Brennan A. Kato, Cheryl Simbulan Beach & Christopher Han, *Artificial Intelligence and the Illusion of Privacy: Social Norms, Disclosure, and Legal Uncertainty* (MCLE materials, webinar presented for the California Lawyers Association, May 1, 2026).

While the law emphasizes clear rules and procedures, many legal questions surrounding the rapid expansion and evolving use of generative artificial intelligence (AI) tools have yet to be answered. The rise of social media produced a similar disconnect between rapidly evolving social norms and the law's ability to fully grasp those changes, leaving a generation exposed to consequences they neither anticipated nor intended. Judges in several recent cases have concluded that the use of publicly available generative AI platforms can result in the loss of protection under both the attorney-client privilege and the work product doctrine. As courts across the country now reach divergent conclusions on the subject of generative AI and legal confidentiality, a pressing question becomes apparent: Has our doctrine of confidentiality kept pace with how people actually interact with technology in modern society?

The Prevalence of Artificial Intelligence Use

A recent study on the use of generative AI in March of 2024 showed that generative AI platforms receive almost 3 billion visits, with Chat GPT commanding 2 billion visits and 500 million users per month.[1] According to a Pew Research Center survey conducted in June 2025, about 50% of adults under the age of 50 report using artificial intelligence approximately once a day or more often.[2] Generative artificial intelligence has become woven into daily life for millions of people. People turn to AI platforms to draft emails, summarize documents, brainstorm ideas, and, increasingly, to think through complex problems, now including legal ones.

The legal profession itself has not been immune to this shift to artificial intelligence. The share of workers with at least a bachelor's degree who use generative artificial intelligence on the job rose from 20% in 2024 to 28% in 2025, and approximately one in five U.S. workers now

[1] Yan Liu & He Wang, *Who on Earth Is Using Generative AI?* (World Bank Working Paper, 2024).
[2] Pew Research Center, *Americans' Views of Artificial Intelligence* (October 2025).

reports using AI in their work.[3] For clients facing legal trouble, the temptation to use these widely available and normalized tools to simply make sense of a confusing legal situation is entirely understandable. Yet as recent case law demonstrates, this reasonable reflex can carry heavy legal consequences.

A Reasonable Expectation of Privacy

A key point in the legal debate over AI and confidentiality is the concept of "reasonable expectation of privacy." Although it is central to Fourth Amendment jurisprudence and privilege analysis, there is no universally accepted definition of what constitutes a reasonable expectation of privacy. Courts evaluate it on a case by case basis, weighing the specific factual circumstances surrounding the communication at issue.

Part of the nuance in this question comes from the fact that a single interaction with an AI tool does not clearly fit within existing, well-established legal categories. There are many arguments presented for different interpretations: a journal entry, search query, conversation with a confidant, etc. The law has well developed rules for each of these categories individually, but it has no clear guidance for a single act that combines all of them at once.

Looking at the terms of service that govern most public AI platforms, research consistently demonstrates that the overwhelming majority of users do not read these agreements before accepting them. In a study of user behavior when joining a social networking service, 74% of participants skipped the privacy policy entirely, and those who did click through spent an average of only 73 seconds on a document designed to take nearly half an hour to read.[4] The terms explicitly permit the platform provider to collect, retain, and in some cases disclose user inputs to third parties, yet users agree without meaningful engagement with their content.

Meanwhile, social norms around AI use are shifting fastest among younger demographics,

[3] *Id*. Pew Research Center.
[4] Obar, J.A., & Oeldorf-Hirsch, A., *The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services*. Information, Communication, & Society, 23(1), 128-147 (2020).

who are adopting these tools at high rates without a corresponding understanding of the privacy implications. A nationwide survey found that roughly 63% of teenagers use generative AI at least weekly, yet about 40% rarely or never consider avoiding the upload of personal data to these systems.[5] Research on youth and AI similarly indicates that younger users often lack the critical skills to evaluate how their data is being used, even as they engage with generative AI tools with increasing frequency.[6] The law has long held that sharing information with a third party extinguishes any claim of confidentiality over that information. Applying that same third party doctrine to AI interactions produces a result that conflicts with how people actually perceive their use of these tools. Research confirms that users consistently disclose sensitive personal information to chatbots based on an intuitive sense of privacy, not on a careful assessment of the platform's data handling practices.[7]

