SpencerFane®

Prompts and Production: Your Chatbot May Not Be Your Friend in Litigation

Generative artificial intelligence (AI) has quietly become one of the most prolific creators of business records inside modern organizations. Employees draft memos, summarize meetings, analyze data, and brainstorm strategy by typing prompts into chatbots and large language models, and every one of those exchanges leaves a trail. What too few companies appreciate is that the trail is discoverable. As litigation over AI systems and AI-generated content matures, courts have made clear that prompts, outputs, and usage logs are electronically stored information (ESI) like any other, subject to the same rules of preservation and production. The comfortable assumption that a private chat with an AI assistant is ephemeral or off the record is turning out to be wrong. For businesses, the message is direct: your chatbot may not be your friend when litigation arrives.

AI Content Is ESI, and the Old Rules Apply

The starting point is unglamorous but decisive. Courts confronting AI-generated material have not created a special exemption for it. They have applied the familiar framework of the Federal Rules of Civil Procedure, treating prompts and outputs as ESI that is discoverable when relevant and proportional to the needs of the case. Novelty is not a shield. When a category of information bears on a party's claims or defenses, its digital origin in a chatbot rather than an email does not remove it from the reach of discovery.

The most vivid illustration comes from the copyright litigation against OpenAI, where a magistrate judge ordered the preservation and segregation of vast quantities of output-log data and, over objection, cleared the way for production of millions of anonymized ChatGPT logs. The district court left the preservation order standing, and

privacy concerns were addressed through anonymization and protective orders rather than by withholding the material outright.[1] The broader lesson transcends that particular dispute: courts will weigh relevance against privacy and expect de-identification and protective safeguards, not wholesale exemption. If a company's AI records are relevant, the presence of sensitive or personal content will shape how they are produced, not whether.

Businesses should therefore assume that when they are a party to litigation, their internal AI usage – the prompts employees typed, the outputs the model returned, and the logs the system generated – can be requested, and can be ordered produced. Treating those materials as invisible is a planning failure waiting to become a sanctions problem.

Preservation and the Spoliation Trap

The most acute risk is not production but preservation. The duty to preserve relevant evidence attaches once litigation is reasonably anticipated, well before a complaint is filed, and it extends to relevant AI-generated ESI. The problem is that AI tools frequently are not designed to retain data by default. Many consumer and enterprise chatbot configurations delete conversation history automatically, and some offer ephemeral or temporary sessions that vanish by design. A company that anticipates litigation and does nothing to override those defaults may find that relevant prompts and outputs have been destroyed – the classic setup for a spoliation claim.

Avoiding that trap requires affirmative steps that many legal-hold processes have not yet incorporated. When a hold is triggered, counsel must consider whether relevant employees are using AI tools, whether those tools auto-delete, and whether preservation requires disabling deletion settings or exporting conversation data. This is harder than it sounds, because AI usage is often decentralized and informal – an employee may be using a personal or unsanctioned tool that IT does not control and legal does not know about. The preservation duty does not bend to that reality; if the data is relevant and within the organization's control, the obligation stands. Under Rule 37(e), a party that fails to take reasonable steps to preserve ESI it should have kept can face curative measures, and where the loss was intentional, severe sanctions including adverse-inference instructions.[2] AI logs are squarely within that

regime.

Privilege Is Not Guaranteed – and Courts Are Split

Companies sometimes assume that an AI exchange conducted by or for counsel is automatically protected. It is not. Courts have begun to divide on whether and when AI prompts and outputs qualify for attorney-client privilege or work-product protection, and the emerging split should temper any confidence.

A central fault line concerns the use of public, third-party AI tools. Some courts have suggested that feeding otherwise-protected material into a public model can jeopardize protection, on the theory that disclosure to an outside system undercuts confidentiality; this reasoning has surfaced most sharply in the criminal context.[3] Other courts, in civil matters, have declined to find waiver merely because a litigant used a public AI platform, analogizing the tool to ordinary software rather than to disclosure to an adversary.[4] The result is genuine uncertainty. A prompt drafted by a lawyer reflecting litigation strategy may be shielded as work product in one court and exposed in another, and the analysis will turn on the particular tool, the terms under which it operates, and the purpose of the communication.

The practical consequence is that privilege cannot be assumed for AI interactions. Organizations should assume that a chatbot session is presumptively discoverable business information unless it was created under conditions genuinely designed to preserve privilege – counsel involvement, a confidential and controlled platform, and a documented legal purpose. Even then, the protection is contestable, and the safer course is to avoid placing sensitive privileged content into AI tools whose confidentiality posture is uncertain.

