part of the bedrock on which both litigation and transactional practices are built. Lawyers' duty to protect client communications and information is variously embodied and enforced: the attorney-client privilege is a critical component of evidence law, the work-product doctrine provides important immunity against the discovery of attorneys' files and mental impressions, and state ethics rules make confidentiality a professional responsibility concern.").

⁶ This is despite the fact that many who use generative AI perceive it as having human-like characteristics. See Raphaël Millière & Charles Rathkopf, *Why it's important to remember that AI isn't human*, VOX (Nov. 23, 2023), <https://www.vox.com/future-perfect/23971093/artificial-intelligence-chatgpt-language-mind-understanding>.

⁷ See Jennifer A. Gundlach & Zeus Smith, *Expanding the Federal Work Product Doctrine to Unrepresented Litigants*, 30 GEO. J. POVERTY L. & POL. 49, 71 (2022) (noting that the Federal Rules of Civil Procedure fail to adequately protect the opinion work-product of self-represented litigants, and thus proposing that "Federal Rule of Civil Procedure 26(b)(3)(B) should be amended to allow for the absolute protection from discovery of unrepresented litigants' mental impressions.").

⁸ See Anthony V. Alfieri, *Law Firm Malpractice Disclosure: Illustrations and Guidelines*, 42 HOFSTRA L. REV. 17, 36 (2013) ("The client-lawyer relationship imposes a duty of confidentiality upon a lawyer."). See also MODEL RULES OF PRO. CONDUCT R. 1.6(A) (Am. Bar Ass'n, 2025) ("A lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent, the disclosure is impliedly authorized in order to carry out the representation or the disclosure is permitted by paragraph (b).").

⁹ See Gundlach & Smith, *supra* note 7, at 53. ("[T]he federal rules of discovery include important protections for represented parties that are not afforded to unrepresented parties. Thus, an unrepresented litigant . . . can be ordered to

ability to procure legal counsel,¹⁰ this gap represents a modern-day access to justice issue that is increasingly important given the widespread emergence and adoption of generative artificial intelligence.

For example, in a civil case, an attorney on the other side could ask a self-represented party in a deposition if they consulted ChatGPT regarding the lawsuit. If the self-represented party responds “yes,” the attorney could send a document production request under Federal Rule of Civil Procedure 34 to the self-represented litigant.¹¹ To avoid production, the self-represented party would have to know to assert a work-product objection.¹² Even if they were sophisticated enough to object, the doctrine may not fully protect them, as it is not universally resolved whether self-represented litigants enjoy opinion work-product protection.¹³ A further area of uncertainty is how to treat the AI communications that are responsive to the self-represented litigant.

Part I of this Article details the rising phenomenon of laypersons relying on ChatGPT and other AI systems for legal advice. Part II explores how traditional confidentiality doctrines, specifically the attorney-client privilege, work-product doctrine, and the duty of confidentiality, fail to adequately shield communications that a self-represented litigant has with ChatGPT.

Part III moves the conversation from the descriptive to the prescriptive. It proposes modifying the Federal Rules of Civil Procedure to expand work-product to more fully cover non-lawyers and AI. It also outlines a discovery management protocol for courts to handle AI communications involving legal advice in a practical way. Finally, it

produce materials that contain their mental impressions, case strategy and the like.”).

¹⁰ *Id.* at 57. (“In 2021, there were 344,567 total civil cases filed in federal district courts, 145,970 of which involved unrepresented litigants, roughly forty-two percent. In 2020, there were 470,581 civil case filings, 267,373 of which were filed by unrepresented litigants, roughly fifty-seven percent. Thus, in only two years, close to half a million cases were filed involving unrepresented litigants[.]”).

¹¹ See Fed. R. Civ. Pro. 34(a) (allowing document requests by litigants in civil discovery on any matter “within the scope of [Federal Rule of Civil Procedure] 26(b)”; Fed. R. Civ. Pro. 26(b) (“Parties may obtain discovery regarding any nonprivileged matter that is relevant to any party’s claim or defense and proportional to the needs of the case . . .”).

¹² See Fed. R. Civ. Pro. 34(b)(2) (setting a 30-day deadline for objecting to document requests in discovery); Fed. R. Civ. Pro. 26(b)(3) (setting conditions for applying and overcoming the work-product doctrine in federal civil litigation).

¹³ See *infra* SECTION III; Gundlach & Smith, *supra* note 7, at 53.

encourages judges to exclude such materials under Federal Rule of Evidence 403.¹⁴

Part IV evaluates the advantages and disadvantages of implementing more complicated solutions, including expanding the attorney-client privilege and the duty of confidentiality. Part V explores the possibility of enacting a new rule specific to AI communications. Focusing on work-product is the more practical fix in the near term and is the thrust of this Article. That said, it may be advisable in the interest of justice to reconceptualize attorney-client privilege in some way or pass AI-specific legislation for confidentiality of legal communications.

Part VI offers both philosophical and sociological justifications for robust AI confidentiality. Part VII concludes and provides a roadmap for future scholarship.

I. THE PUBLIC IS TURNING TO AI FOR LEGAL ADVICE

In March 2023, just four months after ChatGPT was released to the public, LexisNexis conducted a survey to gauge how attorneys, law students, and members of the public were using generative AI to help with legal issues.¹⁵ The LexisNexis survey indicated that “a stunning 48% [of respondents] have already used some form of generative AI tool to obtain legal advice or assistance with a legal question.”¹⁶

The American Bar Association has asserted that there are various advantages to AI, including the use of “AI-powered chatbots [that help users] navigat[e] . . . complex court systems [that] can be challenging even for experienced litigators[.]”¹⁷ AI may also help a non-attorney “understand difficult legal jargon.”¹⁸ Further, AI can provide feedback on a legal matter if lawyers are not available in the area,¹⁹ and can

¹⁴ Fed. R. Ev. 403.

¹⁵ Carroll, *supra* note 1.

¹⁶ *Id.*

¹⁷ Nicole Black, *Access to Justice 2.0: How AI-powered software can bridge the gap*, ABA (Jan. 24, 2025), <https://www.americanbar.org/groups/journal/articles/2025/access-to-justice-how-ai-powered-software-can-bridge-the-gap/>.

¹⁸ Ljubiša Metikoš, *Access to Justice and LLMs*, THE DIGITAL CONSTITUTIONALIST (last accessed Aug. 11, 2025), <https://digi-con.org/access-to-justice-and-llms/>.

¹⁹ *Id.* (observing that “LLMs can also be accessed by anyone with an internet connection, reaching even citizens who live in ‘rural legal desserts [sic]’”). The issue of legal deserts has been pointed out as a critical access to justice issue. See Lisa R. Pruitt et al., *Legal Deserts: A Multi-State Perspective on Rural Access to Justice*, 13 HARV. L. & POL’Y. R. 15, 21 (2018) (“If a lack of rural lawyers results in a disproportionate percentage of rural legal issues going unaddressed, then these already-disproportionate rural social problems will be compounded. This is

communicate legal information immediately and at any hour of the day.²⁰ It can also summarize large swaths of information, such as reducing a lengthy contract to key bullet points, and provide relevant tips.²¹

Access to counsel is a severe problem. Estimates indicate that “92% of substantial civil legal needs among low-income Americans are unmet” while “legal aid organizations can only serve about 30% of the

a problem at both the individual level—meaning not enough lawyers or other forms of legal assistance are available to support the legal services needs of individuals—and at the policy level, where lawmakers might better understand how rural problems and their solutions differ from those in urban areas.”). *See also* Daria F. Page & Brian R. Farrell, *One Crisis or Two Problems? Disentangling Rural Access to Justice and the Rural Attorney Shortage*, 98. WASH. L. R. 849, 894 (2023) (observing that it is critical to identify a shared definition of “access to justice,” because “[t]he development of better and more diverse measures of access to justice, which reflect the multilayered nature of the concept, would hopefully guide more targeted, efficient, and effective interventions.”). The authors also note that simply relying on the number of attorneys is an incomplete perspective of access to justice, particularly in the rural context, and observe that “these two issues—the rural attorney shortage and the justice gap—[are often] viewed as one and the same by so many key stakeholders [and that] this conflation seems to disproportionately shape rural access to justice interventions.”). *Id.* The authors posit that a focus on “legal vulnerability” might be a more complete picture. *Id.* at 893.

Further, some scholars have observed that to address rural access to justice, it is less desirable to recruit private practitioners to a location, but rather, to “subsidize specific types of service based on evidence of legal need,” which could “include providing health care benefits to lawyers for part-time public employment; raising the rates for criminal defense, guardian ad litem, and other public appointments; paying lawyers to provide on-site assistance to self-represented parties at county courthouses; and training other types of providers to provide specific types of assistance (such as eviction and debt collection defense).” Elizabeth Chambliss, *Rural Legal Markets*, 12 TEX. A&M L. REV. 961, 969 (2025). Chambliss also recommends “consider[ing] new ways to market flat-fee, routine legal services to make them more visible and accessible to consumers” such as by “allowing private practitioners to set up in Walmart and other retail locations and encouraging the development of online service and referral platforms.” *Id.* The Covid-19 pandemic spurred a scholarly interest in how technology might be able to improve access to justice. *See* Lisa R. Pruitt, *The Rural Lawyer Shortage as (Surprising) Scholarly Impetus*, 69 S.D. L. REV. 352, 367 (2024).

²⁰ *See* Metikoš, *supra* note 18 (“LLMs could increase access to justice by offering informative, immediate, easy-to-use, and free legal guidance to those unable to afford, understand or otherwise receive, traditional human-based legal services”).

²¹ *See id.* (“Scholarship has also identified the usefulness of LLMs in making legal texts more intelligible for those whose access to justice is contingent on the use of more plain language... LLMs can be used to summarize privacy policies and make them more understandable to non-lawyers.”).

clients that come through their doors.”²² Chief Justice John Roberts has stated that, “[f]or those who cannot afford a lawyer, AI can help. It drives new, highly accessible tools that provide answers to basic questions, including where to find templates and court forms, how to fill them out, and where to bring them for presentation to the judge—all without leaving home.”²³ Furthermore, a group of law faculty and researchers have programmed chatbots to illustrate how AI can provide potential benefits to assist with legal matters, such as criminal expungement and understanding rights in landlord-tenant disputes.²⁴ This group suggested that “AI assistants, used appropriately, and with human supervision, have the potential to make legal processes and information more accessible to low-end consumers.”²⁵

Technological developments often do not square with existing legal rules. In 2023, controversy emerged surrounding the AI legal startup called DoNotPay, whose goal was to assist self-represented litigants in contesting traffic tickets.²⁶ The company proposed that a “person challenging a speeding ticket would wear smart glasses that both record[ed] court proceedings and dictate[d] responses into the defendant’s ear from a small speaker.”²⁷ Then, “[t]he system relied on a few leading AI text generators, including ChatGPT and DaVinci” to feed answers to the self-represented litigant.²⁸ The initiative was called off after multiple state bar associations raised objections regarding the unauthorized practice of law.²⁹ The founder of DoNotPay bemoaned the controversy, asserting that “[t]he truth is, most people can’t afford lawyers,” and that “[t]his could’ve shifted the balance and allowed people to use tools like ChatGPT in the courtroom that maybe could’ve helped them win cases.”³⁰

²² Colleen V. Chien & Miriam Kim, *Generative AI and Legal Aid: Results from a Field Study and 100 Use Cases to Bridge the Access to Justice Gap*, 55 LOY. L.A. L. REV. 903, 911–12 (2025).

²³ *Id.* at 916.

²⁴ Colleen V. Chien et al., *How Generative AI Can Help Address the Access to Justice Gap Through the Courts* at 2 (2025), <https://ssrn.com/abstract=4683309> (earlier version of Chien, *supra* note 22) (on file with the Duke Law & Technology Review).

²⁵ *Id.*

²⁶ Bobby Allyn, *A robot was scheduled to argue in court, then came the jail threats*, NPR (Jan. 25, 2023), <https://www.npr.org/2023/01/25/1151435033/a-robot-was-scheduled-to-argue-in-court-then-came-the-jail-threats>.

²⁷ *Id.*

²⁸ *Id.*

²⁹ *Id.*

³⁰ *Id.*

That said, real risks exist for self-represented litigants who use AI.³¹ In *Kruse v. Karlen*, a pro se party relied on AI to file an appellate brief which contained primarily fabricated cases, with “twenty-two of the twenty-four cases cited [being] entirely fictitious or a fictitious citation using a real case name.”³² The court sanctioned the self-represented party and ordered him to pay \$10,000 in fees.³³ Of course actual lawyers have also used fake citations in legal briefs and have received discipline.³⁴ Even judges have relied on fabricated cases in issuing their opinions.³⁵

Another concern is that AI could provide incorrect legal advice. An AI chatbot featured on the official website of New York City “falsely suggested that it is legal for an employer to fire a worker who complains about sexual harassment, doesn’t disclose a pregnancy or refuses to cut their dreadlocks.”³⁶ One lawsuit alleges that ChatGPT assisted a self-represented litigant with a number of frivolous filings.³⁷

While one version of ChatGPT scored in the 90th percentile on the bar exam, some critics assert that this may not translate effectively to counseling actual clients.³⁸ Legal scholarship also indicates that ChatGPT

³¹ See *Kruse v. Karlen*, 692 S.W.3d 43 (Mo. Ct. App. 2024) (sanctioning a self-represented party for submitting a brief with fictitious sources generated by AI).

