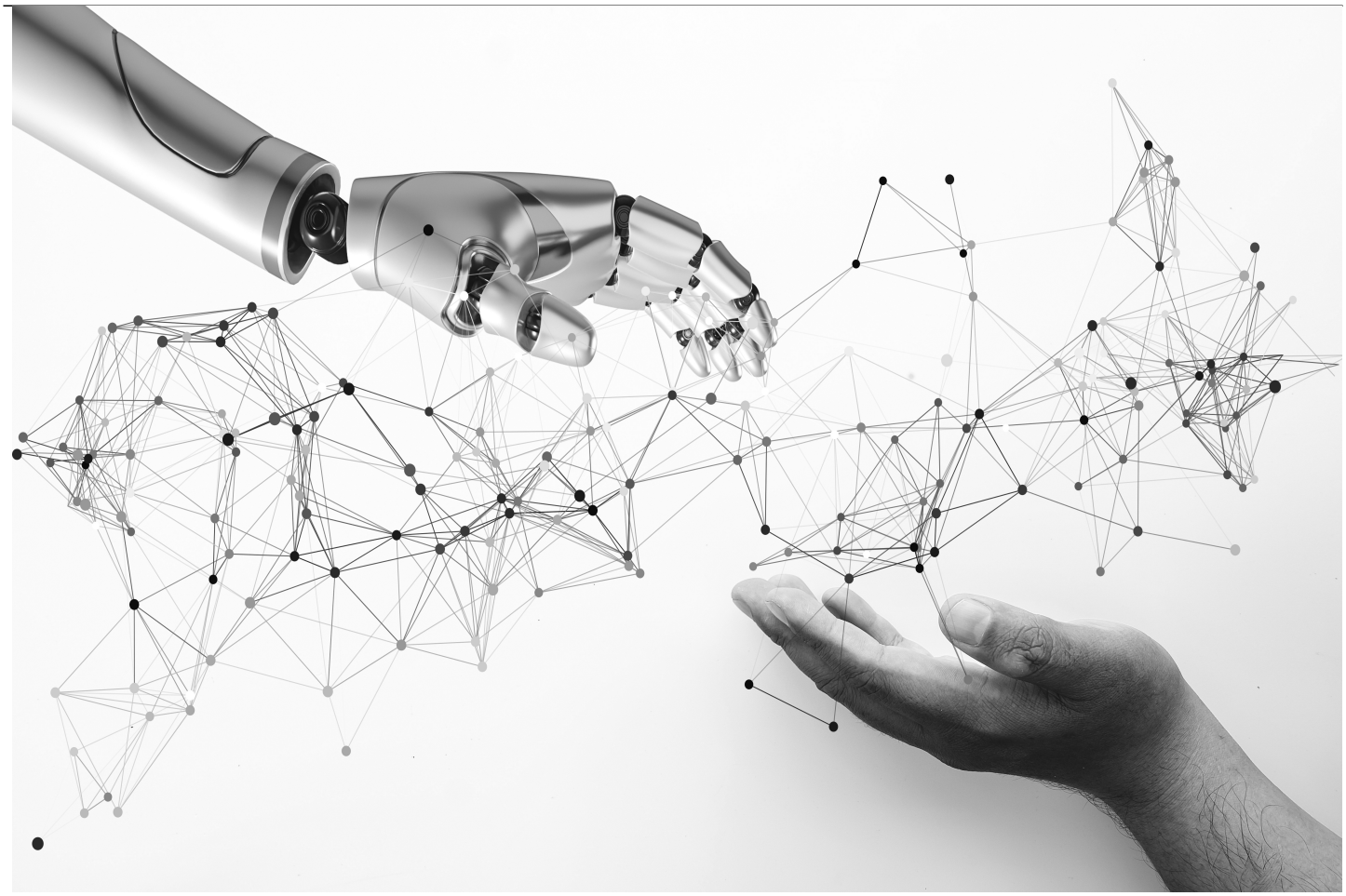
6 Questions You Need to Ask About AI Discoverability and Data Retention

Information Governance

Technology, Privacy, and eCommerce



Banner artwork by Toey Andante / Shutterstock.com

Cheat Sheet

- **Data volume outpaces oversight.** AI tools generate exponential volumes of data, often with no clear view of what’s captured, where it's stored, or how long it’s kept.
- **Retention is a trade?off.** Selective retention is cheaper but risks spoliation; full retention is safer but expands discovery — either way, the choice must be documented and defensible.
- **The legal landscape remains unsettled.** Courts have split on privilege and discoverability for AI?generated content, leaving organizations unsure what’s actually protected.
- **Stakeholders must align early.** IT, legal, external counsel, and business leaders each bring different priorities; only close collaboration produces AI data policies that hold up under scrutiny.

