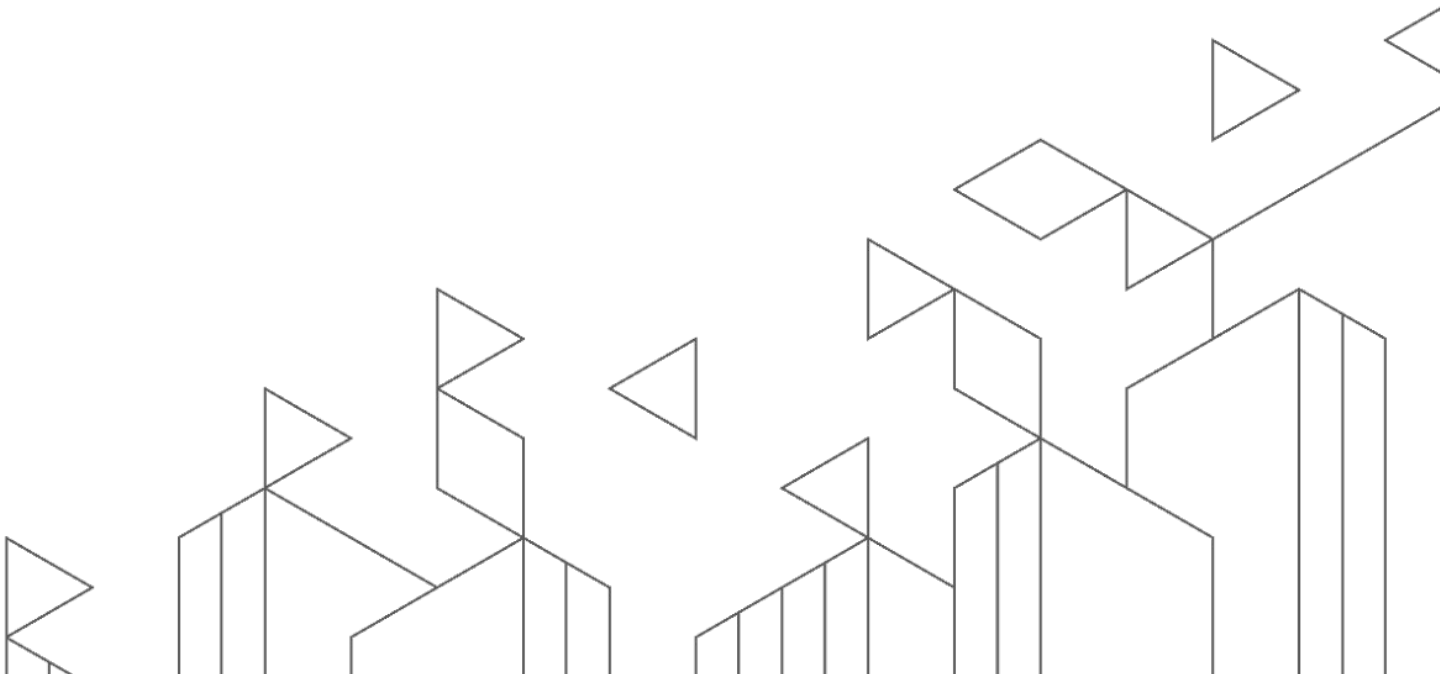
SCALE LLP

AI Privilege

How to Build the Right Architecture



Executive Summary

Generative AI ("Gen AI") is embedded in legal practice. On average, 71% of companies use Gen AI in at least one business function,[1] and 76% of legal professionals in corporate legal departments use Gen AI at least once a week.[2] This trend has no doubt accelerated with the introduction of Claude Legal in early May.

But widespread adoption exposes an unresolved risk: whether these AI conversations are discoverable in litigation or other proceedings. This raises several key questions: Can a user's conversations with Gen AI platforms be protected as privileged under the attorney-client privilege doctrine?

What makes Gen AI different from any other technological platform or tool? What guidance does *United States v. Heppner* provide? Part I below distills the post-*Heppner* analytical framework; Parts II and III build on that framework to provide practical guidance for structuring AI deployments that preserve privilege.

---

1. Gen AI's Broad Reach, McKinsey & Co., https://www.mckinsey.com/featured-insights/week-in-charts/gen-ais-broad-reach (last visited June 29, 2026).

2. Latest Future Ready Lawyer Survey from Wolters Kluwer Finds Broad Adoption of Gen AI by Legal Professionals, LawNext (Oct. 2024), https://www.lawnext.com/2024/10/latest-future-ready-lawyer-survey-from-wolters-kluwer-finds-broad-adoption-of-gen-ai-by-legal-professionals.html.