³² Megan E. Boyd & Brian L. Frye, *The Duty of Efficiency & Generative AI Pedagogy*, 77 WASH. UNIV. J. LAW & POL’Y. 96, 106 (2025).

³³ *Kruse*, 692 S.W.3d at 54.

³⁴ *Id.* See also Jordyn C. Tye, Comment, *Exploring the Intersections of Privacy and Generative AI: A Dive into Attorney-Client Privilege and ChatGPT*, 64 JURIMETRICS J. 309, 318 (2024), (“In June 2023, New York attorneys . . . submitted a legal brief ‘filled with fake judicial opinions and legal citations, all generated by ChatGPT.’”).

³⁵ Debra Cassens Weiss, *Who—or what—is to blame for 2 federal judges’ error-filled withdrawn opinions?*, A.B.A. JOURNAL (Aug. 7, 2025), <https://www.abajournal.com/web/article/who-or-what-is-to-blame-for-2-federal-judges-error-filled-withdrawn-opinions> (“The blame for substantial errors in opinions recently withdrawn by two federal judges likely lies in the first instance with artificial intelligence[.]”).

³⁶ Allen Smith, *NYC AI Chatbot Gives Faulty Legal Advice, Prompting Review*, SHRM (Apr. 5, 2024), <https://www.shrm.org/topics-tools/employment-law-compliance/nyc-ai-chatbot-faulty-legal-advice>. The AI chatbot indicated that it was only to be used for business and could not give legal advice, and also mentioned that some of its answers could be incorrect. *Id.*

³⁷ Nippon Life Ins. Co. Am.’s Complaint, *Nippon v. OpenAI*, No. 1:26-cv-02448 (N.D. Ill. Mar. 4, 2026), <https://fingfx.thomsonreuters.com/gfx/legaldocs/dwpkydrqapm/Nippon%20Life%20v%20OpenAI%2020260304.pdf>.

³⁸ David Alexander, *Why ChatGPT-4’s Score on the Bar Exam May Not Be So Impressive*, N.Y. STATE BAR ASS’N. (Apr. 16, 2024), (Conceding that while ChatGPT scored in the 90th percentile on the bar exam, that result may paint an incomplete picture, and further, “[t]here is also the difficulty of assessing the

struggles in providing accurate, comprehensive answers related to “specific legal text[s].”³⁹ Security risks are also a concern.⁴⁰ The New York legislature recently introduced a bill that would prevent services such as ChatGPT from impersonating lawyers.⁴¹

Troublingly, members of the public may think their communications with AI systems like ChatGPT regarding legal matters are confidential.⁴² This is concerning because the public often uses ChatGPT as a lawyer or therapist, without realizing their AI communications lack confidentiality provisions.⁴³ Sam Altman, the CEO of OpenAI, the organization behind ChatGPT, issued a warning that “unlike a therapist or lawyer, ChatGPT doesn’t come with confidentiality protections.”⁴⁴

II. EXISTING CONFIDENTIALITY DOCTRINES ARE INADEQUATE

This Section provides foundational background regarding how the attorney-client privilege, work-product doctrine, and the duty of confidentiality fail to protect a layperson’s statements to AI, and the AI’s statements in response. It then catalogues recent case law exploring how courts have mapped these confidentiality provisions to litigants using AI. Finally, it explores other bulwarks against disclosure, such as Rule 403 of the Federal Rules of Evidence and the Stored Communications Act.

A. Attorney-Client Privilege

Attorney-client privilege shields communications between a client and an attorney, with “[t]he purpose [] to promote full and frank

capability of an AI system as compared to those of a practicing attorney. The content of the Uniform Bar Exam is not based on the laws of a particular jurisdiction and so may not translate to understanding specific laws.”).

³⁹ Lisa Larrimore Ouellette et al., *Can AI Hold Office Hours?*, J. LEGAL EDUC. (forthcoming 2025) (“[o]nly 49% of ChatGPT’s responses were scored as ‘good,’ meaning that they answered the question correctly and did not contain any wrong information. Claude was slightly higher, at 56%, while only 37% of NotebookLM’s responses met that standard. On the other hand, we scored 26% of ChatGPT’s responses as ‘unacceptable’ due either to completely failing to answer the question or because they contained substantial errors of law. The percentages of “unacceptable” responses were 14% and 31% for Claude and NotebookLM, respectively.”).

⁴⁰ See Tye, *supra* note 34, at 319. (“ChatGPT also poses risks related to hacking and cybersecurity. For example, ChatGPT presents new methods ‘for hackers to potentially breach advanced cybersecurity software.’”).

⁴¹ S. 7263, 2025–2026 Leg., Reg. Sess. (N.Y. 2025), <https://www.nysenate.gov/legislation/bills/2025/S7263>.

⁴² See NAIROBIAN, *supra* note 4.

⁴³ See *id.*

⁴⁴ *Id.*

communications between attorneys and their clients[,] thereby encourag[ing] observance of the law and aid[ing] in the administration of justice.”⁴⁵ The privilege applies to clients, “prospective clients [who seek] legal advice,” and those “individual[s who] reasonably believe [] that the lawyer represents the individual, even if no attorney-client relationship ever materializes[.]”⁴⁶ The attorney-client privilege applies in all types of legal proceedings, including both civil and criminal matters.⁴⁷

The attorney-client privilege is related, though distinct, from the duty of confidentiality. The attorney-client privilege is an evidentiary privilege that prevents communications from being admissible in a legal proceeding.⁴⁸ The duty of confidentiality, in contrast, is broader.⁴⁹ It prevents the attorney from revealing information regarding a client both inside and outside the context of a legal proceeding.⁵⁰

Disclosure, without the client's consent, of communications that normally would be protected by attorney-client privilege is allowed only in limited circumstances, such as when a client indicates they are about to commit a crime.⁵¹ Unlike the work-product doctrine, an opposing party cannot access protected communications, such as by asserting substantial need for factual work-product.⁵²

⁴⁵ *Guy v. United Healthcare Corp.*, 154 F.R.D. 172, 177 (S.D. Ohio 1993) (quoting *Commodity Futures Trading Comm’n v. Weintraub*, 471 U.S. 343, 348 (1985)).

⁴⁶ Thomas E. Spahn, *Attorney-Client Privilege: Overview (Federal)*, Practical Law Practice Note w-023-4677 (Westlaw Jan. 2026).

⁴⁷ See Cary Bricker, *Revisiting the Crime-Fraud Exception to the Attorney-Client Privilege: A Proposal to Remedy the Disparity in Protections for Civil and Criminal Privilege Holders*, 82 TEMP. L. REV. 149, 156 (2009) (observing that while the privilege applies both in civil and criminal contexts, the application of the privilege, particularly in the crime-fraud context, differs in part depending on the type of case at issue, as well as the jurisdiction).

⁴⁸ See *Law. Disciplinary Bd. v. McGraw*, 194 W. Va. 788, 797 (1995).

⁴⁹ Douglas R. Richmond, *Lawyers' Duty of Confidentiality and Clients' Crimes and Frauds*, 38 GA. ST. U. L. REV. 493, 496–97 (2022) (“As Model Rule 1.6(a) makes clear, lawyers' duty of confidentiality is very broad. It is broader than the attorney-client privilege. It is also broader than the confidentiality protections afforded by the work product doctrine. So extensive a duty of confidentiality is necessary to encourage clients to trust their lawyers and to be candid with them.”).

⁵⁰ See *McGraw*, 194 W. Va. at 798.

⁵¹ See TEX. R. EVID. 503(d) (noting that the “privilege does not apply . . . [i]f the lawyer’s services were sought or obtained to enable or aid anyone to commit or plan to commit what the client knew or reasonably should have known to be a crime or fraud.”).

⁵² Thomas E. Smallman, *Civil Procedure: A Corporation's Waiver of the Attorney-Client Privilege and Work-Product Doctrine*, 37 VILL. L. REV. 917, 935 (1992) (“unlike the absolute protection afforded by the attorney-client privilege,

While the attorney-client privilege traditionally involves communications between a client and their attorney,⁵³ it has also been extended in some situations to apply to other individuals, such as the lawyer's administrative assistant,⁵⁴ or a non-attorney whose presence is needed to assist with comprehension.⁵⁵

Sometimes a lawyer will communicate with a client on matters that do not involve legal advice, like when giving business advice.⁵⁶ In that situation, the communication between the lawyer and the client is not privileged.⁵⁷ However, other times a lawyer may communicate with a client and offer both legal advice as well as non-legal advice.⁵⁸ These types

the work-product doctrine is a qualified privilege that may be overcome by a showing of 'substantial need.'").

⁵³ *United States v. Sadler*, 24 F.4th 515, 557 (6th Cir. 2022) (noting eight factors for the attorney-client privilege to apply, with one specifically being a statement to a "professional legal advisor."). The eight factors include "(1) [w]here legal advice of any kind is sought (2) from a professional legal adviser in his capacity as such, (3) the communications relating to that purpose, (4) made in confidence (5) by the client, (6) are at his instance permanently protected (7) from disclosure by himself or by the legal adviser, (8) unless the protection is waived." *Id.* Further, "[w]hile this definition seemingly only applies to the client's statements, courts generally agree that an attorney's statements to a client can also fall within the privilege if that communication would reveal client confidences or legal advice." *Id.*

⁵⁴ See Paul S. Milich, *Georgia Rules of Evidence* § 21:3 (2024-2025 ed.) ("The scope of the attorney-client relationship includes, by necessity, the network of agents and employees of both the attorney and client, acting under the direction of their respective principals, to facilitate the legal representation. The attorney's privileged network includes administrative and legal assistants, from file clerks to law clerks.").

⁵⁵ See *Walsh v. CSG Partners, LLC*, 544 F. Supp. 3d 389, 392 (S.D.N.Y. 2021) (privilege is not waived despite the presence of a third-party "if the purpose of the third-party's participation is to improve the comprehension of the communications between attorney and client."); Michael D. Maicher & Gregory Habiak, *Are You Maintaining Privilege When Using Zoom?*, IMAGINE THAT IP LAW BLOG (Apr. 19, 2021), <https://www.vklaw.com/ImagineThatIPLawBlog/are-you-maintaining-privilege-when-using-zoom>.

⁵⁶ See *In re Polaris, Inc.*, 2021 WL 5913633, at *408 (Minn. Dec. 15, 2021) (holding that "when a document contains both legal advice and business advice, for the attorney-client privilege to apply to the document in its entirety, the predominant purpose of the communication must be legal advice. The privilege does not protect the entirety of the document if legal advice is merely one purpose and not the primary purpose of the communication.").

⁵⁷ See *id.*

⁵⁸ See *id.*

of communications are referred to as “dual-purpose communications.”⁵⁹ Various tests have been put forth, such as the “primary purpose” test, to help guide whether such dual-purpose communications are covered by the privilege.⁶⁰ Under one version of the primary purpose test, if a significant reason for the communication is to confer legal advice, then the whole communication will be considered privileged.⁶¹

The attorney-client privilege protects communications involving a lawyer.⁶² While the privilege extends to communications made to certain representatives of the lawyer, the lawyer remains an essential component of the attorney-client relationship, without whom the privilege does not exist.⁶³

⁵⁹ David Kim, *Ethical Lawyering, Attorney-Client Privilege, and Dual-Purpose Communications in Light of in Re Grand Jury*, 36 GEO. J. LEGAL ETHICS 707 (2023) (observing that “not all attorney-client communications are purely legal in nature. In many contexts, attorneys provide advice that has both legal and non-legal components and purposes. These dual-purpose communications occupy a no-man’s land when it comes to privilege because there is no clear statement in the Constitution, federal statutes, or binding precedent that treats dual-purpose communications uniformly in federal courts.”).

Scholars have noted that great lawyers often provide more to the table than simply a recitation of the black letter law. See Praveen Kosuri, *Beyond Gilson, The Art of Business Lawyering*, 19 LEWIS & CLARK L. REV., 461, 465–66, (2015) (identifying several roles for a business lawyer, including that of a transaction cost engineer, a reputational intermediary, as well as a controller of regulatory costs, but nonetheless observing that those traditional conceptualizations are incomplete and fail to capture the nuance of what lawyers bring to a deal). Kosuri asserts that, “[r]ather than focusing exclusively on technical lawyering skills, the best business lawyers embrace the role of full-fledged advisors to their clients and employ a robust skillset for problem solving that goes beyond what is typically taught in law schools. In many ways, what exceptional business lawyers practice is much more art than science[, with great lawyers demonstrating several characteristics, including] understanding business, allegiance to their client, understanding people, becoming part of the team, creativity, problem solving, and adding value[, which] go beyond technical skills to define the art of lawyering.” *Id.* at 466.

⁶⁰ See Kim, *supra* note 59, at 708–09 (observing that the D.C. Circuit adopted “a primary purpose test” and the Ninth Circuit adopted “the primary purpose test,” while the Seventh Circuit, refused to adopt the test and therefore does not confer privilege unless the communication is only legal in nature).

⁶¹ See *id.*

⁶² See Tex. R. Evid. 503 (noting it is the “Lawyer Client Privilege,” and that the privilege applies for communications with a lawyer or a lawyer’s representative). In a somewhat analogous way, courts have held that self-represented litigants cannot be awarded attorneys’ fees. See *Jackson v. State Off. of Admin. Hearings*, 351 S.W.3d 290, 300 (Tex. 2011).