The discoverability of data generated by AI tools and Large Language Models (LLMs) has emerged as a critical issue for businesses. As the technology drives exponential growth in the volume of material potentially subject to discovery, organizations struggle to define policies around preserving, storing, and retrieving it. This challenge is further compounded by recent landmark cases that have reached inconsistent conclusions regarding the admissibility of AI-derived data.

Amid this complex and rapidly evolving environment, legal teams must navigate growing pressure to create and support workflows that effectively capture and preserve relevant data. Keys to success are alignment with business requirements, fulfillment of goals, and assurance of strict compliance by users across the organization.

Multiple stakeholder groups — each with distinct priorities — play a role in defining these policies and ensuring the effective execution of related workflows. Internal IT departments focus on the selection, configuration, and everyday use of tools that determine what data is created and where it is stored. Corporate legal teams and external counsel, meanwhile, emphasize defensibility in litigation and regulatory contexts, including how retention practices may be evaluated. Finally, advisors translate legal and regulatory requirements into practical, business-aligned approaches.

To help teams tackle the challenges posed by AI discoverability, below are six basic questions to consider asking when assessing AI tools for your organization. While the questions are focused on the legal team collectively, each stakeholder group has a different level of responsibility and area of focus for each of the questions.

Start with the basics

1. What data, and how much, are we generating?

This seemingly basic question contains several layers of nuance. For one thing, businesses are constantly rolling out new AI-driven solutions, often without a necessary connection between the departments and users the tool is intended for. This can lead to pockets of AI-generated data escaping governance oversight. Moreover, AI-generated content is generative and changes rapidly to reflect new prompts, summaries, and outputs. As a result, organizations may underestimate the pace at which data is accumulating.

Moreover, most organizations retain more data than they intend to or need. Often, when a legal hold is placed within a platform like Microsoft 365, in addition to capturing “standard” e-mail and documents, that hold will apply to *all* chats, transcriptions, and recordings — including content where the user is a participant, rather than owner or creator. The often-inadvertent over-preservation that results amplifies the AI-generated data volume problem and requires a deeper knowledge of these systems by the IT department than what typically exists now.

As of now, case law has largely glossed over the implications of the ever-increasing volume of data being generated by AI tools. However, future rulings may take a hard line on AI discoverability, which could result in organizations needing to look for data that is difficult to find or that may not exist.

Responsibility for addressing this challenge goes beyond the IT or legal department, and must be managed collaboratively *across* an entire organization. Tools must be selected, configured, and rolled out through a multi-faceted workflow that requires a foundational understanding of what data is being generated, as well as what is actually being captured. Establishing this foundation early can prevent a great deal of pain at later stages.

> Tools must be selected, configured, and rolled out through a multi-faceted workflow that requires a foundational understanding of what data is being generated, as well as what is actually being captured.

2. Where is the AI-generated data stored?

Answering this question can be unexpectedly complex. To illustrate, consider this example: An executive assistant enables recording and transcription for a meeting they have scheduled on behalf of the senior executive they support. They do not attend the meeting. This transcript is saved *not* in the senior executive's OneDrive, but in that of the assistant — who is unlikely to be on legal hold, and even less likely to have their data collected during a litigation or investigation. Such gaps are easy to miss and difficult to remediate after the fact. This is just one example of why a clear mapping and understanding of where AI-generated content is stored must be a baseline exercise, not an afterthought.

3. What data do we retain, and for how long?

Do you retain records of *all* prompts, responses, AI-generated summaries, and other outputs, in the event that, at some point, they *may be* relevant? Or do you assume the risk that only the most recent version is sufficient? The answers to these questions will vary depending on the business type; for example, highly regulated industries, such as pharmaceuticals and financial services, have strict compliance standards and record-keeping obligations.

Trade-offs are often necessary. *Selective retention*, while cheaper and easier to manage, can expose an organization to allegations of evidence spoliation if a prompt is deemed to be relevant and is no longer available. *Full retention* insulates an organization against that risk but creates its own burden and increases the surface area of what is potentially discoverable. Regardless of the situation, the "correct" answer must be a defensible one, defined and supported by the legal team and documented in a manner that ties retention policies to the organization's legal risk profile. Ultimately, the legal team, often in conjunction with external counsel, will need to defend the decisions around how much data to keep and for how long.

> Ultimately, the legal team, often in conjunction with external counsel, will need to defend the decisions around how much data to keep and for how long.

Drilling down to specifics

4. What is the organization's legal risk profile?

Here again, industry context is key. Drug manufacturers must keep detailed records of research and testing. Financial firms must preserve audio recordings of traders making trades, as well as text messages with clients. This suggests that similarly demanding standards may soon apply to the preservation of AI-generated data.