PART I

The Post-*Heppner* Framework

Attorney-client privilege is a foundational doctrine of law intended to encourage full and frank communication between attorneys and their clients. Judge Jed S. Rakoff's recent decision in *United States v. Heppner* — in a matter of first impression — applied long-standing elements of attorney-client privilege to a client's conversations with the Gen AI platform, Claude.

Judge Rakoff held that a criminal defendant's conversations (consisting of 31 pages of chats) through Claude's consumer product in connection with a pending criminal investigation were not protected under the attorney-client privilege or work product doctrines.

First, what *Heppner* is not. *Heppner* is not a blanket holding that data shared with an AI platform can never be privileged. Nor did the court deviate from a well-established framework on determining whether communications qualify as attorney-client privilege when they are:

(1) between a client and his or her attorney;
(2) kept confidential; and
(3) conducted for the purposes of obtaining or providing legal advice.

Rather, in applying these principles to conversations in a consumer Claude platform, Judge Rakoff held that the AI chats in those circumstances lacked at least two, if not all three elements of the Second Circuit's test for attorney-client privilege:

01
Claude is not an attorney and so the communications were not between Heppner and his counsel;

02
Communications were not confidential because the communications to a "third-party AI platform" were subject to contractual terms in Anthropic's consumer Privacy Policy that permitted training and third-party disclosure;

03
The communications with Claude were not used at counsel's direction for the purpose of obtaining legal advice.

A Gen AI — Third-Party or Tool?

In *Heppner*, Judge Rakoff held that the AI communications failed the first prong because Claude is "not an attorney." Judge Rakoff treated Claude as more akin to a human third party than a passive communication tool. Courts have consistently held that platforms like Gmail and Slack do not, by themselves, destroy privilege.[3] The distinction Judge Rakoff draws is between communicating *with* or *to* Claude versus communicating *through* Claude. In other words, the core distinction is the medium — a passive conduit for privileged communications versus a comprehending recipient of those communications.

We expect this comparison and subsequent arguments will remain prevalent in future court decisions. Indeed, in the same week of the *Heppner* decision, Magistrate Judge Anthony P. Patti in the E.D. of Michigan stated that "ChatGPT (and other generative AI programs) are tools, not persons."[4] The *Gilbarco* case grappled with a fact pattern distinct from *Heppner* in that it involved a pro se plaintiff whose use of ChatGPT to draft litigation materials was protected under the work product doctrine. Critically, *Gilbarco* equated ChatGPT as a tool, rather than a third-party recipient of communications. The divergence matters: if a Gen AI platform is a tool, its use does not implicate the first prong of the privilege test at all. These competing characterizations will likely resurface in future litigation.

Judge Rakoff instead leaves us with a narrower view — privilege may apply to Gen AI communications if "directed by counsel" akin to an "agent" of an attorney as established through the *Kovel* doctrine.

B Confidentiality — Contractual Terms Matter

A key determination in the *Heppner* decision turned on the terms of Anthropic's Privacy Policy as it applied to Claude's consumer products. Judge Rakoff found that Anthropic's then-applicable Privacy Policy permitted Anthropic to train on users' input and output data and reserved the right to disclose such data to a host of third parties, including governmental regulatory authorities. This meant there was "no reasonable expectation of privacy."

Notably, the court's analysis turned on the contractual terms themselves — not on any particular feature within the platform (e.g., a toggle to disable training). While Judge Rakoff left open whether enterprise-grade AI agreements preserve privilege, his decision strongly implies that they could.

---

3. See, e.g., *United States v. Nunez*, 2013 WL 4407069 (S.D.N.Y. Aug. 16, 2013) (Gmail privilege analysis); *United States v. Sharma*, 2019 WL 3802223 (S.D.N.Y. Aug. 13, 2019) (Slack); and *United States v. Sullivan*, 2020 WL 1815220 (D. Haw. Apr. 9, 2020) (iCloud).