⁶³ See Tex. R. Evid. 503(a)(4).

B. *Work-Product Doctrine*

One may think that at least in civil cases, if attorney-client privilege is inapplicable, then the work-product doctrine will operate to protect a layperson's interactions with ChatGPT regarding legal issues. However, legal scholars Jennifer A. Gundlach and Zeus Smith have sharply observed that the work-product doctrine may not fully protect self-represented litigants, particularly with regards to their mental impressions, conclusions, opinions, and legal theories.⁶⁴ Gundlach and Smith also point out that courts have been inconsistent in extending opinion work-product protection to self-represented litigants.⁶⁵

An example of work-product would be an attorney's notes from an interview with a third-party witness.⁶⁶ The third-party witness is not the attorney's client, so the interview would not be protected by attorney-client privilege.⁶⁷ The scope of protection is different for so-called "fact work-product" versus "opinion work-product."⁶⁸ For example, if a lawyer's notes contain simply a summary of the witness' statements, that would be fact work-product.⁶⁹ However, if the attorney makes special notes regarding their belief about the witness' credibility, that would be considered "opinion work-product."⁷⁰ These types of work-product are treated differently, as explained more fully below.

The Supreme Court formally recognized the work-product doctrine in 1947 in the seminal case of *Hickman v. Taylor*.⁷¹ Rule 26(b)(3)

⁶⁴ Gundlach & Smith, *supra* note 7, at 65-66 (2022) ("[Some] courts have ruled [there is] no [opinion] work product protection for unrepresented litigants. Still others have suggested, without deciding, they would not extend such protection under the plain language of Rule 26(b)(3)(B). The few courts that have explained their reasoning for expanding the Rule's protections to unrepresented litigants have rooted their decisions in the fundamental equivalency of an unrepresented litigant and an attorney.").

⁶⁵ *Id.*

⁶⁶ *See id.* at 63-64.

⁶⁷ *See id.*

⁶⁸ *See id.*

⁶⁹ *See id.*

⁷⁰ *See id.*

⁷¹ *Hickman v. Taylor*, 329 U.S. 495 (1947). In *Hickman*, a tugboat ship accident killed several workers. *Id.* at 498. The estate of one of the workers brought suit against the tugboat company. *Id.* The tugboat workers who escaped alive were interviewed at a public hearing by a safety commission. *Id.* Later, the attorney for the tugboat company interviewed these tugboat workers on his own and made various notes regarding his interviews. *Id.* The District Court indicated that the communications were discoverable and must be produced. *Id.* at 499. It reasoned that attorney client privilege did not apply because the statements by the workers

of the Federal Rules of Civil Procedure sets forth the basis for the modern work-product doctrine.⁷²

The plain text of Rule 26(b)(3) protects “documents or tangible things” prepared “in anticipation of litigation” by an attorney *or a party*.⁷³ These documents or tangible things can still be accessed by the other side upon a showing of “substantial need” as well as the inability to obtain substantial equivalents by other means without “undue hardship.”⁷⁴

That said, when it comes to “opinion work-product,” the Rules suggest that these matters are essentially not discoverable, regardless of substantial need. However, opinion work-product may not extend to a *party*. Specifically, while the text of the Rule fully protects the mental

were not communications between a lawyer and their client. *Id.* at 502. The District Court ordered the attorney for the tugboat in contempt of court. *Id.* at 505.

The Supreme Court held that the communications were not discoverable. *Id.* at 508. The Court stated that work-product would prevent the discovery of information the lawyer obtained from interviewing witnesses in anticipation of litigation. *Id.* The Court also went on to assert that “the memoranda, briefs, communications, and other writings prepared by counsel for his own use in prosecuting his client's case” were work-product. *Id.* Moreover, “writings which reflect an attorney's mental impressions, conclusions, opinions, or legal theories” were likewise protected as work-product. *Id.*

In justifying its holding, the Court stated that there was a “general policy against invading the privacy of an attorney's course of preparation.” *Id.* at 512. Furthermore, the Court stated that “[i]t is not without reason that various safeguards have been established to preclude unwarranted excursions into the privacy of a man's work.” *Id.* at 497. Indeed, “it is “essential to an orderly working of our system of legal procedure that a burden rests on the one who would invade that privacy to establish adequate reasons to justify production through a subpoena or court order.” *Id.* at 512.

⁷² FED. R. CIV. P. 26(b)(3). The relevant part of the Rule states:

(3) Trial Preparation: Materials.

(A) Documents and Tangible Things. Ordinarily, a party may not discover documents and tangible things that are prepared in anticipation of litigation or for trial by or for another party or its representative (including the other party's attorney, consultant, surety, indemnitor, insurer, or agent). But, subject to Rule 26(b)(4), those materials may be discovered if:

- (i) they are otherwise discoverable under Rule 26(b)(1); and
- (ii) the party shows that it has substantial need for the materials to prepare its case and cannot, without undue hardship, obtain their substantial equivalent by other means.

(B) Protection Against Disclosure. If the court orders discovery of those materials, it must protect against disclosure of the mental impressions, conclusions, opinions, or legal theories of a party's attorney or other representative concerning the litigation.

⁷³ *Id.*

⁷⁴ *Id.*

impressions, conclusions, opinions or legal theories of an *attorney* or the party's *representative*, the plain text of the Rule does not extend that protection to a *party*.⁷⁵ Gundlach and Smith have remarked on how this difference could impact self-represented litigants.⁷⁶ In 2023, the Eastern District of Pennsylvania observed that “federal and state courts are in disagreement” over whether a self-represented litigant can claim work-product protection for opinion work-product.⁷⁷

While Federal Rule of Civil Procedure 26(b)(3) establishes work-product protection in the civil litigation context, the Supreme Court has observed that the work-product doctrine can also apply in the criminal context.⁷⁸ Specifically, in *U.S. v. Nobles*, the Supreme Court stated that, “[a]lthough the work-product doctrine most frequently is asserted as a bar to discovery in civil litigation, its role in assuring the proper functioning of the criminal justice system is even more vital.”⁷⁹

In justifying the work-product doctrine in criminal proceedings, the Supreme Court similarly focused on the importance of protecting the work-product of attorneys.⁸⁰ Specifically, the Court stated, “at its core, the work-product doctrine shelters the mental processes *of the attorney*, providing a privileged area within which he can analyze and prepare his client's case.”⁸¹ The Court also described the importance of protecting the opinion work-product of representatives and agents of the attorney, though it did not address work-product of the actual defendant.⁸²

Fed. R. Crim. Proc. 16(b)(2)(A) protects a criminal defendant from having to make certain disclosures.⁸³ Specifically, the Rule protects “reports, memoranda, or other documents made by the defendant, or the defendant's attorney or agent, during the case's investigation or defense[.]”⁸⁴ Relying on the text of the Federal Rule, one district court considered “whether a criminal defendant can claim protection via the work product doctrine for documents and notes created, by himself, in

⁷⁵ *Id.*

⁷⁶ Gundlach & Smith, *supra* note 7, at 71.

⁷⁷ *Johnson v. SmithKline Beecham Corp.*, 2023 WL 8817639, at *1 (E.D. Pa. Dec. 5, 2023).

⁷⁸ *See U.S. v. Nobles*, 422 U.S. 225, 236 (1975) (“Respondent contends further that the work-product doctrine exempts the investigator’s report from disclosure at trial. While we agree that this doctrine applies to criminal litigation as well as civil, we find its protection unavailable in this case.”).

⁷⁹ *Id.* at 238.

⁸⁰ *Id.*

⁸¹ *Id.* (emphasis added).

⁸² *See id.*

⁸³ Fed. R. Crim. Proc. 16(b)(2)(A).

⁸⁴ *Id.*

anticipation of litigation,” and decided to “answe[r] that question in the affirmative.”⁸⁵ One commentator has stated that there is a “fundamental unfairness of denying a pro se defendant in a criminal case the protection of [] work product.”⁸⁶

Work-product disputes for self-represented litigants may not arise as often in the criminal context because defendants may have a constitutional right to an attorney when facing imprisonment.⁸⁷ Nevertheless, the issue may still arise where criminal defendants choose to represent themselves, do not face imprisonment, or seek post-conviction relief on a pro se basis.⁸⁸

C. *The Duty of Confidentiality*

Rule 1.6 of the Model Rules of Professional Conduct outlines an attorney’s duty of confidentiality.⁸⁹ This duty is broader than the attorney-client privilege and applies regardless of whether there is a legal proceeding or anticipated litigation.⁹⁰ Model Rule 1.6 dictates that “[a] lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent, the disclosure is impliedly authorized in order to carry out the representation or the disclosure is [otherwise permitted by the Rule.]”⁹¹ Thus, by the text of the rule, the duty of confidentiality binds only attorneys,⁹² though confidentiality can apply in some circumstances to third-parties, such as lawyers who serve as neutral mediators or arbitrator.⁹³

⁸⁵ *United States v. Milk*, 2020 WL 7209763, at *8 (D.S.D. Aug. 14, 2020).

⁸⁶ J. Vincent Aprile II, *The Pro Se Litigant and the Work Product Privilege*, Crim. Just., Fall 2016, at 35, 36.

⁸⁷ *See Gideon v. Wainwright*, 372 U.S. 335, 344 (1963) (“reason and reflection require us to recognize that in our adversary system of criminal justice, any person haled into court, who is too poor to hire a lawyer, cannot be assured a fair trial unless counsel is provided for him. This seems to us to be an obvious truth.”); *see also Alabama v. Shelton*, 535 U.S. 654, 661-62 (2002).

⁸⁸ *See Wainwright*, 372 U.S. at 344; *Shelton*, 535 U.S. at 661-62.

⁸⁹ Model Rules of Pro. Conduct r. 1.6.

⁹⁰ 7 C.J.S. Attorney & Client § 83 (“An attorney’s ethical and contractual duty to maintain client secrets is distinguishable from the attorney-client privilege and may apply in circumstances when the attorney-client privilege is not available. A professional conduct rule regarding client confidences extends to all communications between clients and attorneys and is therefore broader in both scope and protection than the attorney-client privilege.”).

⁹¹ Model Rules of Pro. Conduct r. 1.6(a).

⁹² *See id.*

⁹³ Model Rules of Pro. Conduct r. 1.12. (covering confidentiality for a “Former Judge, Arbitrator, Mediator or Other Third-Party Neutral”); *see also* Comment 3 (“Although lawyers who serve as third-party neutrals do not have information

D. Cases Involving AI Communications of Parties

The issue of accessing a party's AI communications has surfaced in at least two recent cases: *U.S. v. Heppner* and *Warner v. Gilbarco*.

1. U.S. v. Heppner

First, in *U.S. v. Heppner*, the Southern District of New York held that the government could access a defendant's self-directed AI communications, ruling that the AI communications were not protected by attorney-client privilege or work-product.⁹⁴ *Heppner* is, at the time of this article, an ongoing criminal case where a defendant charged with securities and wire fraud utilized Claude, an AI system from the company Anthropic.⁹⁵

The defendant consulted Claude after he "had received a grand jury subpoena [and] after it was clear with discussions with the government that [he] was the target of [the government's] investigation."⁹⁶ By his own initiative, and not at the urging of counsel, the defendant "prepared reports that outlined defense strategy, [indicating] what he might argue with respect to the facts and the law[.]"⁹⁷ The government conducted a search at the defendant's home in connection with the defendant's arrest, and seized documents and electronic devices that included the defendant's various conversations with the AI.⁹⁸

The Court ruled that neither attorney-client privilege nor work-product protection shielded these AI conversations.⁹⁹ First, with regards to attorney-client privilege, the Court reasoned that key elements of the privilege were missing. The Court noted that Claude was not an attorney.¹⁰⁰ Further, the Court went on to say that the defendant's conversations were not confidential given that they were revealed to Anthropic, the third-party AI company, whose terms of service disclaimed confidentiality.¹⁰¹ Additionally, the Court stated that the defendant did not communicate with Claude for the purpose of obtaining legal advice.¹⁰² The

concerning the parties that is protected under Rule 1.6, they typically owe the parties an obligation of confidentiality under law or codes of ethics governing third-party neutrals.").

⁹⁴ *U.S. v. Heppner*, 1:25-cr-00503-JSR, 2026 WL 436479 at *1 (S.D.N.Y. Feb. 17, 2026).

⁹⁵ *Id.*

⁹⁶ *Id.*

⁹⁷ *Id.*

⁹⁸ *Id.*

⁹⁹ *Id.* at *4.

¹⁰⁰ *Id.* at *2–3.

¹⁰¹ *Id.*

¹⁰² *Id.* at *3.