An effective risk strategy for AI data should therefore include clear and detailed definitions around what is being preserved and what is not. For example, how should prompts, queries, updates, and summaries be managed? Are they stored? If so, for how long? What is the destruction policy and how is that determined? Addressing such questions and documenting the company's position can better define criteria around what to retain and why, as well as defensible arguments for what data is discarded and when.

> An effective risk strategy for AI data should therefore include clear and detailed definitions around what is being preserved and what is not.

This process can resemble a traditional risk assessment as applied to, say, the retention of different versions of Microsoft Word documents. What's different now is that the outputs created by generative AI raise entirely new questions and considerations, such as why certain prompts are used, and what those prompts might reveal that is potentially relevant to a discoverability exercise.

A related and increasingly urgent dimension of the risk profile for a given organization is privilege. What constitutes privilege when content is generated by AI is, at this stage, an open question for many. Discovery requests that specifically target prompts and responses generated through tools like Copilot, including prompts that might in another context resemble internal mental impressions or work product, are starting to appear. The line between work product, attorney-client communications, and ordinary-course AI output is a live issue.

Findings on AI-generated data and privilege have to date been inconsistent, suggesting that the courts are still unsettled on the issue, particular given the variability of AI use cases. For example, in United States v. Heppner, a federal court ruled that a defendant's written communications generated by Anthropic's Claude are not protected by attorney/client privilege. In Warner v. Gilbarco, meanwhile, a Michigan federal court ruled that AI-generated queries and responses *were* protected. Finally, Morgan v. V2X, Inc., allowed protection for AI-assisted litigation work, while requiring disclosure of the AI tool used.

That said, the trajectory suggests that AI-generated work product will increasingly be treated as discoverable – unless the privilege analysis has been thought through deliberately and in advance. In any event, rather than taking a reactive posture in response to an action, counsel should be prepared to defend a position and advise their clients on responsible use that would impact privilege calls and consideration of AI-generated data.

5. Does the AI policy do what it promises?

Once a policy for managing AI data is in place, consistent validation and auditing can ensure that outcomes align with objectives. At a basic level, this involves assessing whether enough data, the right data, or unnecessarily large volumes of data are being captured. More specifically, the business

must confirm that it’s retaining the data it intends to retain. Capturing too much unnecessary data, and not enough critical data, signals the need for a reset.

Auditing retention requires technical literacy and a willingness to look at the operational reality at a granular level. Rather than technology problems in the strict sense, policy failures are generally gaps between what users actually do and what the policy assumes they will do. It is the organization’s responsibility to surface those gaps before counsel discovers them in the middle of a discovery dispute.

6. Why is it important to address this now?

Kicking the can down the road can be tempting. However, demand for clarity around how to manage AI-generated corporate data is increasing. RFPs are asking for details on AI prompts and responses, ESI protocols are included scoping requirements on AI tools, and POs are referencing the use of LLMs.

When it comes to traditional sources of data, moreover, courts have historically shown little sympathy for companies that lack awareness with respect to how that data is generated. In the case of AI, organizations that cannot articulate or defend their data retention policies are finding themselves on the wrong side of spoliation arguments before the merits are ever reached.

This uncharted territory compounds the challenges facing legal teams. These tools and platforms did not exist a year ago, and they continue to change rapidly. Litigators are being asked to defend retention practices that have not yet been tested by appellate courts. Organizations are tasked with bridging the constant competing interest of business efficiency and legal risk.

Join ACC for more AI insights!

Tarek Ghalayini

Co-head of Americas Risk Advisory and Global Leader of eDiscovery

AlixPartners

Tarek Ghalayini is a recognized expert in complex litigation matters where he guides and prepares clients for large, cross-border matters involving data mining, analytics-drive document review, and other related data-centric issues. A licensed attorney with a strong background in accounting and finance, Ghalayini works closely with counsel, corporate clients, and financial advisors to quickly and efficiently tackle complex electronic data problems in global litigation, financial investigations, regulatory inquiries, and restructuring matters. He is known for developing and executing effective and straightforward solutions that bring clarity and consistency to cross-border and local matters alike.

Danielle T. Davidson

Head Strategic Advisor, Advanced E-Discovery and AI Strategy Group,

Ropes & Gray LLP

Danielle Davidson is a leader in developing and adopting advanced technology and analytics, she works with clients — including the heads of E-Discovery at global corporations — to create complex E-Discovery strategies. With nearly 20 years of experience, she has deep experience piloting and adopting advanced technologies such as concept analytics and Technology Assisted Review (TAR). Davidson leads projects on short message review strategies, the future applied use of AI, and technological solutions for mobile data content capture and monitoring. Davidson is at the forefront of advising law firms and their clients on the strategic adoption of Generative AI, including its use in early case assessment, investigations, document review optimization, and litigation strategy, with a focus on building scalable, defensible, and auditable workflows.