

AI, Confidentiality, and Privilege: What Every Business Should Know

Description

By [Travis J. McDermott](#)

Artificial intelligence (“AI”) is transforming the legal industry by automating many routine tasks, shifting billing models, and forcing courts to establish new rules on privacy and the use of AI tools. While AI is not yet “replacing lawyers,” it can already empower legal professionals to work more efficiently, shifting their focus toward high-level strategy and advising clients, and handle larger workloads. But AI presents novel challenges as well, including with respect to privilege and the practical limits of machine learning.

A key decision earlier this year highlights the privilege concerns AI presents. In February 2026, the United States District Court for the Southern District of New York issued a notable decision in *United States of America v. Bradley Heppner*, 2025 WL 436479 (S.D.N.Y. Feb. 17, 2026), holding that materials a party in a legal proceeding created using a publicly accessible AI platform were not shielded by either the attorney-client privilege or the work product doctrine.

The case addressed a new and increasingly relevant issue: whether communications between an individual and a public AI tool (here, made in connection with an ongoing criminal investigation) qualify for traditional privilege protections. Generally, the attorney-client privilege protects confidential communications between a client and counsel made for the purpose of seeking or providing legal advice. Its companion protection, the work product doctrine, covers materials prepared by or for an attorney in anticipation of litigation. While both doctrines are subject to limitations, together they are the foundation on which candid communications between attorneys and their clients rest.

In *Heppner*, the defendant faced charges including securities fraud, wire fraud, conspiracy, false statements to auditors, and falsification of corporate records. Following his arrest in November 2025, the F.B.I. executed a search warrant at his residence and seized various materials, including approximately 31 documents reflecting his interactions with Anthropic’s AI platform “Claude.”

The defendant had used Claude to generate written analyses outlining potential defense strategies. Through counsel, he asserted that these materials were privileged, arguing that his inputs included information derived from discussions with his attorneys and that the outputs were intended to facilitate further legal advice. However, his attorneys had not instructed him to use Claude.

The court rejected the privilege claim. It emphasized that communications between non-attorneys, even those concerning legal topics, do not fall within the attorney-client privilege. Because neither the defendant nor Claude was legal counsel, the communications were not protected by the privilege. The court further found that the defendant was not “seeking legal advice” from the AI tool and, importantly, that any expectation of confidentiality was unreasonable given that the communications were made through a third-party, public platform. Indeed, Claude’s privacy policy stated that it (1) collects user inputs, as well as the platform’s outputs; (2) uses inputs and outputs to further train itself; and (3) reserves the right to disclose both inputs and outputs to a “host of third parties,” including “governmental regulatory authorities.”

Despite the defendant sharing his AI-generated documents with his counsel, the court found those documents were not “somehow alchemically changed into privileged [documents] upon being shared with counsel.” *Heppner*, 2026 WL 436479, at *3.

The court similarly declined to extend work product protection. Although the defendant argued that the

materials were prepared in anticipation of litigation, they were neither created by counsel nor at counsel's direction, and did not reflect attorney mental impressions or strategy. As a result, the defendant could not be viewed as acting as an agent of his attorneys when engaging with Claude.

The *Heppner* court did not foreclose the possibility that different facts could lead to a different outcome. It left unresolved whether privilege or work product protection *might* apply where an attorney directs a client's use of an AI tool or where the tool operates under confidentiality obligations that could support a reasonable expectation of privacy.

The decision was expressly limited to the facts of the case, which included the use of a publicly available AI platform which lacked any contractual confidentiality protections. The court did not address whether that analysis might change if a user were to rely on a service that affirmatively guaranteed the confidentiality of inputs and outputs—an issue likely to gain significance as AI tools evolve and become more integrated into legal practice.

Regardless of how future cases may refine this area of law, *Heppner* underscores a clear present risk: unsupervised use of public generative AI tools to analyze legal issues, test strategies, or reinterpret legal advice may result in waiver of privilege over those communications. Entering sensitive information, such as legal theories, attorney communications, or case-specific facts, into such platforms likely renders that information discoverable in litigation or subject to third-party subpoena.

While obtaining a second legal opinion remains an appropriate and protected practice, *Heppner* highlights that using publicly available AI tools—such as those offered by Anthropic, OpenAI, or Google—to obtain that “second opinion” carries substantial risk. To remain protected by privilege, any such second opinion should be derived from privileged communications with a legal practitioner.

This decision represents an early development in what will be a rapidly evolving legal landscape. Courts and legislatures may eventually establish clearer standards governing privilege and AI use, but at present, no such safeguards exist. Accordingly, both attorneys and clients should proceed with caution and assume that unsupervised engagement with public generative AI tools could result in a devastating loss of confidentiality protections.

For more information about managing legal risks associated with AI or developing internal best practices, please contact PSH's [AI Task Force](#).

Date Created

August 4, 2026