Court conceded that this last point was a closer question, but indicated that the AI communications were not made at the direction of counsel.¹⁰³

Second, the Court rejected the work-product argument because the AI documents were not prepared by or at the direction of an attorney.¹⁰⁴ Recognizing that other jurisdictions had extended work-product protection to parties, the Court nonetheless rebuffed such reasoning, stating that other courts have “concluded that the work product doctrine is not limited to materials prepared by or at the direction of an attorney. But that conclusion undermines the policy animating the work product doctrine, which [] is to ‘preserve a zone of privacy in which a lawyer can prepare and develop legal theories and strategy with an eye toward litigation.’”¹⁰⁵ The Court also refused to credit the defendant’s argument that the text of the Federal Rules of Criminal Procedure prevented discovery or inspection of “reports, memoranda, or other documents made by the defendant, or the defendant’s attorney or agent, during the case’s investigation or defense[.]”¹⁰⁶ Rather, the Court stated that the documents were seized pursuant to a lawful arrest, and not pretrial discovery.¹⁰⁷

2. *Warner v. Gilbarco*

In different case, this one civil and not criminal, the Eastern District of Michigan held that a self-represented litigant could assert work-product for their AI communications.¹⁰⁸ In *Warner v. Gilbarco*, the Court reasoned that parties, not just attorneys, are permitted to claim work-product, and that AI systems could be conceptualized as tools.¹⁰⁹

In *Gilbarco*, a defendant issued both requests for production of documents as well as interrogatories to a self-represented litigant in connection with the self-represented litigant’s AI usage in the lawsuit.¹¹⁰

¹⁰³ *Id.*

¹⁰⁴ *Id.* at *3.

¹⁰⁵ *Id.* at *4.

¹⁰⁶ *Id.*

¹⁰⁷ *Id.*

¹⁰⁸ *Warner v. Gilbarco, Inc.*, No. 2:24-CV-12333, 2026 WL 373043, at *4 (E.D. Mich. Feb. 10, 2026) (“Even if this [AI] information were discoverable, it is subject to protection under the work-product doctrine, which Plaintiff is permitted to assert”).

¹⁰⁹ *See id.* at *4 (“ChatGPT (and other generative AI programs) are tools, not persons... [d]efendants’ motion asks the Court to compel [p]laintiff’s internal analysis and mental impressions) (quotations omitted).

¹¹⁰ Plaintiff’s Answers to Defendant’s Second Set of Interrogatories, *Warner v. Gilbarco, Inc.*, No. 2:24-CV-12333, 2026 WL 373043 at *1–3 (E.D. Mich. Feb. 10, 2026); *see also* Plaintiffs Responses to Defendants Second Request for Production of Documents, *Warner v. Gilbarco, Inc.*, No. 2:24-CV-12333, 2026 WL 373043 at *1–3 (E.D. Mich. Feb. 10, 2026).

The requests for production sought the self-represented litigant's queries to AI systems, the responses from the AI systems, and any documents that the self-represented litigant uploaded to the AI systems.¹¹¹

The interrogatories asked the self-represented litigant, among other matters, to “state whether you have used any artificial intelligence application, software, resource or chatbot in connection with this lawsuit in any way, and, if so, please identify them, including the name, web address, date accessed, and brief summary explaining exactly what you have used these tools.”¹¹² Further, the self-represented litigant was asked to identify “queries (i.e. searches) you have entered into any artificial intelligence application” as well as identify all documents, emails and things that you have uploaded to any artificial intelligence application.”¹¹³ The defendant also sought from the self-represented litigant “all responses and output you received as a result of your queries[.]”¹¹⁴

Significantly, the Court rejected the discovery requests for “all documents and information concerning [the self-represented litigant's] use of third-party AI tools in connection with this lawsuit,” and held that this request inappropriately sought the *party's* work-product made in anticipation of litigation.¹¹⁵

In a compelling passage, the Court sustained the work-product assertion by the self-represented litigant and approvingly quoted the self-represented litigant's objections identified in her brief in opposition:

Defendants' motion asks the Court to compel Plaintiff's internal analysis and mental impressions—*i.e.*, her thought process— rather than any existing document or evidence, which is not discoverable as a matter of law. The motion seeks intrusive post-discovery production based on speculation about what might exist in Plaintiff's internal drafting process, untethered from Rule 26 relevance, disregarding the heightened protection afforded to opinion work product, and improperly attempting to manufacture a waiver where

¹¹¹ For example, one relevant document production request provided: “Produce copies of all queries that you have submitted to Chat-GPT, Claude, Gemini, Perplexity or [other AI tools] regarding: your employment [with defendant], your claims in this lawsuit, your communications with [your current and former lawyers], and pleadings or notices that you have filed in this lawsuit.” *Id.* (cleaned up).

¹¹² *See id.*

¹¹³ *Id.*

¹¹⁴ *Id.*

¹¹⁵ *Warner v. Gilbarco, Inc.*, No. 2:24-CV-12333, 2026 WL 373043, at *4 (E.D. Mich. Feb. 10, 2026).

none exists. At its core, Defendants' request is a fishing expedition[.]¹¹⁶

Further, the Court stated that using generative AI involved “*tools, not persons*,” even if they may have administrators somewhere in the background,” so waiver discussions were not germane.¹¹⁷ Additionally, the Court believed that the discovery requests were neither relevant nor proportional.¹¹⁸

The self-represented litigant’s brief in opposition to the defendant’s motion to compel is higher quality than what might normally be expected of a pro se party.¹¹⁹ This raises the possibility of whether the self-represented litigant employed AI to respond to the discovery requests about AI.

E. Federal Rules of Evidence

While work-product impacts discoverability, other rules besides privilege can impact admissibility. For example, Fed. R. Ev. 403 indicates that “[a] court may exclude relevant evidence if its probative value is

¹¹⁶ *Id.* (quotations omitted).

¹¹⁷ *Id.*

¹¹⁸ *Id.*

¹¹⁹ For example, consider the following excerpt from the self-represented litigant’s brief:

Defendants devoted substantial deposition time to speculative, hypothetical concerns regarding Plaintiff’s use of ChatGPT, yet identified no specific withheld, discoverable document. Rather than timely seeking Court intervention during discovery, Defendants now attempt to conflate a third-party drafting tool with an attorney or litigation agent, recasting Plaintiff’s internal drafting activity as attorney communications, ChatGPT prompts as “case evidence,” and the absence of compelled production or a privilege log as grounds for post-discovery relief. That framing mischaracterizes both the record and the governing law[.]

Defendants’ purported ChatGPT issue, the alleged risk of uploading confidential documents into ChatGPT, was already raised, addressed, and resolved by the Court’s October 30, 2025 Order. (ECF No. 66). Defendants now advance a retroactive theory that improperly equates the use of a private drafting tool with disclosure to a third-party, collapsing the protections of Rule 26(b)(3). Plaintiff’s deposition established only that she used a drafting interface to prepare protected work product; it identified no disclosure to an adversary, no communication to any human third-party, and no factual inconsistencies, admissions, contradictions, or waivers, only conjecture, which is insufficient as a matter of law.

Plaintiff’s Opposition to Defendants’ Motion to Compel and for Sanctions, Warner v. Gilbarco, Inc., No. 2:24-CV-12333, 2026 WL 373043 at *3–5 (E.D. Mich. Feb. 10, 2026).

substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.”¹²⁰ In *United States v. Segui*, the defendant searched “when it is legal to kill a person and justifiable homicide.”¹²¹ The Court excluded this search under Fed. R. Ev. 403.¹²² However, the Court upheld the admissibility of searches involving ways to kill another person.¹²³

F. Stored Communications Act

Suppose an adversary subpoenas OpenAI for the contents of a self-represented litigant’s communications with ChatGPT. In this scenario, the Stored Communications Act may operate to prevent the disclosure, if OpenAI is considered a “remote computing service” or “electronic communication service.”¹²⁴ Some courts have invoked the Act to bar or restrict content disclosures in civil lawsuits from covered providers.¹²⁵ Either the covered provider or litigant whose communications are at issue may have standing to quash such subpoena.¹²⁶ Of course, even if a judge blocks discovery through the SCA, the adversary may still be able to obtain such content directly from the self-represented litigant via normal discovery procedures.

¹²⁰ Fed. R. Ev. 403.

¹²¹ *United States v. Segui*, No. 19-CR-188, 2019 WL 8587291, at *10 (E.D.N.Y. Dec. 2, 2019).

¹²² *Id.* at *13 (“The other searches should be excluded either because they bear fewer parallels to the threat made by Mr. Segui and are therefore less probative, such as Mr. Segui’s search about making and buying poisons, or because they suggest a consciousness of guilt in a manner untethered to Mr. Segui’s specific threats, such as articles about Rikers Island, the legality of murder, and justifiable homicide[.]”). One scholar, Chisup Kim, has examined how courts treat the admissibility of Google searches by defendants in criminal cases. Chisup Kim, *Google Searching for the Truth: Examining the Admissibility of Internet Search History*, 19 WASH. J.L. TECH. & ARTS 46, 47 (2024). Kim argues that whether the search occurs before, during, or after the crime is an important consideration in evaluating admissibility. *Id.* at 59.

¹²³ *Segui*, 2019 WL 8587291 at *13.

¹²⁴ See 18 U.S.C. § 2701.

¹²⁵ See *Crispin v. Christian Audigier, Inc.*, 717 F. Supp. 2d 965, 980 (C.D. Cal. 2010) (holding that private Facebook wall posts could be covered under the Stored Communications Act).

¹²⁶ See *id.* at 974 (approvingly citing “at least two district courts [that] concluded that individuals have standing to move to quash a subpoena seeking personal information protected by the SCA.”); See also Fed. R. Civ. Pro. 45(d) (establishing the procedures for objecting to a third-party subpoena and grounds on which the court may quash sustain such objections).

III. THE SIMPLER FIX: FOCUS ON THE WORK-PRODUCT DOCTRINE

This Section offers a simpler fix, at least compared to alternatives, that focuses on the work-product doctrine. Specifically: (A) courts should interpret the Federal Rules of Civil Procedure to permit opinion work-product for self-represented litigants; (B) the Rules should be amended to make this explicit; (C) a discovery management protocol for AI communications is desirable; and (D) as a backstop to work-product, courts should not admit AI communications involving legal assistance under Fed. R. Ev. 403.

A. Interpreting the Current Rules as Permitting Opinion Work-Product for Self-Represented Litigants Who Use AI Tools

The Eastern District of Michigan in *Warner v. Gilbarco* indicated that self-represented litigants enjoy opinion work-product protection for their AI communications.¹²⁷ Courts across the nation who are presented with this issue should invoke *Gilbarco* as persuasive authority from a federal district court that opinion work-product is indeed available to self-represented litigants who use AI for legal matters.

Saying that AI usage does not reflect one's mental impressions seems incomplete. A user chooses which facts are relevant and what questions to pose that are applicable to the legal matter. Further, the AI's responses to the self-represented litigant can spur further mental impressions from the user. Further, the user is provided with AI strategies, but ultimately must endorse one. Even if the user doesn't fully comprehend all (or even most, or even a majority, or even a lower threshold) of what AI counsels on a legal matter, a user likely exercises some judgment as to some part of the legal decision at hand. Additionally, simply because these AI communications are directed to a third-party company, the user should not automatically waive their confidentiality rights, similar to other technologies.¹²⁸

Courts could further take a proactive, as opposed to a reactive, approach by issuing standing orders for AI communications involving self-represented litigants. This approach has the advantage of being implemented immediately without needing to resort to the time-

¹²⁷ *Warner v. Gilbarco, Inc.*, No. 2:24-CV-12333, 2026 WL 373043, at *4.

¹²⁸ See Elizabeth X. Guo, *United States v. Heppner*, Harvard Law Review Blog <https://harvardlawreview.org/blog/2026/03/united-states-v-heppner/> (Mar. 23, 2026) (noting that for attorney-client privilege, "courts do not treat communication with these third-party platforms [such as Gmail, Slack and Apple] as vitiating a client's reasonable expectation of confidentiality.[] It would be anomalous to treat generative AI systems differently.").

consuming process of formal rule modification. Of course, these standing orders must adhere to the Federal Rules.

Courts could also hold, on a case-by-case basis, that interrogatories and documents requests for generalized AI usage could be irrelevant, overbroad, or disproportionate.¹²⁹

B. Amending the Federal Rules of Civil Procedure

Another fix, though more complicated, would involve amending the Federal Rules of Civil Procedure to explicitly clarify that work-product doctrine applies to self-represented litigants' mental impressions, conclusions, opinions, and legal theories. This fix involves two changes.

The first change would make it explicit that a self-represented litigant enjoys protection for their opinion work-product, to the same degree as an attorney, an idea first proposed by Jennifer A. Gundlach and Zeus Smith.¹³⁰ For example, Rule 26(b)(3)(B) could be changed to the following:

(B) *Protection Against Disclosure.* If the court orders discovery of those materials, it must protect against disclosure of the mental impressions, conclusions, opinions, or legal theories of a party, a party's attorney or other representative concerning the litigation.¹³¹

While this adequately protects the communications of a self-represented party to ChatGPT, it arguably leaves discoverable the responses from ChatGPT back to the self-represented party. Given this complication, another change to Rule 26(b)(3)(B) could be made:

(B) *Protection Against Disclosure.* If the court orders discovery of those materials, it must protect against disclosure of the mental impressions, conclusions, opinions, or legal theories of a party, a party's attorney or other representative concerning the litigation. A generative artificial intelligence system is considered such "other representative."

This would protect an AI's legal theories that are generated in response to the self-represented party's statements and prompts. That said, if an AI's

¹²⁹ See *Gilbarco*, 2026 WL 373043, at *4 (holding that the discovery requests were not proportional).

¹³⁰ Gundlach & Smith, *supra* note 7, at 71.

