ALM | LAW.COM

CORPORATE COUNSEL

Your AI Prompts Aren't 'Thinking.' They're Documents the Other Side Will Ask For.

By Trey Cox & Alan Dabdoub
March 6, 2026

Most in-house teams are debating the wrong AI risk.

It's not whether the tool "hallucinates." It's whether your organization is quietly manufacturing a new category of discoverable records—outside legal's controls—then acting surprised when a court treats those records like evidence.

AI doesn't just accelerate work. It accelerates record creation. And record creation has owners, retention periods, subpoena pathways, privilege fights, and cross-examination consequences.

Two recent litigation flashpoints make the point in a way every GC and head of litigation should internalize—because they're not "AI cases." They're discovery and privilege cases with AI facts.

The Emerging Reality: Courts Are Applying Old Doctrines, Aggressively

1. "AI testing" can become a discovery trap: the numerator/denominator problem

In *Concord Music Group v. Anthropic PBC*, No. 24-CV-03811-EKL (SVK), 2025 WL 3677935, at *1 (N.D. Cal. Dec. 18, 2025), the court treated prompts and related settings used in an investigation as attorney work product—specifically opinion work product—because they reflect counsel's mental impressions about how to interrogate an AI system. That sounds like comfort. It isn't.

Courtesy photos

Alan Dabdoub (L) of Lynn Pinker Hurst & Schwegmann, and Trey Cox (R) of Gibson Dunn.

The operational punchline is what happened next.

The court held that once a party plans to use an AI-driven "investigation" as evidence—especially through witness testimony—fairness requires the opposing party to have what it needs to cross-examine effectively. That meant producing all prompts and outputs used by investigators and client witnesses in a post-suit investigation, including "negative" (unsuccessful) attempts,

because selective disclosure distorts the truth-testing function of cross-examination.

The court's framing is the one that should keep litigation leaders up at night: **"If you're supplying the numerator, the other side is entitled to discover the denominator."**

This is not a narrow copyright quirk. It's a predictable extension of basic trial instincts. If your witness is going to say "our prompts were simple" or "anyone could have done this," the other side gets to explore whether that's accurate—what the successful prompts looked like, what the unsuccessful prompts looked like, how many tries it took, and what was filtered out.

There's also a clean, practical line in the order that matters for how you run internal "testing" going forward:

- Prompts drafted by lawyers but **never used by testifying witnesses** stayed protected.
- Prompts drafted by lawyers but **given to investigators or client witnesses to use** were swept into the waiver created by testimony and had to be produced.

In-house takeaway: the moment you decide to use AI-based testing affirmatively, you are not just "doing analysis." You are committing to a discoverability posture. If you don't control who runs the tests, what gets logged, and how completeness will be defended, you're betting leverage on a process you can't explain.

2. Privilege doesn't attach because you "meant" to talk to counsel later

In *United States v. Heppner* (S.D.N.Y.), Case No. 25-cr-00503, prosecutors filed a motion seeking a ruling that documents generated by a defendant using the AI platform "Claude" were not protected by attorney-client privilege or work product.

The government's argument is straightforward:

- The AI tool is not a lawyer and no lawyer was involved when the documents were created.
- Using a third-party platform undermines confidentiality.
- You can't "retroactively" cloak unprivileged materials by later sending them to counsel.
- Work product generally requires preparation by or at the direction of counsel (or counsel's agent), not a layperson's independent research.

U.S. District Judge Jed S. Rakoff agreed and ruled from the bench during a pre-trial conference that the defendant (a Texas financial services executive accused of a $150MM fraud) could not claim privilege or work product protection over documents he prepared (after he knew he was a law enforcement target) using Claude that he sent to his lawyers. Judge Rakoff's Feb. 17, 2026, written ruling sets forth his rationale for the ruling, which included the fact that Claude contains a written privacy policy to which users of Claude consent that information inputted is not confidential. The government leaned hard on the practical point most businesses ignore until discovery: platform terms, privacy policies, retention, and disclosure pathways can be used to argue diminished confidentiality and waiver.

In-house takeaway: **intent is not a privilege element.** "I was preparing to talk to legal" is not the same thing as a privileged communication. If employees use public AI tools as a substitute for counsel, you can end up with business records and waiver arguments—then spend real money trying to put the toothpaste back in the tube.

Why This Matters to Enterprise Legal Teams Right Now

This is not theoretical.

AI is already embedded across the enterprise—sales, HR, finance, engineering, compliance. The common pattern is predictable: someone is trying to move fast, reduce friction, and look helpful. They paste sensitive facts into a tool, ask for a "legal risk summary," and forward the output to legal with a subject line like "FYI—thoughts?"