The waiver question is closely tied to the terms under which a given tool operates. Enterprise AI deployments frequently include contractual commitments that customer inputs will not be used to train models and will remain confidential; consumer-grade tools often reserve broad rights to retain and use inputs. A court weighing whether confidentiality was preserved may look to exactly those terms, which means the same prompt can carry very different privilege consequences depending on which product an employee happened to open. That reality argues for channeling any AI use that might touch privileged material into vetted,

contractually protected enterprise tools, and for treating public chatbots as unsuitable for confidential work. It also argues for documenting the confidentiality posture of approved tools in advance, so that a privilege claim can rest on more than an after-the-fact assertion.

AI Output as a Business Record

Beyond preservation and privilege lies a category question that many organizations have not confronted: when does an AI output become a business record with its own retention obligations? Where employees rely on AI outputs to make decisions, or where those outputs support audit, compliance, or client-facing functions, the outputs can take on the character of business records. In regulated industries, AI-generated content that informs communications, marketing, or operational decisions may trigger the same documentation, supervision, and retention requirements that apply to other records.[5]

That reframing matters because it shifts AI content from an afterthought to a governed information asset. A financial services firm whose advisers use AI to draft client communications, or a healthcare organization whose staff use AI to summarize records, may have retention and supervision duties that reach the AI layer. Ignoring those duties creates exposure not only in civil discovery but in regulatory examinations. The organizations best positioned are those that understand how their AI systems create, retain, and delete data, and that align those practices with records-management and legal-hold processes rather than leaving AI usage in an ungoverned shadow.

The records-management lens also reframes a risk that is easy to overlook: over retention. Just as failing to preserve relevant AI data invites spoliation claims, indefinitely retaining every prompt and output builds a growing reservoir of material that can be searched, subpoenaed, and used against the organization in future disputes. A hastily typed prompt reflecting an employee's candid assumptions, or a model output that the business declined to follow, may read very differently to a jury than it did to the person who generated it. The goal, therefore, is not maximal retention but deliberate retention – keeping what the business genuinely needs and what the law requires, for defined periods, under a policy applied consistently. A defensible retention schedule, followed in the ordinary course, is itself a form of

litigation protection.

Practical Guidance for Businesses

The path forward is governance, not avoidance. Organizations should begin by inventorying how AI tools are actually used across the enterprise, including unsanctioned "shadow AI" that employees adopt on their own, because a preservation duty cannot be met for data no one knows exists. From that inventory, companies should set retention policies that make deliberate choices about what AI data is kept, for how long, and where—choices that balance operational value, storage cost, and litigation risk rather than defaulting to whatever the vendor's settings happen to be.

Legal-hold procedures should be updated to expressly address AI ESI, with a checklist that prompts counsel to identify AI usage, suspend auto-deletion, and capture relevant prompts, outputs, and logs when a hold attaches. ESI protocols negotiated at the outset of litigation should account for AI-generated material, addressing sources, formats, search methodology, and the privacy and privilege safeguards – protective orders, anonymization, and staged production – that courts increasingly expect. Employee training and acceptable-use policies should tell workers plainly that their AI interactions are not private, may be preserved, and may become evidence, and should steer sensitive or privileged material away from tools that cannot protect it.

None of this requires abandoning generative AI, which delivers real value. It requires treating AI-generated content as what the courts have already decided it is: ordinary ESI, governed by ordinary rules, carrying ordinary risk. The organizations that internalize that reality now – before a preservation demand or a discovery request forces the issue – will be the ones whose chatbots do not become the most damaging witnesses in their own cases.

This article was written by Arnold D. Lee, an attorney in the Phoenix, Arizona office of Spencer Fane. For more information, visit spencerfane.com.

———————————————

[1] *In re OpenAI, Inc. Copyright Infringement Litigation*, No. 25-md-3143 (S.D.N.Y.) (May 13, 2025 preservation order, available at https://cases.justia.com/federal/district-courts/new-york/nysdce/1%3A2023cv11195/612697/551/0.pdf); the consolidated docket is available at https://www.courtlistener.com/docket/69879510/in-re-openai-inc-copyright-infringement-litigation/.

[2] Fed. R. Civ. P. 37(e), available at https://www.law.cornell.edu/rules/frcp/rule_37.

[3] See *United States v. Heppner*, 2026 WL 436479 (S.D.N.Y. Feb. 17, 2026) (holding that a criminal defendant's exchanges with a consumer AI tool were protected by neither the attorney-client privilege nor the work-product doctrine).

[4] See *Warner v. Gilbarco, Inc.*, 2026 WL 373043 (E.D. Mich. Feb. 10, 2026); *Morgan v. V2X, Inc.*, 2026 WL 864223 (D. Colo. Mar. 30, 2026).

[5] See, e.g., FINRA Regulatory Notice 24-09 (confirming that existing recordkeeping, supervision, and communications rules apply to member firms' use of generative AI), available at https://www.finra.org/rules-guidance/notices/24-09.

*Click **here** to subscribe to Spencer Fane communications to ensure you receive timely updates like this directly in your inbox.*