¹³¹ See Fed. R. Civ. P. 26(b)(3)(B). See also Gundlach & Smith, *supra* note 7, at 71 (recommending the phrase "unrepresented party," specifically changing the language to: "(B) *Protection Against Disclosure.* If the court orders discovery of those materials, it must protect against disclosure of the mental impressions, conclusions, opinions, or legal theories of an unrepresented party, a party's attorney or other representative concerning the litigation.")).

response is considered as prepared by a self-represented party, then such a provision may be superfluous.

Alternatively, the language specifying that AI is a “representative” could be included earlier, in Rule 26(b)(3)(A), which would provide even broader work-product protection, as it would not only protect the mental impressions, opinions, conclusions, or legal theories proposed by ChatGPT, but it would specify that “documents and tangible things” created by ChatGPT in anticipation of litigation could not be accessed unless there were a showing of “substantial need” and “undue hardship.” Specifically, Rule 26(b)(3)(A) could be amended as such:

(A) *Documents and Tangible Things.* Ordinarily, a party may not discover documents and tangible things that are prepared in anticipation of litigation or for trial by or for another party or its representative (including the other party’s attorney, consultant, surety, indemnitor, insurer, ~~or~~ agent, or generative artificial intelligence system utilized by that party). But, subject to Rule 26(b)(4), those materials may be discovered if:

- (i) they are otherwise discoverable under Rule 26(b)(1); and
- (ii) the party shows that it has substantial need for the materials to prepare its case and cannot, without undue hardship, obtain their substantial equivalent by other means.

However, this may take matters too far, as it arguably could shield all responses that an AI provides, subject to the “substantial need” and “undue hardship” limitation.¹³² Further, the discussion of AI need not be in the actual text of the Rules, but perhaps in the comments, indicating that an AI can be construed as a party’s “consultant.”

Expanding work-product is not a perfect solution for self-represented litigants who rely on AI for legal advice. If a layperson uses AI to draft a contract, then that might not be considered in anticipation of litigation, leaving the back and forth that went into drafting that contract discoverable. Further, many self-represented litigants are likely unaware that the work-product doctrine exists.¹³³ They may fail to assert the

¹³² See Fed. R. Civ. P. 26(b)(3)(A).

¹³³ See *Woodruff v. City of Pelham*, 1 So. 3d 157, 159 (Ala. Crim. App. 2008) (“The law is well settled that, “[g]enerally, parties acting pro se should be treated as parties represented by counsel are treated. . . . In particular, pro se litigants ‘must comply with legal procedure and court rules.’”) (internal citations omitted); *but see James v. Servicesource, Inc.*, No. CIV.A. 3:07CV317, 2007 WL 4190794, at *2 (E.D. Va. Nov. 21, 2007) (“The Court is mindful of the need to afford due deference to the complaints and motions of pro se litigants. Accordingly, James’ motion to stay will be construed liberally so that his arguments are given appropriate consideration, no matter how inartfully presented.”).

doctrine at the appropriate stage in litigation. While this problem may be alleviated to some extent by pre-printed forms and “know-your-rights” campaigns, a real concern is that the self-represented litigant will fail to argue work-product, even if the doctrine is expanded.¹³⁴

C. Discovery Management Protocol

As such, one regime could be to use a discovery management protocol. For example, at the outset, parties (and their counsel, if they are represented) could have to fill out a form that indicates in what ways they employed generative AI, for initial classification, which could be categorized into red, yellow, and green categories. This triaging could serve as a case-management sorting tool to facilitate judicial workflow.

Red categories could include those communications made from a party or counsel *to the AI* that reflect mental impressions, conclusions, opinions, or legal theories. Further, communications from the AI *back* to the party or counsel that reflect or respond to these inputs also could be conceptualized as part of a “red” category.

In contrast, those communications made to the AI that are purely factual and do not reflect mental impressions, conclusions, opinions, or legal theories, or involve AI responses that merely restate or organize such facts, could be conceptualized as part of the “green” category.

Closer situations could involve a “yellow” category, which might later warrant in camera review by judges, or special masters who could be appointed to resolve the discovery issue. One example of a yellow scenario could be if a user uploads a document containing information that would otherwise be discoverable, and then asks ChatGPT for legal feedback on how to proceed. The user’s query asking for legal feedback, as well as ChatGPT’s response, could be conceptualized as part of the red category, but the uploaded document could be construed as falling within the green category, hence the yellow classification.

Thus, red reflects what is initially classified as opinion work-product and therefore likely protected. Green would reflect factual material provided by a self-represented litigant, or AI responses that merely summarize or organize such facts, and would be ordinarily discoverable. That said, if green materials qualify for factual work-product, then the adverse party would only be able to obtain them upon a showing of substantial need and undue hardship. Yellow classification would be for mixed or difficult to categorize scenarios which may necessitate judicial review. Such an approach has the advantage of being

¹³⁴ If the other side sought materials from the self-represented party, then the pro se litigant would have to invoke Fed. R. Civ. Pr. 26(b)(5), which permits them to withhold certain privileged and trial preparation materials.

practical and also immediately implementable without any changes to the Federal Rules of Civil Procedure. Again, judges could issue standing orders or local rules could be adopted implementing such discovery management protocols. Of course, if this regime is conceptualized as somehow changing presumptions related to discoverability, it might warrant formal amendments to the Rules. Perhaps one day AI could also assist with this color triaging analysis. Alternative protocols could be adopted; the point is that a plan for how to deal with AI communications at the outset is advisable.

D. Exclusion Under Rule 403 as Backstop

An imperfect last resort would be to focus not on discoverability under work-product, but rather admissibility through the Federal Rules of Evidence. Rule 403 permits a trial court judge the discretion to exclude evidence on account of unfair prejudice, among other matters.¹³⁵ Revealing a party's legal strategies could prove quite prejudicial.

Thus, similar to *Segui*, a judge might view communications where a party asks AI for legal information as highly prejudicial and decide to exclude them from evidence. The problem with this approach is that allowing AI communications to be discoverable (and then later ruling on admissibility) can prejudice a self-represented litigant at the settlement phase. If a self-represented litigant must produce highly sensitive AI legal communications, then the adverse party is in a position to exploit that information to obtain a better settlement offer. Further, these AI communications may help the adverse party prevail in summary judgment. As such, in the interests of justice, it makes sense to focus first on discoverability, and then later on admissibility.

IV. ATTORNEY-CLIENT PRIVILEGE IS MORE COMPLICATED

Other solutions, such as attorney-client privilege, may not cleanly handle AI communications in their current state, and there are hurdles in extending the privilege. Ira P. Robbins offers compelling points as to why an AI privilege is unadvisable.¹³⁶ That said, Elizabeth X. Guo provides an intriguing and cogent perspective as to why attorney-client privilege should apply for AI communications, at least in some circumstances.¹³⁷

This Article notes that focusing on work-product doctrine is more easily implementable in the near-term. That said, extending privilege under certain circumstances may be advisable in the interests of justice,

¹³⁵ Fed. R. Ev. 403.

¹³⁶ Ira P. Robbins, *Against an AI Privilege*, HARV. J. L. TECH. DIGEST (Nov. 7, 2025), <https://jolt.law.harvard.edu/digest/against-an-ai-privilege>.

¹³⁷ Guo, *supra* note 128.

particularly because work-product is narrower, as it only applies in anticipation of litigation and thus still leaves many communications discoverable. This Section attempts to make the case for expanding attorney-client privilege before observing drawbacks.

A. Benefits of Expanding Attorney-Client Privilege

Many of the policy rationales undergirding the attorney-client privilege also extend to communications between self-represented litigants and AI.

1. Determining What the Law Is

A primary policy rationale for the attorney-client privilege is encouraging individuals to seek counsel to ascertain what the law is, and thus, how the law might apply to them.¹³⁸ For example, with regards to attorney-client privilege, the Model Rules of Professional Conduct state that “[a]lmost without exception, clients come to lawyers in order to determine their rights and what is, in the complex of laws and regulations, deemed to be legal and correct.”¹³⁹ While ChatGPT may provide erroneous feedback on the law,¹⁴⁰ that does not diminish the layperson’s motivation, which is to understand the law.

2. Full and Frank Communications

A key justification for the attorney-client privilege is to encourage full and frank communications between a client and an attorney.¹⁴¹ A complete picture of a legal matter is essential for an attorney to render competent legal counsel.¹⁴² Suppose a self-represented litigant uses ChatGPT to assist them. Encouraging that individual to provide as many details as possible regarding the legal matter would seem beneficial for ChatGPT to provide the most effective legal analysis.¹⁴³

3. Privilege Has Been Applied to Non-Attorneys and Representatives

Despite the fact that an AI can display certain lawyer-like characteristics, an AI system is not an attorney. However, the privilege has

¹³⁸ Richmond, *supra* note 5, at 385 (“recognizing the privilege encourages the public to seek early legal assistance”).

¹³⁹ Model Rules of Pro. Conduct r. 1.6, Comment 2.

¹⁴⁰ Smith, *supra* note 36.

¹⁴¹ Model Rules of Pro. Conduct r. 1.6, Comment 2. (“The client is thereby encouraged to seek legal assistance and to communicate fully and frankly with the lawyer even as to embarrassing or legally damaging subject matter. The lawyer needs this information to represent the client effectively and, if necessary, to advise the client to refrain from wrongful conduct.”).

¹⁴² *See id.*

¹⁴³ *See id.*

also been extended to various representatives of the attorney, such as an administrative assistant or paralegal.¹⁴⁴ Moreover, it has also been extended to certain representatives of the client.¹⁴⁵

4. *Access to Justice*

Expanding the privilege theoretically improves access to justice, because it allows individuals to consult legal tools to more fully prepare their case. The Wisconsin Supreme Court contemplated that if the attorney-client privilege did not exist, then “everyone would be thrown upon his own legal resources.”¹⁴⁶ A layperson who cannot afford to hire a lawyer, who then relies on ChatGPT for legal insights, faces a difficult double bind if their conversations are not shielded. On the one hand, the layperson could refuse to use an AI system, for fear of incriminating themselves. On the other hand, use of the AI system may assist with strategic insights regarding how to proceed. Perhaps self-represented litigants should not be placed in such a difficult position.

5. *Expectation of Confidentiality*

If a layperson uses ChatGPT and believes that the communications are confidential, then justice may support the contention that the attorney-client privilege should apply. ChatGPT does not always indicate to a user that discussions involving legal matters are confidential.¹⁴⁷ When an attorney fails to communicate to a client that a conversation is subject to the attorney-client privilege, then normally the communication remains privileged.¹⁴⁸ If ChatGPT does not tell a user

¹⁴⁴ See Milich, *supra* note 54; Tex. R. Evid. 503(a)(4).

¹⁴⁵ See Tex. R. Evid. 503(a)(2).

¹⁴⁶ *Harold Sampson Children's Tr. v. The Linda Gale Sampson*, 271 Wis. 2d 610, 629–30 (1975). The court further opined that “[d]eprived of all professional assistance, a man would not venture to consult any skillful person, or would only dare to tell his counselor half his case.” *Id.* at 629 (quotations omitted). Therefore, a person must feel secure that when they give documents or information to an attorney that the “attorney will [not] release the documents to an adversary who will use the documents against the client.” *Id.* Otherwise, “[c]lients aware that an attorney's disclosure waives the privilege may keep critical information from their attorney, thus thwarting the policy of the free flow of information that lies behind the attorney-client privilege.” *Id.* at 629–30.

¹⁴⁷ See Anoo D. Vyas, ChatGPT, “I have a question. If I got into an accident can what I say be used against me right after the car accident? What's the law on that,” (Oct. 19, 2025) (on file with the Duke Law & Technology Review) (the AI did not indicate that the conversation lacked confidentiality).

¹⁴⁸ See Tex. R. Evid. 503(a)(5) (noting that “[a] communication is ‘confidential’ if not intended to be disclosed to third persons other than those: (A) to whom disclosure is made to further the rendition of professional legal services to the client; or (B) reasonably necessary to transmit the communication.”).

about the lack of confidentiality, then perhaps the matter should indeed be treated as confidential. However, it could be argued that confidentiality disclaimers exist in the terms of service. That said, users may not read these terms carefully, if at all.

6. *The Privilege Naturally Applies Both Ways*

With work-product doctrine, arguably, a split analysis must occur. First, are the statements from the layperson to the AI considered work-product? Second, are the statements back from the AI to the layperson construed as work-product? (These inquiries may collapse into one if the statements generated by the AI in response to user input are construed as being “prepared” by the layperson.) The attorney-client privilege has been interpreted to shield communications both from the attorney to the client and vice versa.¹⁴⁹

7. *Litigants Are More Familiar with the Privilege*

Members of the public are more familiar with attorney-client privilege than work-product doctrine.¹⁵⁰ If know-your-rights campaigns publicize that communications with an AI regarding legal matters fall under attorney-client privilege, and courts provides sample forms that help self-represented litigants assert the privilege, that may resonate more with laypersons than similar initiatives involving work-product.

8. *AI Communications are Conceptually Different from Traditional Search Engines*

It seems that no court has held that a self-represented party’s internet or browser search history related to a legal matter is protected by attorney-client privilege.¹⁵¹ Is a query made to a traditional search engine involving a legal matter the same or different from a conversation with ChatGPT? Arguably, the conversation with ChatGPT feels more

¹⁴⁹ *Id.* at 503(b)(1)(A).

