

AI and Privilege – Key Risks Your Organization Needs to Know

Why This Matters Now

The rapid adoption of artificial intelligence tools — from public chatbots and autonomous agents to AI-enabled transcription, translation, and document review platforms — has created a new and often overlooked category of legal risk: the erosion of attorney–client privilege and work-product protection.

AI is not an attorney. Sharing information with an AI tool can waive privilege, and any guidance it provides does not constitute legal advice. When your employees or legal team use AI tools to draft, analyze, summarize, or discuss privileged matters, those interactions can generate a permanent, potentially discoverable record. In some cases, the very act of using AI can destroy privilege that would otherwise have applied. This alert identifies the principal risk areas and sets out practical steps your organization should take now.

Key Privilege Risks from AI Use

There are four distinct categories of privilege risk that arise from the use of AI in a legal or compliance context:

- ▼ **AI prompts and outputs as discoverable, non-privileged records.** Every query you or your company’s employees type into an AI tool — and every response it generates — is a record. Unlike a conversation with your lawyer, these records are not automatically privileged. Prompts submitted to AI systems by non-lawyers can reveal litigation strategy, legal theories, or sensitive business concerns, yet courts are unlikely to protect them unless they were (1) created at counsel’s express direction and (2) within a properly controlled environment. In litigation or a government investigation, opposing counsel or a regulator can request AI prompt histories, chat logs, transcripts, and agent “memory” as part of document discovery. Your organization may have created a documentary record of its legal thinking without realizing it.
- ▼ **Waiver of privilege through use of AI in open or public systems.** Attorney–client privilege and work-product protection can be permanently waived when privileged communications or documents are disclosed to a third party without the client’s control. Public AI platforms, including widely used consumer chatbots and free tools, typically process user inputs on external servers, retain data for extended periods, and may use that data to train future models. Uploading a privileged memorandum, feeding in legal advice, or even describing the substance of confidential communications into a public AI tool can constitute disclosure to a third party, triggering waiver. Privilege lost in this way cannot be restored retroactively. Courts have already begun to hold that communications processed through AI tools with permissive data retention policies do not meet the confidentiality requirements for privilege protection.

- ▼ **Risks of waiver in closed or enterprise AI systems.** Even where an organization deploys a nominally “private” or “enterprise” AI system, privilege risk does not disappear. Enterprise AI tools are often hosted by third-party vendors with their own data access rights, and contractual restrictions on training or data use are only as strong as the vendor’s compliance posture and the technical controls in place. Persistent memory features, audit logs, and usage telemetry within enterprise systems may themselves be discoverable. Additionally, if access controls are inadequate and the AI system is accessible to employees or contractors beyond the privileged group, the “need to know” requirement for privilege may not be satisfied, potentially destroying privilege across all communications in the system.
- ▼ **AI in privileged meetings and proceedings.** AI transcription and meeting assistant tools (including features now embedded by default in video conferencing platforms) create verbatim records of conversations that would otherwise be protected by privilege. If a meeting with counsel is transcribed by an AI tool that sends data to an external server, the transcript is a third-party record — not a protected communication. Similarly, AI tools that silently “listen” to meetings, generate follow-up summaries, or maintain persistent memory across sessions can inadvertently capture and store privileged legal strategy in an uncontrolled format.

What the Courts Are Saying

Courts are now actively addressing issues relating to AI and the attorney-client privilege and work product doctrine, and emerging case law reinforces the need for caution. Key themes from the decisions include:

- ▼ **AI-generated work product can be protected — but only when counsel controls it.** Courts have held that a lawyer’s AI prompts and outputs are opinion work product where they are crafted by counsel and reflect counsel’s legal strategy. This protection extends to using AI as a drafting tool where counsel controls the process. The critical variable is attorney direction and oversight — without it, AI-generated content is unlikely to receive any protection.
- ▼ **Using AI without attorney direction and a confidential system destroys privilege.** A non-lawyer employee who uses AI to discuss or develop a legal position is not automatically entitled to privilege protection. Courts have held that whether privilege attaches turns on two questions: (1) was the AI use directed by an attorney? and (2) does the platform’s architecture ensure true confidentiality or can the system “learn” from inputs in ways that break it? Even employees using a closed enterprise tool can create unprotected, discoverable records if the use was not attorney-directed and the system’s design does not isolate each user’s inputs.
- ▼ **Courts are already restricting AI use in litigation — and your contracts may be next.** At least one court has prohibited parties and their counsel from uploading confidential information into AI systems that store or use inputs for model training or that allow third-party access. This reflects a growing judicial awareness that public AI platforms are incompatible with confidentiality obligations — a constraint that is likely to appear with increasing frequency in litigation hold notices, protective orders, and discovery protocols.

The bottom line: Employees who use AI independently to seek or develop legal guidance are potentially creating unprotected, discoverable records, regardless of whether the tool is public or enterprise-deployed. Privilege protection for such AI-related content turns on two factors: (1) whether counsel directed the AI use, and (2) whether the environment preserved confidentiality. Public AI tools fail on both counts in most scenarios. Even closed systems carry risk if confidentiality cannot be technically guaranteed.

Steps to Minimize Risk

▼ Establish a Clear AI Use Policy for Legal Matters

- Prohibit use of public or consumer AI tools (including free chatbots, browser-based assistants, and third-party transcription apps) for any matter involving legal advice, litigation, investigations, regulatory inquiries, or other privileged communications.
- Require that any AI use touching legal matters be attorney-directed, logged, and conducted within an approved platform subject to documented confidentiality controls.
- Disable AI transcription and meeting assistant features by default in all meetings involving legal counsel, and require explicit legal sign-off before enabling them in sensitive proceedings.

▼ Vet Your AI Vendors and Platforms Before any AI tool is used in connection with privileged or sensitive legal matters, obtain contractual commitments covering:

- A prohibition on using your data to train or improve the vendor's models
- Defined data retention limits and enforceable deletion rights
- Technical security controls, including role-based access, encryption, audit logging, and breach notification obligations
- Customer control over all data sharing and processing settings, with no default sharing with third parties

▼ Train Your People

- Train legal, compliance, HR, finance, and executive personnel on: (i) which AI uses create discoverable records; (ii) the difference between protected and unprotected AI-assisted work; (iii) when to involve counsel before using AI; and (iv) the risk of permanent privilege waiver.
- To reduce an inadvertent privilege waiver, implement technical controls, such as blocking access to non-approved AI platforms on corporate devices.

▼ Prepare for Discovery and/or Retention of AI Records

- Treat AI prompt histories, chatbot outputs, AI-generated transcripts, and agent logs as potentially discoverable documents—and manage them accordingly from the outset of any legal matter. Document retention policies should be reviewed and updated to expressly address what AI-generated content is being retained, for how long, in what format, and by which systems or vendors.
- When a legal hold obligation arises, it extends to AI systems, not just traditional document repositories. Hold notices must expressly identify and cover AI platforms, including enterprise tools, transcription services, meeting assistants, and any system that may have processed, generated, or stored relevant content. Companies should not assume that standard hold procedures automatically capture AI-generated records; in many cases they will not.

AI and privilege is a rapidly developing area of law. Organizations that act now — by establishing clear policies, vetting their tools, and training their people — are far better positioned to preserve privilege and limit discovery exposure in future litigation.

Contact Us

Please contact us to discuss how these risks apply to your organization, or for assistance reviewing your current AI governance policies.



Joni S. Jacobsen

Global Co-Chair and Partner,
Securities and Complex Litigation
New York / Chicago
joni.jacobsen@dechert.com
+1 212 698 3680



Angela M. Liu

Partner,
Securities and Complex Litigation
New York
angela.liu@dechert.com
+1 212 698 3678