The way people are interpreting their usage of AI highly contrasts the perspective of the law. People are not using their AI chatbots as casual search engines or social media. They are using them as confidants. Eighty two percent of study participants rated their chatbot conversations as sensitive or highly sensitive, and a significant portion reported disclosing health conditions, financial details, and other personal matters they would describe as information they would not like disclosed.[8] This ambiguity creates a potentially harmful trap for the unsuspecting. Without the proper education, people speak their minds in conversations with AI chatbots, sharing sensitive thoughts and strategies without the fear of it being used against them in court, at least in the general population's subjective experience. As one court recently observed, it may be "entirely reasonable for a person to expect some privacy and confidentiality" when interacting with (AI) tools.[9]

---

[5] Child Welfare League Foundation, *2025 Survey on Taiwanese Children and Adolescents' Digital Privacy in the AI Era* (2025).
[6] Skobo, M., & Sovic, N., *Artificial Intelligence in Education: The Role of AI Chatbots in Developing Critical Literacy Among University Students*. International Journal of Childhood and Education, 4(2), 32-45.
[7] Tran, S., Lu, H., Slaughter, I., Herman, B., Dangol, A., Fu, Y., Chen, L., Gebreyohannes, B., Howe, B., Hiniker, A., Weber, N., & Wolfe, R., *Understanding Privacy Norms Around LLM based Chatbots: A Contextual Integrity Perspective* (2025).
[8] *Id.* Tran
[9] Herbert Smith Freehills*, US Courts Find Privilege Applies to Use of Public AI Tools by Self Represented Litigants*. HSF Kramer Notes (April 7, 2026).

Recent Cases

In the *United States V. Heppner* [10]case, Bradley Heppner, the former CEO of Beneficient, a Texas based financial services company, found himself the target of a federal grand jury investigation into allegations of securities and wire fraud. According to the government, Heppner misappropriated approximately $150 million for his personal benefit before a partner company filed for bankruptcy, resulting in roughly $1 billion in losses to investors[11]. Heppner has denied guilt and contested the charges.

After receiving a grand jury subpoena and knowing he was under federal investigation, Heppner took a step that many would consider entirely reasonable. He turned to Claude, a generative artificial intelligence platform developed by Anthropic, to help organize his thoughts and prepare for his legal defense. Acting on his own initiative and without direction or oversight from his attorney, Heppner used Claude to prepare approximately 31 documents outlining potential defenses and legal strategy. Some of the information he uploaded into the AI tool reflected private and confidential discussions he had already had with his attorney.

When FBI agents executed a search warrant at his home in November 2025, they seized these documents from his electronic devices. Heppner's counsel asserted that the AI generated documents were protected by both the attorney-client privilege and the work product doctrine, adding all 31 documents to his privilege log. Judge Jed S. Rakoff disagreed on both counts.

The court's analysis of attorney-client privilege was supported through three grounds. First, and most fundamentally, Judge Rakoff held that Claude is not an attorney. The attorney-client privilege, the court emphasized, exists to protect a trusting human relationship with a licensed professional who owes fiduciary duties and is subject to discipline. That relationship does not exist between a user and an AI platform.[12] As the court later mentioned, that fact alone disposed of

[10] *United States v. Heppner*, No. 1:25-cr-00503-JSR (S.D.N.Y. Feb. 17, 2026).
[11] Barnes & Thornburg, *Federal Court Rules Documents Prepared Using Public AI Tools Not Protected By Attorney Client Privilege* (February 20, 2026).
[12] Cleary Gottlieb. *Courts Grapple with Privilege Implications of AI*. Cleary Gottlieb Publications (February 27, 2026).