¹⁵⁰ See Armen Adzhemyan & Susan M. Marcella, “Better Call Saul” If You Want Discoverable Communications: The Misrepresentation of the Attorney-Client Privilege on Breaking Bad, 45 N.M. L. REV. 477, 478-79 (2015) (observing how the crime-fraud exception would serve to vitiate the privilege in the popular television show Breaking Bad).

¹⁵¹ Some courts have excluded this type of evidence under Fed. R. Ev. 403. See Chisup Kim, *Google Searching for the Truth: Examining the Admissibility of Internet Search History*, 19 WASH. J. L. TECH. & ARTS 46, 58 (2024).

Thus, suppose that hours after a murder, an accused individual conducts a traditional Google search asking, “how many years in prison does someone normally serve for murder?” First, this search may be discoverable, given that it is non-privileged information that could be relevant. Second, the search could be admissible. Arguably it is both non-privileged and relevant to a claim or defense.

“human.”¹⁵² The humanlike-conversation that AI offers may lure an individual into a sense that they are in fact talking to an attorney, or at least some attorney-like persona. ChatGPT often asks follow-up questions, which is different from a traditional internet search engine that simply displays clickable links.¹⁵³ A user may not volunteer certain information at first, but may do so later based on additional questions that the AI poses.

9. *The Privilege Has Evolved with Technological Developments*

The attorney-client privilege has responded to changes in technology.¹⁵⁴ For example, communications between a lawyer and a client may take place via technological solutions provided by third parties, such as Zoom, provided that certain precautions are taken.¹⁵⁵

10. *Breaking a Self-Regulated Monopoly*

Some have asserted that the practice of law is akin to a legal monopoly, where members of the bar formulate special rules that grant them privileges that other professions do not enjoy.¹⁵⁶ When the startup DoNotPay proposed that laypersons wear glasses equipped with AI to guide self-represented litigants through arguing traffic tickets in court, bar associations objected that this assistance would constitute the unauthorized practice of law.¹⁵⁷ Holding that AI communications lack privilege could be perceived as the profession maintaining its hegemony over legal practice.

11. *Sets the Groundwork for Limited Bar Licensure of AI*

The future of legal services may look drastically different several decades from now. Currently, AI systems can pass the bar exam, often scoring in the 90th percentile or greater.¹⁵⁸ Could AI systems one day rival,

¹⁵² See Damian K.F. Pang, *Why Does ChatGPT Feel So Human?*, PSYCHOLOGY TODAY (May 4, 2023), <https://www.psychologytoday.com/us/blog/consciousness-and-beyond/202305/why-does-chatgpt-feel-so-human>.

¹⁵³ See Vyas, ChatGPT conversation, *supra* note 147.

¹⁵⁴ See Eric Shaffer, Legal Issues in Zoom Meetings, AM. BAR. ASS’N. (Aug. 2, 2021), https://www.americanbar.org/groups/law_practice/resources/law-technology-today/2021/legal-issues-in-zoom-meetings/.

¹⁵⁵ See *id.*

¹⁵⁶ For example, non-competes involving attorneys are considered unethical. However, non-competes involving physicians may be enforced, though are subject to scrutiny for reasonableness. Why the difference? One cynical perspective behind the dichotomous treatment is that perhaps it is in an attorney’s financial interest to avoid non-competes.

¹⁵⁷ Allyn, *supra* note 26.

¹⁵⁸ Alexander, *supra* note 38.

or even outperform, actual lawyers in providing legal services? The scenario is worth considering, even if perhaps decades into the future.¹⁵⁹

In such scenario, should AI be granted the full ability to practice law?¹⁶⁰ Or should a limited license be conferred in certain domains, such as completing court forms, reviewing contracts, or counseling within specific practice areas? At a certain point, if AI progresses to a particular level of skill, might it be unethical for licensing entities not to extend such practicing privileges, given the widespread problem of lack of access to counsel? To be clear, there are many reasons why granting AI some limited license to practice at present would be ill-advised.¹⁶¹ Further, testing these systems to ensure they meet particular quality control standards is imperative. Grading how AI responds to a bar exam is a much different task than evaluating how AI may respond to an actual client in need of services.¹⁶² Additionally, AI may reinforce discriminatory practices which may make bar licensure unwise.¹⁶³

B. The Harms of Expanding Attorney-Client Privilege

Professor Ira P. Robbins has offered what seems to be the most robust case to date against an AI privilege, stating that extending privilege would be “neither doctrinally justified nor normatively sound.”¹⁶⁴ This Section explores some disadvantages and complications to expanding attorney-client privilege to AI communications.

¹⁵⁹ See All-In with Chamath, Jason, Sacks & Friedberg: *Trump AI Speech & Action Plan, DC Summit Recap, Hot GDP Print, Trade Deals, Altman Warns No Privacy* at 59:05 (Apple Podcasts, Aug. 1, 2025) (“Why don't we let AI models get bar certified and get medically certified? So if the AI models, it turns out, are actually proving to be more accurate, more thoughtful, more responsive, more reasonable... and they pass the same criteria as one would need to pass to qualify for the bar or to qualify for a doctor certificate, why don't we do that for the AI? If that then happens, then the same privilege accrues to the AI as it does to the individual human that does it. And now if you extrapolate from where that takes us, if we're suddenly giving AI the same sort of privileged rights that we give to privileged humans, where is that going to take us ultimately with respect to the overall rights for AI?”).

¹⁶⁰ See *id.*

¹⁶¹ Charlotin, *supra* note 2.

¹⁶² Would the AI know to ask appropriate questions to identify the correct legal issue? Would the AI handle the matter with professionalism? A host of questions arise.

¹⁶³ See Lorenzo Belenguer, AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry, 2 AI AND ETHICS 771, 771–87 (2022).

¹⁶⁴ Robbins, *supra* note 136.

1. *Poor Advice and the Need for Quality Control*

AI could provide poor legal feedback to the detriment of a layperson. Thus, bar licensing agencies have an understandable motivation to communicate that feedback given by AI is not at the same level as advice from an attorney. Expanding privilege may signal to the public that AI effectively operates at the level of attorneys.

2. *Agency Issues*

Lawyers are agents for their clients.¹⁶⁵ If an AI provides legal feedback to a user, then does the AI become an agent from a legal perspective, and further, one that owes a fiduciary duty to its principal? What would be the scope of this agency relationship? Could the AI enter into settlement agreements on behalf of the layperson? How might such authority be granted, and how might such authority be limited?¹⁶⁶ These types of fiduciary questions would need to be resolved.

3. *Procedural Complexity*

An AI algorithm is often controlled by a company.¹⁶⁷ In the case of ChatGPT, that entity is OpenAI.¹⁶⁸ Imposing a privilege on an entity, as opposed to an attorney, may invite complexities.¹⁶⁹ For example, could the entity behind the AI still use the information for training purposes? Also, existing law does not seem to adequately cover extending privilege to an AI. Thus, in contrast to work-product doctrine, where the presence of a third-party is often contemplated, implementing privilege may be more difficult.¹⁷⁰

4. *Distinguishing Dual-Purpose Communications*

The primary purpose test is one test used by courts to determine whether to shield conversations an attorney has with a client involving

¹⁶⁵ See *Hayes v. Nat'l Serv. Indus.*, 196 F.3d 1252 (11th Cir. 1999).

¹⁶⁶ In *Hayes v. Nat'l Serv. Indus.*, a Georgia court held that a settlement offer accepted by a lawyer was binding on a client, even though the client did not agree expressly to the settlement terms. *Id.* Extending privilege to conversations that an individual conducts with an LLM may make it likelier that a layperson feels comfortable utilizing an LLM for agency-like matters, such as entering into a settlement agreement.

¹⁶⁷ See OpenAI, *Introducing ChatGPT* (Nov. 30, 2022), <https://openai.com/index/chatgpt>.

¹⁶⁸ *Id.*

¹⁶⁹ In normal circumstances the lawyer would be a fiduciary, but the entity (i.e., OpenAI) would simply be providing services in a non-fiduciary capacity.

¹⁷⁰ See generally, Fed. R. Civ. P. 26(b)(3); see also *Hickman v. Taylor*, 329 U.S. 495 (1947).

both legal and non-legal issues.¹⁷¹ Identifying precisely how dual-purpose communications enjoy privilege could be challenging in the AI context.

Relatedly, there may be difficulty in splitting statements that a user provides to ChatGPT. Consider a situation where a user first confesses to a murder and then later inquires about the legal consequences of that murder. Should the confession be shielded given that it is related to legal advice? Or in the interests of justice, should the confession be admitted, but the later conversations involving legal implications of that murder be protected?

5. *Reducing the Need for Lawyers*

If a user knows that their communications with an AI are confidential, then they may not see the need to hire a lawyer in the first place. This could lead to a slippery slope whereby the market to hire attorneys is reduced.

C. *Duty of Confidentiality*

The attorney-client privilege naturally stresses the confidentiality of communications between a client and a lawyer.¹⁷² However, there is also a duty of confidentiality that is distinct from attorney-client privilege and which provides a separate basis to bind an attorney to keep information disclosed by a client as confidential.¹⁷³ The attorney-client privilege applies in a court proceeding to prevent the disclosure of privileged information.¹⁷⁴ The duty of confidentiality, in contrast, exists both inside and outside the courtroom and is imposed upon the attorney.¹⁷⁵ Some have conceptualized it as part of a broader duty of loyalty that the attorney owes the client that emanates out of the fiduciary relationship.¹⁷⁶ The duty of confidentiality would prevent a lawyer from speaking to a journalist, or some other member of the public, about a legal issue

¹⁷¹ See Kim, *supra* note 59, at 708-09.

¹⁷² See *id.*

¹⁷³ Model Rule of Professional Conduct 1.6, Comment 3, provides that, “[t]he principle of client-lawyer confidentiality is given effect by related bodies of law: the attorney-client privilege, the work-product doctrine and the rule of confidentiality established in professional ethics.” Further, the Model Rules state that “attorney-client privilege and work-product doctrine apply in judicial and other proceedings,” but the “rule of client-lawyer confidentiality applies in situations other than those where evidence is sought from the lawyer through compulsion of law.” *Id.* Thus, this independent duty of confidentiality “applies not only to matters communicated in confidence by the client but also to all information relating to the representation, whatever its source.”

¹⁷⁴ See *id.*

¹⁷⁵ See *id.*

¹⁷⁶ See *id.*

disclosed by a client.¹⁷⁷ Thus, while the confidentiality imposed by the attorney-client privilege is in some sense related to the duty of confidentiality, the duty of confidentiality is in fact distinct and could serve as a separate basis to find that a client's communications with an AI should be guarded.¹⁷⁸

Jonah E. Perlin recently has examined how an attorney's use of generative AI intersects with the duty of confidentiality.¹⁷⁹ The duty of confidentiality does not squarely comport with a self-represented litigant's use of generative AI. After all, an attorney is not involved. Rather, an entity such as OpenAI, provides the AI service. As such, extending the duty of confidentiality could face barriers. Further, because the duty of confidentiality is broader than attorney-client privilege, it could operate to shield too much information from disclosure.¹⁸⁰

V. EXPLORING AN AI-SPECIFIC RULE

Rather than applying existing law regarding work-product, attorney-client privilege, or the duty of confidentiality, jurisdictions could implement statutes or rules that separately address this issue. For example, Robbins argues that instead of an AI privilege, "the right mechanisms are statutory privacy rules or platform-based safeguards[.]"¹⁸¹ That said, such rules and procedures still invite complexity.

A. *Prismatic Framing: A Separate "Horse Law" vs. Existing Bodies of Law That Happen to Apply to Horses*

In Judge Easterbrook's famous essay, *Cyberspace and the Law of the Horse*,¹⁸² Easterbrook observed that the law of the internet could be conceptualized in at least two discrete ways. First, internet legal issues could be resolved simply by applying existing doctrines, without the need to create a separate body of "internet law."¹⁸³ On the other hand, the

¹⁷⁷ *See id.*

¹⁷⁸ *See id.*

¹⁷⁹ Jonah E. Perlin, *Client Confidentiality and Generative AI*, 40 HARV. J.L. & TECH. (forthcoming 2027) (on file with the Duke Law & Technology Review).

¹⁸⁰ Model Rule of Professional Conduct 1.6, Comment 3, *supra* note 173.

¹⁸¹ Robbins, *supra* note 136.

¹⁸² Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207 (1996).

¹⁸³ *Id.* at 207-08 ("Lots of cases deal with sales of horses; others deal with people kicked by horses; still more deal with the licensing and racing of horses, or with the care veterinarians give to horses, or with prizes at horse shows. Any effort to collect these strands into a course on 'The Law of the Horse' is doomed to be shallow and to miss unifying principles . . . Far better for most students— better, even, for those who plan to go into the horse trade—to take courses in property, torts, commercial transactions, and the like, adding to the diet of horse cases a

internet could be so qualitatively distinct that it should in fact necessitate discrete rules.¹⁸⁴ Easterbrook observed that in certain domains, the internet was best understood as fitting within existing areas, such as contracts, torts, or civil procedure, but that in others, it was appropriate to demarcate a discrete body of law to govern its complexities.¹⁸⁵ Easterbrook analogized this situation to horses.¹⁸⁶ The legal rules governing horses can often be categorized as falling into existing doctrine, such as contract or tort law.¹⁸⁷ But in some situations, there may in fact be a distinct need to create “Horse Law.”¹⁸⁸

Easterbrook’s piece was written at the dawn of the consumer internet age of the mid-1990s.¹⁸⁹ Decades later, it is worth considering whether there should be a “Law of AI,” or whether AI legal issues can simply fit within existing doctrines.¹⁹⁰ Daniel Solove has stated that “we must be careful to avoid what I call ‘AI exceptionalism’—treating AI as so different and special that we fail to see how the privacy problems with AI are the same as existing privacy problems, just enhanced.”¹⁹¹

There is precedent for laws evolving with advancements in technology. In the mid-aughts, the Federal Rules of Civil Procedure were modified to respond to concerns related to the storage and production of electronically stored information on computers.¹⁹² Various stakeholders

smattering of transactions in cucumbers, cats, coal, and cribs. Only by putting the law of the horse in the context of broader rules about commercial endeavors could one really understand the law about horses.”).

