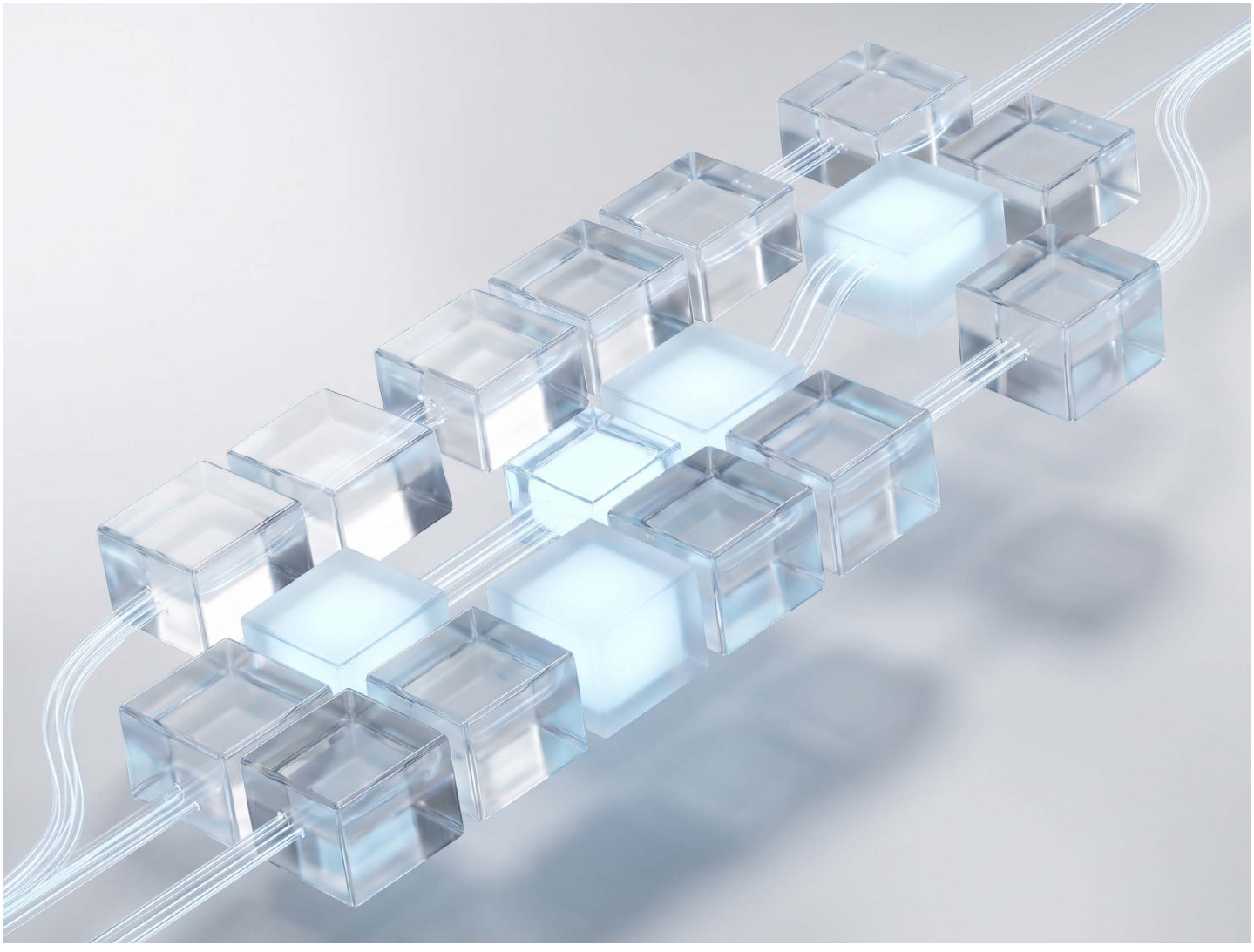
MIT Technology Review
Insights

Produced in partnership with

cyberhaven

Data lineage is the new security perimeter, providing the visibility needed for governance, accountability, and trust, without impeding AI adoption at scale.

Securing the agentic enterprise with data lineage



Key takeaways

1. Agentic AI exacerbates data visibility challenges because it accelerates data replication, fragmentation, and movement across human and machine workflows.
2. Trustworthy AI operations require more than model performance alone; enterprises must be able to explain how data was transformed, what influenced agent outputs, and who or what has interacted with information.
3. To rapidly scale AI adoption, security teams must monitor workflows, provenance, context, and data movement (in addition to the underlying systems).