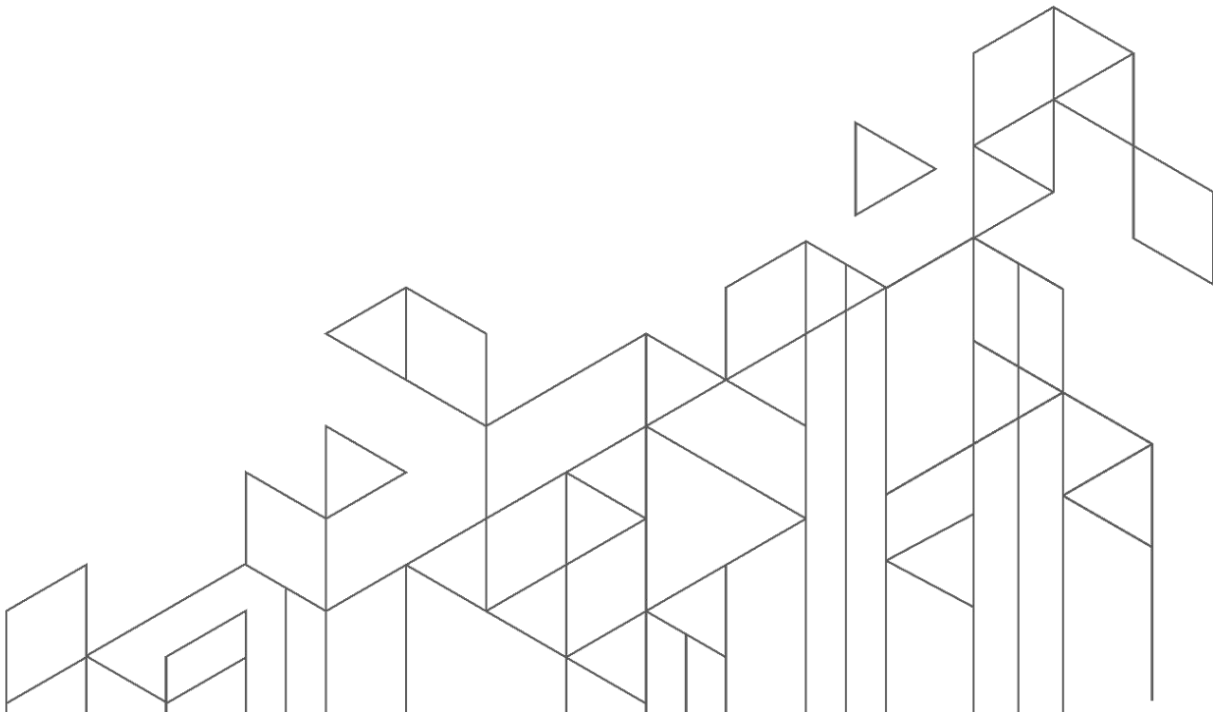
4. Warner v. Gilbarco, Inc., 2026 WL 373043 (E.D. Mich. Feb. 10, 2026) (holding that a pro se plaintiff's use of ChatGPT to draft litigation materials was protected by the work-product doctrine).

C Direction of Counsel / Obtaining Legal Advice

With the third prong — communication for the purpose of obtaining legal advice — Judge Rakoff acknowledged it was the closest call out of the three. Ultimately the decision hinged on Heppner's counsel acknowledging that he did not suggest or direct Heppner to communicate with Claude at counsel's direction. Judge Rakoff recognized, however, that had counsel in fact directed Heppner to use Claude, Claude may have functioned as counsel's agent and preserved privilege.

Citing *United States v. Kovel*, 296 F.2d 918 (2d Cir. 1961), Judge Rakoff indicated that attorney-client privilege could extend to Claude as an attorney's "agent" when the agent's role is to facilitate the attorney's provision of legal advice. The clear example is that of a translator for the attorney or, in the case of Kovel, a tax accountant.

No court has yet extended the Kovel agent theory to AI use, and so the scope of "attorney direction" remains untested. Under Kovel principles, the stronger the showing that an AI tool is "necessary or highly useful" to the attorney's legal advice — coupled with a documented engagement — the more defensible a privilege claim becomes under the *Heppner* framework. In the most cautious approach, those that utilize purposeful, privilege-preserving architectures now, grounded in *Kovel*, will be conservatively positioned to defend discoverability of their AI chats should such a challenge arise.



PART II

Practical Takeaways

A Focus on Key Data Provisions and Consumer Plan Risk

Heppner makes clear that contractual terms are the threshold inquiry for confidentiality — and therefore for privilege. Before any privileged or confidential information enters an AI platform, counsel and clients should confirm that the applicable terms of service do not permit training, third-party disclosure, or other uses incompatible with a reasonable expectation of privacy.