That workflow feels informal. A court may treat it as evidence.

The practical risk is not just "discovery cost." It's leverage.

Discovery fights change settlement economics. They create delay. They create credibility questions. And they force board-level conversations at the worst possible time: when leadership wants clarity and you're explaining why you can't confidently answer what exists, where it lives, or what might have been disclosed to a third party.

If you are managing litigation, AI governance is now part of litigation strategy—not because AI is special, but because record creation is consequential.

A Realistic Scenario: How This Becomes an Enterprise Problem Fast

A product issue surfaces with regulatory interest and plaintiff-side noise.

A business leader wants to be proactive and asks the team to summarize exposure, likely plaintiff themes, and "what we should say internally." A manager uses a consumer AI tool on a personal device, pastes internal emails and a draft legal assessment, and asks for (1) "our risk," (2) "likely allegations," and (3) "deposition Q&A."

The output is polished. It looks like progress. It is forwarded to legal.

Then litigation hits. Plaintiffs serve broad discovery. Someone on the other side asks the question that now shows up in serious cases: "Identify and produce materials reflecting use of generative AI relating to the subject matter, including prompts, outputs, settings, and logs."

Now you're living in three bad places at once:

- First, you are fighting about confidentiality and waiver—because the most sensitive content may have been routed through a third-party platform outside legal's control.
- Second, you are fighting about preservation—because AI content can exist in places your ESI map never contemplated (personal devices, browser histories, exported PDFs, vendor-side logs).
- Third, you are fighting about credibility—because witnesses will be asked what they entered, what they received, and whether the AI influenced decisions. If the company has no controlled system and no consistent training, testimony becomes inconsistent and judges become skeptical.

Even if you ultimately prevail, you've already paid: in distraction, spend, delay, and settlement leverage.

What In-House Counsel Should Do Now (Three Moves That Actually Matter)

Most "AI policies" fail because they're written to feel responsible, not to survive discovery.

Here are three moves that will hold up when the fight is real:

1. **Create a legal-grade lane for AI use—and treat everything outside it as an ordinary business record.**

The first order of business should be ensuring that your company, and its law firms, have an enterprise license with frontier AI models. This should prevent the leakage outside the AI model that contributed to the *Heppner* result.

Then, stop pretending you can govern enterprise-wide behavior with a single paragraph that says "don't input sensitive info." Instead, define a controlled lane for legal work and legal-adjacent work (investigations, dispute analysis, regulatory response, high-risk HR matters). In that lane, approved tools, access controls, and retention rules are clear. Outside that lane, assume prompts and outputs are discoverable business records.

That framing is honest—and it prevents the worst failure mode: "everyone assumes privilege later."

2. **Decide, up front, whether you will ever use AI-driven "testing" as evidence—and if yes, plan for completeness.**

If your litigation team wants to use AI testing affirmatively, assume the *Concord* problem: the other side will demand the denominator behind your numerator, especially if a witness will testify about "simplicity," "ease," "typicality," or results.

That means you need a defensible process: who ran the prompts, what logs exist, whether negative attempts are preserved, and how you will demonstrate completeness. If you cannot defend completeness, don't build the case around the testing.

3. **Update your ESI map and legal-hold playbook to treat AI artifacts as a data source.**

If your legal holds don't contemplate prompts, outputs, and logs, you're behind.

You don't need to over-engineer it. You need clarity: where AI usage happens, what systems retain it, what can be exported, what vendors hold, and how it is preserved when a hold triggers. If you can't answer those questions quickly in a discovery conference, you will lose control of the narrative.

The Hard Truth: "Privilege" Is Not a Label you Apply After the Fact

There is a recurring unforced error in enterprise disputes: people use AI as a substitute for counsel, then forward the output to legal and assume it becomes protected.

That assumption is exactly what adversaries will attack. And courts have every incentive to treat the underlying materials like any other third-party communication—especially when fairness and cross-examination are on the table.

If your organization wants AI to be a force multiplier, that's achievable. But it requires a litigation-grade operating model: controlled tools, clear lanes, preservation discipline, and decisions about when AI "analysis" becomes "evidence."

If you don't build that model, the other side will build it for you—in discovery requests, deposition outlines, and motions to compel.

Trey Cox *is co-chair of Gibson Dunn's global litigation practice group and co-partner in charge of the firm's Dallas office.* **Alan Dabdoub** *is a partner in Dallas-based Lynn Pinker Hurst & Schwegmann where he represents plaintiffs and defendants in high stakes commercial disputes.*