A sales representative asks an AI agent for customer pricing, and the agent pulls data from multiple sources: restricted pricing systems, internal communications, and archived quotes. The output looks accurate to the agent, so a second agent uses the data to draft the proposal for the sales rep. No human checks either handoff. When it's discovered that the proposal includes unapproved, confidential, and sensitive information, no one can trace where the problematic data entered the workflow or how it spread across systems, files, and agents.

This scenario illustrates why AI governance needs to evolve, rapidly and dynamically. As more enterprises begin to use agentic AI in daily workflows, they're realizing that they must manage more than static files; they must manage moving and fragmented data that is continuously ingested, copied, summarized, transformed, reused, and redistributed in new forms. In short, AI security isn't a tool problem; it's a data visibility problem.

Consequently, traditional security approaches built solely around applications, systems, networks, or endpoints are no longer sufficient. Securing agentic AI at scale requires understanding how data moves. Teams, therefore, need to establish data lineage in order to effectively manage **governance, accountability, and trust** across human, agentic, and hybrid workflows.

Losing sight of the data

An organization's data no longer lives in one place or as one complete object. It is copied, summarized, reused, splintered, and incorporated into countless workflows, both human and machine. A single data point can spawn hundreds or thousands of derivatives moving independently throughout an organization. And that movement is often obscured by collaboration platforms, file-sharing tools, cloud applications, distributed endpoints, and mobile workforces.

Such data visibility challenges aren't new, but the rapid adoption of agents has intensified them. "As soon as replication and fragmentation happen, you start losing visibility," says Nishant Doshi, CEO at Cyberhaven.

For example, visibility into how an output was produced often disappears once a task is completed. Ultimately, the erosion of visibility is not only a security problem; it's a trust issue. "If you can't trace what the output came from and where it's going and how it influences data, that's where you lose that trust," says Aman Sirohi, CISO at Cyberhaven.

Establishing data lineage, or the ability to trace data and its fragments across human and machine workflows and throughout its lifecycle, requires teams to understand two key things: data's provenance, or point of origin; and how that data moves and spreads across various workflows.

> "Data lineage is the memory of your data. Without the memory, AI has intelligence but no accountability or trust."
>
> **Aman Sirohi**, Chief Information Security Officer, Cyberhaven

Visibility depends on provenance

While inaccurate or hallucinated outputs have dominated the conversation around AI risks, provenance is an equally important factor. Clear provenance, such as always knowing what data access agents have, adds more trust to agents.

"Strong provenance builds trust into trust, but it is a continuous process," says Justin Lam, senior research analyst at 451 Research.

Organizations must understand which data influenced an outcome, whether appropriate permissions governed access, how information was transformed, and how an AI-generated result was produced. The National Institute of Standards and Technology (NIST) recommends, **as part of an AI risk-management framework**, organizations document training data sources to be able to trace the provenance of AI-generated content.

Sirohi emphasizes that AI outputs shouldn't be trusted unless their lineage can be verified. "Provenance is going to become the foundation of trustworthy AI," he says.

Trust but verify is no longer enough; lineage must be verified before AI-generated outputs can be fully trusted. As enterprise AI deployments scale, executives, auditors, regulators, and customers will increasingly require organizations to explain not merely what an AI system produced, but precisely how it arrived at its result.

Lineage provides context

With greater use of agentic AI comes more fragmentation of data.

Prompts, summaries, snippets, tickets, generated content, recommendations, and workflow outputs together create an ever-expanding collection of derivative information. Because AI is non-deterministic, these fragments often move independently from their original source material. "Agentic AI could unlock many of the risks of that data," says Lam. "There could be unforeseen dangers because we're not considering toxic combinations."

The risks Lam is referring to are security, privacy, and compliance due to the increasing number of paths through which information can leave an organization. For example, AI agents may place sensitive customer data in external documentation, unsanctioned applications and connectors might access restricted files, or an audit trail may be lost because agents failed to authenticate or log activity. Without data lineage, an organization cannot confidently explain to regulators, auditors, customers, or incident-response teams how information was handled. "What data lineage adds to the equation is context," Doshi says.