¹⁸⁴ *Id.*

¹⁸⁵ *Id.*

¹⁸⁶ *Id.*

¹⁸⁷ See U.C.C. Art. 2 §9-102 (44), noting that “goods” means “all things that are movable when a security interest attaches . . . [and even includes] the unborn young of animals.”

¹⁸⁸ Easterbrook, *supra* note 182, at 211.

¹⁸⁹ See *id.*

¹⁹⁰ Legal scholars have observed that Easterbrook’s analogy naturally applies to artificial intelligence. See Captain Christopher J. Lin, *AI-Augmented Targeting and Reining in the Law of the Horse*, 129 PENN ST. L. REV. 483, 493–94 (2025) (“While Judge Easterbrook famously noted that general legal principles can apply to novel technologies . . . there are likely gaps in existing law that fail to adequately consider challenges unique to AI-augmented targeting systems.”).

¹⁹¹ Daniel Solove, *Artificial Intelligence and Privacy*, 77 FLA. L. REV. 1, 15 (2025).

¹⁹² Stephanie S. Rojo, *New Federal Rules Address “Electronically-Stored Information”*, THOMPSON COE (Jan. 15, 2007), <https://www.thompsoncoe.com/resources/publications/new-federal-rules-address-electronically-stored-information> (“On December 1, 2006, several changes to the Federal Rules of Civil Procedure went into effect which will impact litigation pending in federal court. Most significantly, several of the changes to the Rules address the preservation

recognized that responding to discovery requests in the age of computer retained data was different than in a pre-computer era.¹⁹³ Accordingly, Rule 34 of the Federal Rules of Civil Procedure was modified to respond to changing technological realities.¹⁹⁴ Furthermore, in the domain of criminal law, the Supreme Court interpreted the Fourth Amendment to guard against unreasonable governmental searches of cell phones in *Riley v. California*.¹⁹⁵ The Court also protected against extensive surveillance monitoring of cell site location information in *Carpenter v. United States*.¹⁹⁶

B. AI Specific Rules for Confidentiality

What might AI-specific legislation regarding confidentiality for legal matters entail? Practically speaking, how might it be enforced? What are other relevant considerations? This subpart discusses these issues.

1. Forcing AI to Flag Conversations That Involve Legal Issues

AI specific legislation could force the AI to flag conversations where a user seeks feedback on a legal matter. Then, companies such as OpenAI could take steps, such as indicating in a database that the conversation involves an issue that might need shielding. This simple act of identifying and segregating a conversation that involves a legal matter could be helpful if an opposing party later seeks discovery regarding that conversation because the issue of confidentiality already has been flagged.¹⁹⁷

and production of ‘electronically stored information,’ or ‘e-discovery,’ which includes employee e-mails, individual computer hard drives, network drives, backup tapes, optical disks, and other electronic data; and the amendments penalize parties who delete—in compliance with company policy or otherwise—electronic data when a lawsuit can be reasonably anticipated.”).

¹⁹³ *Id.*

¹⁹⁴ *See id.*

¹⁹⁵ *Riley v. California*, 573 U.S. 373, 378 (2014) (holding that “the police may, without a warrant, [may not] search digital information on a cell phone seized from an individual who has been arrested.”).

¹⁹⁶ *Carpenter v. United States*, 585 U.S. 296, 297 (2018) (“Tracking a person’s past movements through [cell site location information] partakes of many of the qualities of GPS monitoring considered in *Jones*—it is detailed, encyclopedic, and effortlessly compiled. At the same time, however, the fact that the individual continuously reveals his location to his wireless carrier implicates the third-party principle of *Smith* and *Miller*. Given the unique nature of cell-site records, this Court declines to extend *Smith* and *Miller* to cover them.”).

¹⁹⁷ In the absence of a comprehensive federal statute that preempts state laws, then determining which state law applies, and how that state’s law might govern the particular conversation, may be difficult. Further, if the conversation starts out with a confidential subject, but then veers toward a non-confidential subject, then

Perhaps an easier solution would be for ChatGPT to warn the user that a particular subject matter seems to involve a legal issue, so they should be aware that the conversations lack confidentiality protections. While ChatGPT sometimes indicates that it is not an attorney and cannot provide legal advice,¹⁹⁸ this declaration could be expanded to inform the user that the conversation is not subject to various confidentiality doctrines, such as attorney-client privilege, work-product, or the duty of confidentiality. Further, ChatGPT could state that the discussion could potentially show up later in a court proceeding.

2. *Storage Limitations*

A different approach would be to command OpenAI not to store the legal conversations permanently. For example, ChatGPT could be forced to delete the conversation immediately after it ends, so that the back-and-forth communication is never stored in a database for later retrieval. This would prevent the conversation from being discoverable or admissible. However, this has the disadvantage of the user not being able to retrieve the material at some later time. Other approaches, such as de-anonymizing a specific conversation from a specific user, may not be effective, given that this again may make it difficult for the user to pull up the conversation at a later time.

One fix would be to warn the user that a specific conversation seemed to implicate a confidential issue, and thus, the conversation would not be stored permanently, and that the user could not access it at a later time. That way, the user could take notes on the conversation for their own reference. Alternatively, the user could be given the option of whether they wished to have the conversation stored permanently.

3. *Removing the Need to Assert Confidentiality*

Self-represented litigants must assert work-product for the doctrine to apply. Given that they are likely unaware that the work-product doctrine exists, they may not effectively exercise their rights, even if work-product is expanded. Perhaps there could be a presumption that the work-product doctrine applies, with the other party needing to rebut the presumption.

it may be difficult to switch back and forth and identify to the user what is or is not covered. For example, the primary purpose test is used in many jurisdictions to extend attorney-client privilege to dual-purpose communications. LLMs would need to be able to navigate this complexity, which may be technologically difficult.

¹⁹⁸ See Smith, *supra* note 37.

4. *Usage Limitations*

AI specific rules could limit how ChatGPT may use confidential information for training purposes.¹⁹⁹ However, limiting the scope of training data may hinder ChatGPT and other AIs from improving.²⁰⁰ Further, too many restrictions on AIs may impede global competitiveness.²⁰¹

5. *One-Way vs. Two-Way Confidentiality*

Should statements made by the self-represented litigant to ChatGPT be protected, and should the shield of privacy also apply to statements generated by ChatGPT in response? An AI specific rule that is aimed at protecting self-represented litigants should shield both.

6. *Identifying Exceptions to Confidentiality*

If a client reveals to an attorney that they are about to commit a crime, or engage in fraud, then the attorney's obligation of confidentiality does not apply.²⁰² Further, the attorney may face a duty to report the information to authorities depending on the jurisdiction and if the

¹⁹⁹ Currently, OpenAI's terms of use do allow for training data to be based on user input. *How ChatGPT and Our Foundation Models Are Developed*, OPENAI, <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed> (last visited March 14, 2026).

("OpenAI's foundation models, including the models that power ChatGPT, are developed using three primary sources of information: (1) information that is publicly available on the internet, (2) information that we partner with third parties to access, and (3) information that our users, human trainers, and researchers provide or generate.").

²⁰⁰ *See id.*

²⁰¹ Some assert that the United States and China are vying for AI primacy. Dewardric L. McNeal, *Silicon Valley is Dictating U.S. AI Policy*, CNBC (Aug. 8, 2025), <https://www.cnbc.com/2025/08/08/silicon-valley-dominant-us-ai-policy-will-benefit-china.html> (observing that "China's approach is not without risk to its users. Its AI systems often come bundled with surveillance tools, opaque financing, and long-term strategic dependencies. But for many global south countries, these are acceptable trade-offs for tangible benefits. Beijing is meeting rising demand for AI access at cost, at scale, and at speed. In doing so, it is turning AI access into a powerful instrument of soft power, one that fuses technological exports with influence over digital norms, governance preferences, and strategic alignment. By contrast, the U.S. approach remains overly anchored in Silicon Valley's commercial interests. American firms continue to lead in foundational models and generative AI, but the strategy often fails to connect technological excellence with global deployment or development impact. Without an accessible, adaptable model that serves diverse needs, the U.S. risks ceding the AI frontier to China by default.").

²⁰² *See, e.g.*, TEX. R. EVID. 503(d)(1).

threatened conduct is sufficiently severe.²⁰³ How might the crime-fraud exception to privilege square with ChatGPT communications?

7. *Enforcement*

Another key issue is determining how to enforce the obligations of confidentiality. Should individuals be able to bring a private cause of action against AI entities? Or should the issue be left to agency enforcement, such as by the Federal Trade Commission? Are money damages appropriate, or should there simply be injunctive relief?

8. *Calls and Wrappers*

Many AIs have application programming interfaces (APIs) where the AI functionality is called by another entity, which then provides certain additional services.²⁰⁴ What might be the confidentiality obligations of these so-called “AI wrappers?”²⁰⁵

9. *Identifying the Appropriate Jurisdiction*

Users may access ChatGPT from a variety of jurisdictions. A question exists regarding what law should apply, particularly if an individual accesses ChatGPT from out of the forum state or out of the country.

VI. PHILOSOPHICAL & SOCIOLOGICAL JUSTIFICATIONS FOR AI CONFIDENTIALITY

This Section explores both philosophical and sociological justifications for AI confidentiality in legal matters.

A. *The Veil of Ignorance*

The contemporary moral philosopher John Rawls proposed using a “veil of ignorance” as a mechanism to evaluate whether laws are morally just.²⁰⁶ The veil of ignorance is a thought experiment where each

²⁰³ Jeffrey Segal & Michael Sacopulos, *The Emotionally Liable Client: Attorney Duties When a Client Threatens Violence*, 4 ELON L. REV. 55, 56 (2012) (“In most jurisdictions, under the ethics rules, disclosure of a client’s threats against third parties is wholly discretionary and must be based upon an assessment that violence is likely imminent. A few states make this duty mandatory if violence is imminent, but the ethics rules do not supply a civil cause of action in tort against non-compliant attorneys.”).

²⁰⁴ Sander Schulhoff, *Exploring ChatGPT (GPT) Wrappers — What They Are and How They Work* (Nov. 15, 2024), https://learnprompting.org/blog/gpt_wrappers.

²⁰⁵ *See id.*

²⁰⁶ John Rawls, *The Idea of Public Reason Revisited*, 64 U. CHI. L. REV. 765, 770 (1997) (“Thus when, on a constitutional essential or matter of basic justice, all appropriate government officials act from and follow public reason, and when all reasonable citizens think of themselves ideally as if they were legislators

individual does not know their actual position in society.²⁰⁷ For example, they do not know whether they are rich or poor.²⁰⁸ Rather, they live behind this “veil of ignorance.”²⁰⁹ Behind this curtain of uncertainty, then, the individual must hypothesize how they might formulate a law.²¹⁰ Rawls argues that this thought experiment is vital as it forces individuals to consider how laws might be drafted.²¹¹ Central to the thought experiment is the concept of the “difference principle,” which asserts that “the social contract should guarantee that everyone has an equal opportunity to prosper” and that “if there are any social or economic differences in the social contract, they should help those who are the worst off.”²¹²

Whether to extend attorney-client privilege, work-product, or other confidentiality protections to conversations that self-represented litigants have with ChatGPT regarding legal issues can be conceptualized

following public reason, the legal enactment expressing the opinion of the majority is legitimate law. It may not be thought the most reasonable, or the most appropriate, by each, but it is politically (morally) binding on him or her as a citizen and is to be accepted as such.”). *See also* STANFORD ENCYCLOPEDIA OF PHILOSOPHY, <https://plato.stanford.edu/entries/original-position> (Oct 24, 2023) (“Rawls’s idea of the original position, as initially conceived, is his account of the moral point of view regarding matters of justice. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls’s impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of all adult) members of an ongoing society. The point of view of justice is then represented as a general ‘social contract’ or agreement by free and equal persons on the basic terms of cooperation for their society.”).

That said, various critiques of Rawls have been raised. Dr. Giancarlo Panagia, *Tot Capita Tot Sententiae: An Extension or Misapplication of Rawlsian Justice*, 110 PENN ST. L. REV. 286 (2005) (“The solution to equality for Rawls is the principle of difference of individuals and groups. One legal author has interpreted Rawls’ theorem to mean that ‘institutions must be designed in such a manner that the expectations of the least advantaged members of society are maximized.’ Once again, it is clear how Rawls’ myopic view is limited only to institutions and fails to reach individuals. Since Rawls is so consumed by the protection of, and the accruing of benefits to, the least advantaged in society, it is somewhat puzzling why he still posits to accept inequalities within individual communities.”).