When evaluating any AI tool for use with client confidential or privileged information, clients and attorneys should examine key contractual provisions on data collection and use: training, disclosure, confidentiality, data retention, data deletion, third-party disclosure, and ownership of inputs and outputs. The review should encompass all applicable contractual terms, including any applicable Terms of Service, Privacy Policy, and/or Data Processing Addendums.

The following chart compares the contractual terms of Gen AI platforms, Claude and ChatGPT, evaluated against the *Heppner* confidentiality framework to meet the reasonable expectation of privacy prong. For each platform, the chart identifies the key data provisions with a high-level assessment based on risk and defensibility.

Category	ChatGPT Consumer (Free/Plus/Pro)	ChatGPT Enterprise	Claude Consumer (Free/Pro/Max)	Claude Teams/ Enterprise
Overall Risk Level	**HIGH RISK**	**DEFENSIBLE**	**HIGH RISK**	**DEFENSIBLE**
Data Training	Trains on user data by default; opt-out required via user settings. Feedback used without restriction. *(Terms of Use, "Our use of content")*	No training on customer data by default. Opt-in only via explicit feedback submission. *(Services Agreement §4)*	Data used to train models unless opted out through account settings. Even with opt-out, data used for training when you provide: (1) feedback or (2) if flagged for safety review. *(Consumer Terms §4; Privacy Policy §2)*	No training on commercial customer data. Training permitted for explicitly reported feedback (e.g. via thumbs up / thumbs down feedback button). Admin can disable feedback features. *(Commercial Terms §B)*
Disclosure to Third Parties	Disclosed to vendors, service providers, marketing partners, government authorities or other third parties. *(Privacy Policy §4)*	Access limited to authorized employees for engineering support, abuse investigation, and legal compliance. Contractors bound by confidentiality. *(Enterprise Privacy Page; Services Agreement §7)*	Disclosed to a variety of third parties including service providers, as part of significant corporate event, governmental regulatory authorities, and to third parties in connection with claims, disputes or litigation (based on a good-faith belief that disclosure is reasonably necessary). *(Privacy Policy §3)*	Disclosure limited to employees, agents, and advisors with a need to know, bound by confidentiality obligations at least as protective as the Commercial Terms. Disclosure permitted when required by law or court order. Commercial Terms §E.2, §E.3)
Confidentiality Obligations	None.	Mutual confidentiality obligations. Customer Content designated as Confidential Information. *(Services Agreement §7; DPA §7)*	None.	Express mutual confidentiality obligations. Customer Content is Confidential Information. SOC 2 Type II, ISO 27001, ISO 42001 certified. *(Commercial Terms §E; DPA; Privacy Center)*
Data Retention	Indefinite until user deletes; 30-day backend purge after deletion. No admin controls. *(Privacy Policy §6)*	Admin-configurable (minimum 90 days). *(Enterprise Privacy Page)*	Data retained for as long as reasonably necessary for purposes identified in the Privacy Policy. Deleted chats (30 days); Opt-in training (up to 5 years); Feedback (up to 5 years); Violations (up to 2 years); Trust & Safety flag (up to 7 years) *(Privacy Policy §6; Privacy Center)*	30-day purge on deletion; Feedback (up to 5 years); Violations (up to 2 years); Trust & Safety flag (up to 7 years). *Enterprise Only*: Custom data retention controls with minimum retention period of 30 days. *Enterprise AI Customers*: ZDR available—no storage beyond real-time abuse screening. *(Commercial Terms §6; Privacy Center)*
Data Deletion	Manual deletion; backend removal within 30 days. Data in trained models cannot be removed. *(Privacy Policy §6)*	Admin-controlled. Customer owns all data. DPA provides data subject rights. Deletion/return on termination. *(DPA §5; Services Agreement §11.3)*	Deleted Chats: Immediate removal from chat history; 30-day backend purge. *(Privacy Policy §6; Privacy Center)*	Admin-controlled; 30-day backend purge. DPA requires deletion/return on termination. *(Commercial Terms §E.4; DPA)*