> "Good security removes the friction from innovation."
>
> **Aman Sirohi**, Chief Information Security Officer, Cyberhaven

Context helps organizations detect inappropriate sharing, and determine whether sensitive information is being used in unsanctioned ways. This approach enhances the traditional file-centric security model, which focuses primarily on documents in storage or attachments in transit, while missing the smaller units of data moving through modern, agentic AI workflows.

Lineage reduces data risk

A lack of data lineage can turn ordinary AI-assisted work into a business, security, or compliance issue. Consider the following examples:

- **A sales assistant combines an outdated playbook with preliminary internal guidance and produces pricing that is no longer approved.**
- **A customer-service chatbot provides a plausible answer that conflicts with contractual restrictions governing the customer's data or the approved model's use.**
- **An employee reuses an AI-generated summary, unaware that it includes sensitive customer data from an unauthorized source.**

In each of these scenarios, the challenge extends beyond inaccurate output. Without data lineage, the organization cannot track the information sources, permissions, and transformations behind the AI output.



“What data lineage adds to the equation is context.”
Nishant Doshi, Chief Executive Officer, Cyberhaven

Context is also essential for understanding and managing exceptions. While a blanket policy might dictate that annual recurring revenue figures should never leave the organization, a CFO may have a legitimate business need to share that information with external stakeholders. Establishing context based on normal workflow patterns alongside data lineage enables a security team to safely allow justified exceptions without defaulting to broad restriction or broad exposure that compromises both productivity and data security.

Turn visibility into risk-based control

Visibility serves as a stepping stone to control. By **using context** to distinguish between low-, medium-, and high-risk workflows, security teams can allow low-risk activity to proceed freely while applying additional controls where warranted. They can provide guidance, just-in-time coaching, or approval steps for medium-risk workflows, while redirecting high-risk workflows to sanctioned environments with suitable guardrails. The goal is to calibrate controls based on data sensitivity, user identity and permissions, workflow purpose, transformations involved, and **possible downstream consequences**.

Organizations must translate visibility into governance decisions without disrupting productivity. If organizations want visibility and control to scale, then security must automate in parallel, Lam says. “As these tasks get more complicated and autonomous, that’s where you’ll see the need to make security more autonomous as part of the AI stack itself.”

Scaling agentic workflows requires oversight

As agentic AI graduates from experimentation to enterprise deployment, organizations increasingly operate in three workflow modes: human-to-human, human-to-AI, and fully agentic. Visibility must extend across all of these workflows **and encompass** human actions, agent actions, data access, transformations, and downstream delivery. The new unit of AI security becomes the workflow itself, says Doshi.

63%

of breached organizations lacked AI governance policies.

Source: Compiled by MIT Technology Review Insights, based on data from IBM, 2026.

Teams once managed a handful of daily interactions involving sensitive data, but now AI systems **can generate thousands** of decisions and information exchanges over the same timeframe. “Visibility gets lost in profound technological changes,” says Lam.

As more decisions become autonomous, organizations will rely increasingly on exception-based intervention rather than manual review of every action. While human oversight remains essential, it must adapt so security can scale in tandem with AI adoption. “As agentic processes become more sophisticated, the gating value will be questioned: ‘Do I really need to check in with humans?,’” Lam says.

Shadow AI adds governance complexity

Achieving complete data visibility requires shining a light on unsanctioned, or “shadow,” AI.

Shadow AI is one of the fast-growing issues for enterprises, extending far beyond unauthorized models. According to research from IBM, **63% of breached organizations** lacked AI governance policies, with 20% experiencing breaches linked to shadow AI.

“Shadow AI is growing at a scale that we haven’t yet wrapped our hands around,” says Sirohi. Shadow AI can impact personal accounts, applications, unofficial connectors, plug-ins, skills, and workflows. But most shadow AI activity is not malicious; rather, it stems from employees adopting unauthorized tools to work more effectively. And organizations are increasingly relying on open-source models, which trade visibility for affordability.

“Someone might switch from Claude Code to Cursor,” Lam explains. “No security team truly knows everything that could happen there.”

Context is essential for understanding how shadow AI helps employees accomplish legitimate work. By gaining visibility into the data flows, prompts, uploads, connectors, and identities involved, organizations can assess the

risk of shadow AI in the context of business objectives and establish a practical path toward safe adoption. As Sirohi says, "Good security removes the friction from innovation."