²⁰⁷ *Id.* at 285–86.

²⁰⁸ *Id.*

²⁰⁹ *Id.*

²¹⁰ *Id.*

²¹¹ *Id.*

²¹² *Veil of Ignorance*, MCCOMBS SCHOOL BUS. (last visited Mar. 29, 2026) <https://ethicsunwrapped.utexas.edu/glossary/veil-of-ignorance>.

through a Rawlsian framework.²¹³ A Rawlsian lens might help support the idea that laypersons should enjoy confidentiality for legal communications with AI.²¹⁴

Suppose that the veil of ignorance exists, and an individual does not yet know their station in society. Given the staggering problem of access to justice, the individual behind the veil may sensibly assume that there is a decent chance they will not be in a position to procure legal counsel for a particular matter.²¹⁵ Further, if they are not in a position to obtain legal counsel, they may reasonably hypothesize that they will have to resort to some type of workaround to obtain legal information. Given these likely realities, it very well may be the case that they would desire to shield these communications, even if ChatGPT provides incorrect legal advice.

Of course, in a society where everyone has access to legal counsel, then the Rawlsian justification for conferring some scope of confidentiality on these AI communications is reduced substantially. In this scenario, an individual who is behind the veil but knows that they will have adequate access to counsel may fail to see the need to have ChatGPT conversations treated as confidential because they can always consult a lawyer. But the staggering division within our society between those who can afford legal counsel and those who cannot supports the conclusion that a layperson should not receive some lesser degree of confidentiality when attempting to navigate a legal issue.²¹⁶

Of course, this Rawlsian framework presupposes a society where ChatGPT exists. Perhaps it would be beneficial for humanity not to have AI, and some have made this argument, alleging that the pursuit of artificial superintelligence may lead to the destruction of humanity.²¹⁷ The counter to this is that AI, practically speaking, is not going anywhere, so it is necessary to consider their existence when conceptualizing morally just laws.²¹⁸

²¹³ Rawls, *supra* note 206.

²¹⁴ *See id.*

²¹⁵ *See* Page and Farrell, *supra* note 19.

²¹⁶ *See id.*

²¹⁷ Frederik Federspiel, *Threats by artificial intelligence to human health and human existence*, BMJ GLOBAL HEALTH (2023).

²¹⁸ Indeed, there is a game theory component to LLMs in the pursuit of superintelligence that may make it difficult to hit the brakes. If one country chooses not to pursue ASI, but another does, then the former country would be at a competitive and security disadvantage. On the other hand, it should be noted that others are so-called “accelerationists” who believe on balance that the pursuit of ASI will unleash tremendous benefits and opportunities to improve the human condition. *See* Lance Eliot, *AI Doomers Versus AI Accelerationists Locked In Battle For*

B. *The Performed Self*

The sociologist Erving Goffman observed that humans often engage in performances, putting on metaphorical masks, as if they were “onstage,” and then removing them, as if they are “backstage.”²¹⁹ Goffman remarked that while an actor may act a certain way while “on stage,” the actor may behave quite differently when they are “backstage.”²²⁰ Social norms may reinforce our presentation of on-stage identity to the detriment of our backstage inner-self.²²¹ Backstage behavior may promote agency and validation of authentic identity.²²² This framing can provide a useful analogy to help justify AI confidentiality.

Future Of Humanity, FORBES, (Feb. 18, 2025), <https://www.forbes.com/sites/lanceeliot/2025/02/18/ai-doomers-versus-ai-accelerationists-locked-in-battle-for-future-of-humanity/> (“AI accelerationists are said to be optimists. They tend to contend that advanced AI such as AGI or ASI is going to solve humanity’s problems. Cure cancer, yes indeed. Overcome world hunger, absolutely. We will see immense economic gains, liberating people from the drudgery of daily toils. AI will work hand-in-hand with humans. This benevolent AI is not going to usurp humanity. AI of this kind will be the last invention humans have ever made, but that’s good, in the sense that AI will invent things we never could have envisioned. Some liken AI accelerationists to perceiving AI as a bright and shiny liberator, it will jumpstart humankind to greater levels never anticipated, AI is good and will be a promulgator of goodness. AI accelerationists who stridently and proudly carry that banner are often assigned the label of being AI evangelists.”).

²¹⁹ See Erving Goffman, *The Presentation of Self in Everyday Life* (1959).

Legal scholarship suggests that Goffman’s critique of the performed self is especially pertinent to the digital forum. Albertina Antognini & Andrew Keane Woods, *Shallow Fakes*, 128 PENN ST. L. REV. 69, 78 (2023) (defining in lieu of deepfakes the term “shallow flakes” as being “fakes as superficial, commonplace deceptions about one’s self-presentation online. Note that this definition does not turn on intent; we are interested both in actors who intend to deceive, for whatever reason, and those who do not. This definition covers a wide range of image-enhancement techniques, including those that were available before the rise of today’s digital networks, such as airbrushing. Importantly, we begin from the premise, as Erving Goffman has observed, that ‘life itself is a dramatically enacted thing.’ All of us perform a kind of ‘presentation of the self’ every day, often by attempting to enhance our image in subtle, and not-so-subtle, ways.”).

²²⁰ *Id.*

²²¹ See *id.*

²²² See *id.* See also Robert C. Post, *On the Popular Image of the Lawyer: Reflections in A Dark Glass*, 75 CAL. L. REV. 379, 387 (1987) (Goffman provided “the insight that the self in modern society can be understood not as something of substance that actually exists, but rather as a series of performances.” Further, “[t]he character attributed by others to an individual is the result of these performances.” Rather, Goffman himself explained that “[i]n our society the character one performs and one’s self are somewhat equated[.]” To Goffman, the

Much of the justification for doctrines preserving confidentiality of legal communications is that they permit the client to be honest.²²³ Eventually, the client may be forced to put on a presentation to a judge or jury.²²⁴ However, prior to that presentation that occurs “on stage,” the client must be able to indulge in revealing their true selves, such that they are “backstage.”²²⁵

When a self-represented litigant seeks to understand the contours of their legal rights, they should not be forced to engage in a performance out of fear that they may implicate themselves. Rather, legal confidentiality doctrines may provide a forum where a client can discard performance and reveal their true backstage persona.²²⁶ They may tell their

presentation of one’s self can be viewed much like an actor on stage who is aware that the audience is observing.).

²²³ See Model Rules of Pro. Conduct r. 1.6, Comment 2 (noting that the client should feel free to communicate embarrassing matters to the lawyer).

²²⁴ Robert F. Blomquist, *Overinterpreting Law*, 116 PENN ST. L. REV. 1081, 1091–92 (2012) (“Teams in the law are theoretically diverse. Some examples include: a client and his or her lawyers in litigational or transactional settings; a corporation’s inside counsel and its outside counsel; a trial court judge and his or her law clerks; judges who coalesce in supporting a majority, concurring, or dissenting opinion in a case; legislators who support or oppose a particular measure (along with their staff); the chief executive of a nation, state, or political subdivision (along with his or her staff); and government administrative agency officials and their staff. As part of their performative roles in the law, these actors must advance and defend an interpretational situation of the law in the course of performances to achieve litigation victories, transactional successes, preferred legal outcomes, legislative wins, and regulatory objectives.”).

²²⁵ *Id.* Blomquist notes that “[i]n the course of these various roles and legal performances, legal actors must strategically decide upon an interpretational approach for each material issue in the conflict situation. Given the generally conservative nature of legal discourse—with arguments tethered by necessity to texts, precedents, intents, traditions, and public policies—common sense and pragmatic interpretational approaches are likely to have the highest probability of success in advancing and defending a team’s legal situation. However, on occasion, more elaborate and complex legal interpretations may be needed in order to prevail, but complex interpretations risk rejection by legal audiences as over-interpretations.”).

Moreover, it is worth observing that Goffman also denotes the role of the associational self, and that the client thereby gains identity via association. See Raymond H. Brescia, *Social Change and the Associational Self: Protecting the Integrity of Identity and Democracy in the Digital Age*, 125 PENN ST. L. REV. 773, 791–92 (2021) (“The associational self emerges to change society through associations and the movements made up of such associations. But the individual also realizes the self in such associations and movements...”).

²²⁶ See Model Rules of Pro. Conduct r 1.6, Comment 2.

attorney that they have committed illegal and awful deeds.²²⁷ That type of disclosure is necessary so the attorney can assist.²²⁸ A self-represented litigant who lacks access to counsel should not be further ostracized by needing to engage in the art of performance. Moreover, they should not be penalized for failing to realize that revealing their true legal situation to ChatGPT may prove deleterious.

VII. CONCLUSION

Many self-represented litigants may not realize that their communications with AI models are not confidential, or may have no choice but to resort to AI communications given the high cost of legal services. It is imperative that legislatures and courts respond thoughtfully to this issue of social and economic justice.

While this Article explores a layperson's use of AI for legal assistance, it also sets the stage for future discussions on related topics. For example, the public often employs ChatGPT as a therapist.²²⁹ Should these conversations enjoy some degree of confidentiality? Many jurisdictions have a psychotherapist privilege that shields statements made to a licensed professional from being admissible.²³⁰ However, ChatGPT of course is not a licensed human professional, so these intimate

²²⁷ *See id.*

²²⁸ *See id.*

²²⁹ Allison Parshall, *Why AI Therapy Can Be So Dangerous*, SCIENTIFIC AMERICAN (AUG. 13, 2025), <https://www.scientificamerican.com/article/why-ai-therapy-can-be-so-dangerous/>. (“Artificial intelligence chatbots don’t judge. Tell them the most private, vulnerable details of your life, and most of them will validate you and may even provide advice. For this reason, many people are turning to applications such as OpenAI’s ChatGPT for life guidance. But AI “therapy” comes with significant risks.”).

²³⁰ *See* Mass. R. Ev. 503 (“Except as hereinafter provided, in any court proceeding and in any proceeding preliminary thereto, and in legislative and administrative proceedings, a patient shall have the privilege of refusing to disclose, and of preventing a witness from disclosing, any communication, wherever made, between said patient and a psychotherapist relative to the diagnosis or treatment of the patient’s mental or emotional condition.”).

conversations may end up finding their way into court.²³¹ That said, employing ChatGPT as a therapist comes with risks.²³²

Further, while many turn to ChatGPT for legal advice, they also turn to it for medical advice.²³³ The Federal Rules of Evidence provide a hearsay exception for statements made by a declarant for the purposes of medical diagnosis or treatment.²³⁴ Should statements made to ChatGPT be admissible under this hearsay exception? They might be admissible in federal court, as authority indicates that the statements need not be directed at a licensed professional, though that is an unsettled question.²³⁵

This Article also invites a conversation regarding granting AI systems a limited license to practice law, at some future point, if AI progresses to the point where it can offer counsel at a level similar to, or even greater than, human attorneys. While this piece does not argue that such licensure should be granted, or that AI will ever progress to such a level of sophistication, the matter is worth at least discussing.

²³¹ See *id.* The Massachusetts rule defines psychotherapist as follows: “(2) A “psychotherapist” is (A) a person licensed to practice medicine who devotes a substantial portion of time to the practice of psychiatry; (B) a person who is licensed as a psychologist by the board of registration of psychologists or a graduate of, or student enrolled in, a doctoral degree program in psychology at a recognized educational institution, who is working under the supervision of a licensed psychologist; or (C) a person who is a registered nurse licensed by the board of registration in nursing whose certificate of registration has been endorsed authorizing the practice of professional nursing in an expanded role as a psychiatric nurse mental health clinical specialist.”

²³² Allison Parshall, *Why ChatGPT Shouldn't Be Your Therapist*, SCI. AM. (Aug. 13, 2025), <https://www.scientificamerican.com/article/why-ai-therapy-can-be-so-dangerous/>.

²³³ See Harriet Walker et al., *Reliability of Medical Information Provided by ChatGPT: Assessment Against Clinical Guidelines and Patient Information Quality Instrument*, 25 J MED. INTERNET RSCH. e47479 (2023).

²³⁴ FED. R. EV. 803(4).

²³⁵ See *id.* Further, a key policy reason for allowing medical statements is the reliability factor: someone who makes this type of statement presumably wants to get better, so they are incentivized to provide honest information. A person who seeks medical advice from ChatGPT may in fact be more forthcoming than with a human given the cloak of perceived anonymity. See Tracy A. Bateman, Annotation, *Admissibility of Statements Made for Purposes of Medical Diagnosis or Treatment as Hearsay Exception Under Rule 803(4) of the Uniform Rules of Evidence*, 38 A.L.R.5th 433 (1996) (“The rationale for the exception is that a patient has a strong motivation to be truthful about information that will form the basis of his diagnosis and treatment, making statements in this regard inherently trustworthy.”).

Ultimately, the public should be aware that their communications with AI, particularly regarding legal matters, may be discoverable and potentially admissible. Policies regulating this issue should be subject to expansive debate. At a high level, AI communications should not be employed in a punitive manner against self-represented litigants who merely seek to understand the scope of their legal rights. As such, self-represented litigants should be able to claim opinion work-product for their AI communications if applicable. Such a solution is more practicable in the near-term than other options, such as reconfiguring attorney-client privilege, which also may be favored through the lens of Rawlsian justice.