Category	ChatGPT Consumer (Free/Plus/Pro)	ChatGPT Enterprise	Claude Consumer (Free/Pro/Max)	Claude Teams/ Enterprise
Overall Risk Level	**HIGH RISK**	**DEFENSIBLE**	**HIGH RISK**	**DEFENSIBLE**
Ownership of Inputs / Outputs	Customer retains rights to inputs and owns output. *(Terms of Use, "Content")*	Customer owns all inputs and outputs. *(Services Agreement §4)*	Customer retains rights to inputs. Output rights assigned to user. *(Consumer Terms §4)*	Customer retains rights to inputs and owns outputs. *(Commercial Terms §B)*
Governing Terms	**OpenAI Terms of Use** (Eff. Jan 1, 2026) openai.com/policies/row-terms-of-use/ OpenAI Privacy Policy (Eff. Mar 18, 2026) https://openai.com/policies/privacy-policy/	**OpenAI Services Agreement** (Eff. Jan 1, 2026) openai.com/policies/services-agreement/ OpenAI DPA (Eff. Jan 1, 2026) openai.com/policies/data-processing-addendum/	**Anthropic Consumer Terms** (Eff. Oct 8, 2025) anthropic.com/legal/consumer-terms Anthropic Privacy Policy (Eff. July 8, 2026) anthropic.com/legal/privacy Consumer Privacy Center (Eff. Mar 16, 2026) https://privacy.claude.com/en/articles/10023548-how-long-do-you-store-my-data	**Anthropic Commercial Terms** (Eff. Jun 17, 2025) anthropic.com/legal/commercial-terms Anthropic DPA (incorporated into Commercial Terms (Eff. Feb 24, 2026) https://www.anthropic.com/legal/data-processing-addendum Commercial Customers Center (Eff. Mar 16, 2026) https://privacy.claude.com/en/collections/10663361-commercial-customers

The Takeaway Is Clear

At a high level, the takeaway is clear: not to use consumer-tier AI products for client confidential or privileged information. Even paid consumer options are problematic. This applies equally to Claude Free/Pro/Max and ChatGPT Free/Plus/Pro.

Ultimately this outcome hinges on the AI platform's specific contractual provisions on data use and training, though Judge Rakoff expressed heavy skepticism that privilege extends to consumer-grade AI products.

The absence of contractual confidentiality and default training exposure is also incompatible with an attorney's duties under ABA Model Rule 1.6 (Confidentiality of Information) and likely fails to meet the reasonable expectation of privacy prong to afford privilege as examined in *Heppner*.

Enterprise-tier products should be treated as the minimum contractual baseline for any AI use in Gen AI platforms involving privileged or confidential information.

B Attorney Direction is Determinative

Of the three prongs, attorney direction may be the most within counsel's control. In *Heppner*, the privilege claim failed, in part, because defense counsel conceded that he did not direct Heppner to run searches in Claude.

Heppner communicated with Claude of his own volition, not at counsel's direction. As Judge Rakoff acknowledged, however, the outcome could have been different had counsel directed Heppner to use Claude.

In that case, Claude may function in a manner "akin to a highly trained professional" who may act as the lawyer's agent, protecting the AI chats as privileged. The takeaway: to maintain privilege, establish and document a framework for counsel to direct, or ratify, the decision to use an AI platform.

- **First,** counsel should affirmatively instruct the client, in writing, to use a particular AI platform for a specific legal purpose prior to the client's use. Engagement letters or email direction should state that the attorney has directed the client to use X platform to facilitate legal advice. Where possible, the direction should be matter-specific.

- **Second,** couple the AI use to a specific legal purpose. General business queries or searches unrelated to seeking legal advice will not satisfy the third prong. Legal teams, or the product itself, should establish protocols that distinguish between AI use for legal analysis — which may be privilege-eligible — and AI use for general business operations.

- **Third,** maintain a documented record. If the direction is later challenged in discovery, evidence that the AI platform was used for the purpose of obtaining legal advice at counsel's direction will strengthen any privilege claim.

PART III

The Solution: Privilege as a Service

The concern with inputting sensitive communications into Gen AI platforms is warranted. Courts are grappling with how to characterize Gen AI — *Gilbarco* treated ChatGPT as a tool; *Heppner* treated Claude as a third party — and this tension will shape future decisions. *Heppner* left open a path to privilege through the Kovel agent doctrine. Even so, the structural elements for defensibility are identifiable:

An AI product that embeds counsel into the client's AI workspace — under enterprise protections that preserve confidentiality — represents the most defensible architecture for maintaining privilege. Specifically:

01

Provisioned and directed for use by counsel, with documented attorney direction

02

Bound by contractual confidentiality obligations, no-training defaults, and restrictions on third-party disclosure

03

Used by the client for the purpose of obtaining legal advice through the attorney

The privilege landscape is developing in real time, but organizations that build the contractual, operational and documentation infrastructure now will be positioned to defend their AI communications if challenged.

SCALE LLP

www.scalefirm.com

©2026 Scale LLP