Heppner's claim of privilege. Second, the court found that the communications were not confidential. Judge Rakoff examined Anthropic's privacy policy, which users agree to when using the platform's public consumer services, and found that the policy explicitly discloses that the company collects both user inputs and AI outputs, may use them to train its models, and reserves the right to share data with third parties, including governmental and regulatory authorities.[13] Given these terms, the court concluded that Heppner could have had no reasonable expectation of confidentiality. He had, in effect, consented to the very disclosure he was now trying to prevent. Third, the court noted that Heppner used Claude on his own initiative, not at his attorney's direction. The fact that he later shared the AI generated documents with his attorneys did not retroactively protect the previously stored logs. What matters for privilege analysis is whether the client sought legal advice from the AI tool itself, not whether the outputs were shared with counsel.

Even assuming the documents were prepared in anticipation of litigation, defense counsel concluded that they were not created at counsel's direction and did not reflect counsel's legal strategy at the time they were created. The court reaffirmed that the work product doctrine exists to preserve a zone of privacy in which a lawyer prepares and develops legal theories and strategies. That purpose is not served by a client's independent, unsupervised AI assisted research.[14]

The same week that Heppner was decided, a federal court in Michigan reached a contrary conclusion on the application of the work product doctrine to AI assisted materials. In *Warner v. Gilbarco*[15], a Magistrate in the Eastern District of Michigan considered a motion to compel discovery of the litigant's ChatGPT assisted materials and held that the materials were protected by the work product doctrine. Magistrate Patti reasoned that AI platforms are tools, not persons, and that waiver of work product protection requires disclosure to an adversary, not to software. Compelling discovery of AI assisted drafts, the court stated, would effectively nullify work product

[13] Ballard Spahr. *AI, Privilege, and the Future of Confidentiality in the Workplace and Beyond*. Ballard Spahr Legal Alerts (April 2, 2026).
[14] American Bar Association. *AI Communications are Not Privileged: What United States v. Heppner means for Corporate Clients*. ABA Business Law Section (2026).
[15] *Warner v. Gilbarco, Inc.*, No. 2:24-cv-12333, 2026 WL 373043, at *4 (E.D. Mich. Feb. 10, 2026).

protection across most documents prepared in the present day.

These two decisions reflect fundamentally different characterizations of AI technology. The *Heppner* court treated the AI platform as a third party recipient, evaluating confidentiality by reference to the platform's terms of service. The *Warner* court treated AI as a tool analogous to word processing software, holding that use of such a tool does not constitute waiver. [16] The conflict remains unresolved, and the question of whether consumer AI use waives work product protection has been identified as one that may ultimately require Supreme Court review.

Future Considerations for the use of Generative AI Data as Evidence

Recently, Professor Abelson[17] reviewed the case of *United States v. Rinderknecht*[18] and its implications for the use of Chatbot queries as evidence. Rinderknecht was charged with setting the spark that resulted in the devastating 2025 Palisades Fire in Los Angeles based on evidence found in ChatGPT communication records. Evidence used in the complaint included that Mr. Rinderknecht allegedly queried the software "Are you at fault if a fire is lift [sic] because of your cigarettes." Using records such as these, the Government established probable cause in its criminal complaint that Mr. Rinderknecht had the requisite state of mind when he started the fire. If these types of queries were made to a lawyer, it would be protected by attorney-client privilege, inadmissible in court, and law enforcement would not be permitted to have access to the communication. The case is scheduled to go to court in the near future, and is anticipated to provide more insight into the laws regarding privileged information in relation to the use of generative AI platforms.

Generative AI use has become commonplace in modern society. As social norms increasingly treat AI interactions as private, courts struggle to articulate implications to privacy doctrines, which are rooted in traditional disclosure.

---

[16] *Id*. Cleary
[17] Abelson, L.G., The Multidimensions of AI Chatbots as Evidence, 60 U.C. DAVIS L. REV. (forthcoming 2027).
[18] Unites States v. Rinderknecht, No. 2:25-mj-06103 (C.D. Cal. Oct. 2, 2025).