Build a roadmap for data lineage visibility

To establish a lineage-first security strategy for scaling agentic AI, organizations should prioritize five core areas:

- **Inventory AI usage:** Track both sanctioned and shadow AI, including personal accounts, connectors, skills, and unsanctioned workflow components.
- **Establish end-to-end lineage:** Extend visibility across endpoints, SaaS platforms, on-premises systems, prompts, outputs, AI agents, and downstream data flows.
- **Define ownership and accountability:** Classify and authorize data use. Ensure AI operators manage execution and transformation, while business owners remain accountable for decisions based on outputs.
- **Apply least-privilege access controls:** Treat AI agent access as a foundational identity-and-access-management (IAM) priority, granting only the permissions needed for each agent's specific role.
- **Implement continuous monitoring and response capabilities:** Identify risky data movement, assess exceptions, simulate possible outcomes, and continuously refine policies using observed workflow data.

"Take stock of what you have," Lam says. "Understand what activities are in place, and map them back to business goals."

Organizations should move toward **shared accountability models** in which data creators, AI operators, business leaders, and security teams all play distinct roles.

"The companies that will win with AI will not be the ones that use the most AI," Sirohi says. "They'll be the ones that built trust in their security architecture to move faster with confidence and reduced risk."

Transitions from systems of record to systems of delivery

Enterprise security typically starts with systems of record, such as SaaS applications, data warehouses, and document repositories. Agentic AI adds another layer: systems of delivery that dynamically assemble information from multiple sources to serve customers, employees, or business processes.

AI-enabled workflows depend on a series of source systems, permissions, prompts, transformations, and outputs rather than a single database or application. Understanding that assembly process is central to establishing trust in AI outputs.

Securing the agentic enterprise requires a new mindset

The rise of agentic AI requires rethinking enterprise security beyond traditional application categories. Traditional perimeters, endpoint-centric controls, and application-focused governance models are increasingly insufficient on their own. Now that information moves through interconnected human and machine workflows, conventional boundaries have blurred.

With agentic AI, the new perimeter is data, Sirohi says. When determining whether to approve a particular AI tool, security leaders must consider how well it can trace data's origin, movement, influence, access, and fragmentation. A lineage-first security model turns visibility into context and control, enabling accountability, governance, and secure AI deployment at scale.

"Data lineage is the memory of your data," Sirohi says. "Without the memory, AI has intelligence, but no accountability or trust."

"As tasks get more complicated and autonomous, that's where you'll see the need to make security more autonomous as part of the AI stack itself."

Justin Lam, Senior Research Analyst, 451 Research

"Securing the agentic enterprise with data lineage" is an executive briefing paper by MIT Technology Review Insights. Virginia Wilson was the editor of this report, and Nicola Crepaldi was the publisher. MIT Technology Review Insights has independently collected and reported on all findings contained in this paper.

We would like to thank the sponsor, Cyberhaven, as well as the following experts for their time and insights:

Nishant Doshi, chief executive officer, Cyberhaven
Justin Lam, senior research analyst, 451 Research
Aman Sirohi, chief information security officer, Cyberhaven

About MIT Technology Review Insights

MIT Technology Review Insights is the custom publishing division of MIT Technology Review, the world's longest-running technology magazine, backed by the world's foremost technology institution – producing live events and research on the leading technology and business challenges of the day. Insights conducts qualitative and quantitative research and analysis in the US and abroad, and publishes a wide variety of content, including articles, reports, infographics, videos, and podcasts. This content was researched, designed, and written by human writers, editors, analysts, and illustrators. This includes the writing of surveys and collection of data for surveys. AI tools that may have been used were limited to secondary production processes that passed thorough human review.

From the sponsor

Cyberhaven is the leader in Data Security for the Agentic Enterprise, tracing the full lifecycle of your data and adapting protection as context changes. Protect workflows, not just data. Take action when and where it matters. Proven by the world's most innovative companies. AI changed work. See how Cyberhaven protects it at **cyberhaven.com**.

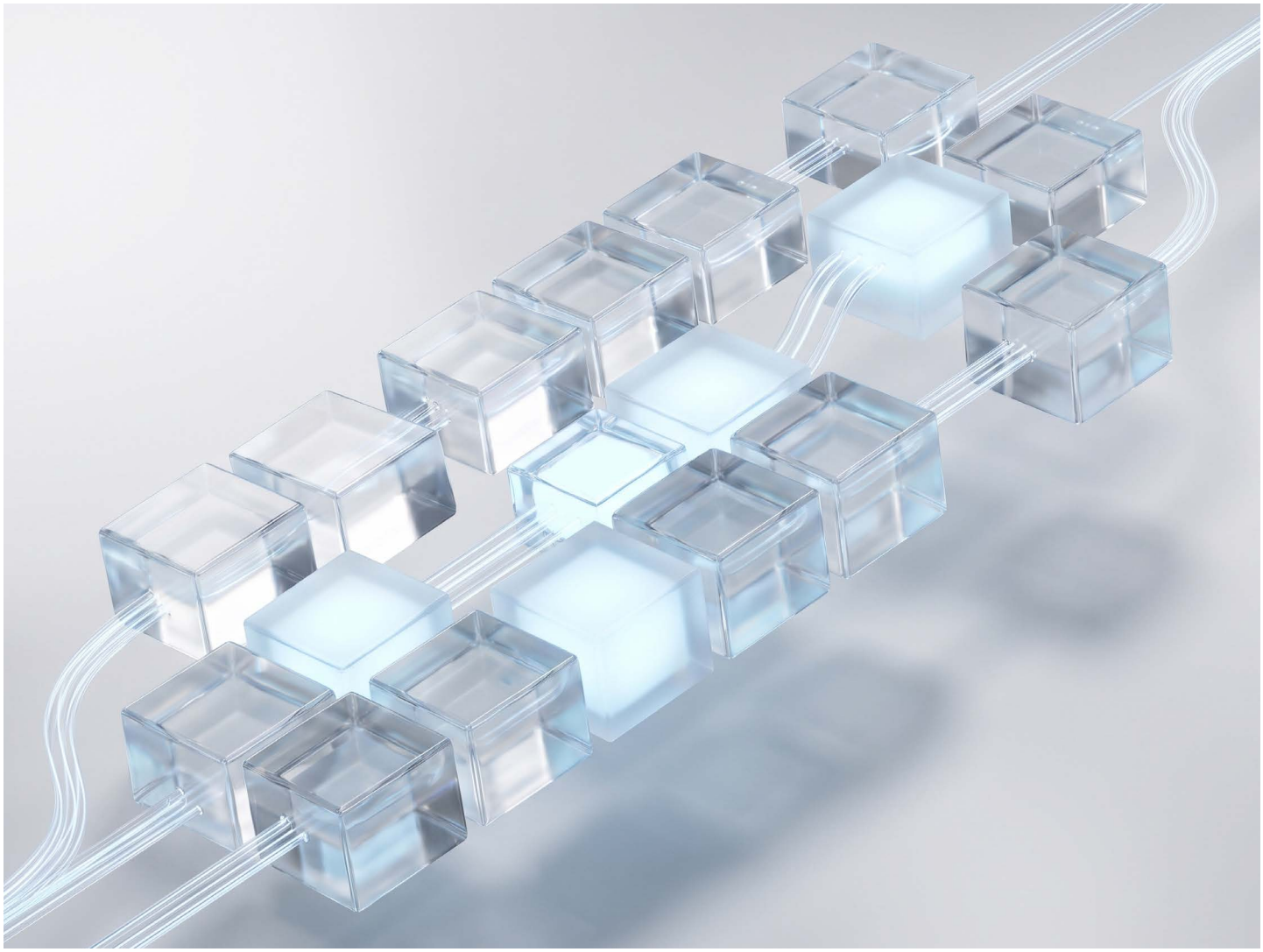


Illustrations
Cover art and spot illustrations by Adobe Stock.

While every effort has been taken to verify the accuracy of this information, MIT Technology Review Insights cannot accept any responsibility or liability for reliance by any person on this report or any of the information, opinions, or conclusions set out in this report.

To cite this report, please use: "Securing the agentic enterprise with data lineage," MIT Technology Review Insights and Cyberhaven, November 2026.

© Copyright MIT Technology Review Insights, 2026. All rights reserved.



MIT Technology Review Insights

www.technologyreview.com

insights@technologyreview.com