CYERA

2025 REPORT

State of AI Data Security

How to close the readiness gap as AI outpaces enterprise safeguards

Research by

Cybersecurity INSIDERS

Executive Summary

Artificial Intelligence (AI) has crossed the tipping point. 83% of enterprises already use AI in daily operations, yet only 13% report strong visibility into how it is being used. The result is a widening gap: sensitive data is leaking into AI systems beyond enterprise control, autonomous agents are acting beyond scope, and regulators are moving faster than enterprises can adapt. AI is now both a driver of productivity and one of the fastest-expanding risk surfaces CISOs must defend.

This report, based on a comprehensive survey of 921 IT and cybersecurity professionals, sets out to answer a critical question: as AI becomes embedded in the enterprise, are CISOs equipped to govern it with the same rigor applied to users, systems, and data? The findings reveal a clear tension: AI adoption has gone mainstream, but visibility, monitoring, and access frameworks remain shallow and fragmented. Left unchecked, AI functions as a shadow identity—powerful, fast, and often unaccountable.

Key Survey Findings

- **AI adoption without oversight:**
  83% already use AI, yet only 13% have strong visibility, leaving most enterprises blind to how AI interacts with their data.
- **Agents are the new shadow risk:**
  76% say autonomous AI agents are the hardest to secure, with 70% pointing to external prompts.
- **AI as a shadow identity:**
  Only 16% treat AI as a distinct identity, while two-thirds have caught AI over-accessing data.
- **Controls lag reality:**
  Nearly a quarter have no prompt or output controls, and only 11% can automatically block risky AI activity.
- **Governance gaps persist:**
  Only 7% have a dedicated AI governance team, and just 11% feel fully prepared for regulation.

The chapters that follow examine these findings across three themes: how AI deployment is outpacing control, why agents and prompts create new exposures, and why identity and access management must be redefined for AI.

To guide the response, each chapter connects survey data with the OWASP Top 10 for LLM Applications—the leading community framework for AI security risks such as prompt injection, excessive agency, and unbounded consumption. By aligning enterprise experience with OWASP's categories, this report offers CISOs not only evidence of the gaps but also a practical roadmap for closing them.

AI Deployment Outpaces Control

AI has become nearly universal in the enterprise, but adoption has surged faster than the guardrails needed to manage it. Organizations are embedding AI into daily workflows, yet most lack visibility, monitoring, or the ability to enforce controls, leaving a governance gap that regulators are beginning to notice.

- 83% of organizations already use AI, yet the majority (57%) remain in early maturity stages.
- Only 13% report good or full visibility into AI usage; nearly half admit they have little to no visibility.
- Logs are treated more as forensics than defense: a third of organizations review them only after incidents.
- Only 11% can automatically block risky AI activity; one-third acknowledge awareness without controls.
- Governance lags behind adoption, with only 7% having a dedicated AI governance team and 11% fully prepared to meet regulatory requirements.

Together these findings confirm a simple truth: the enterprise risk surface created by AI is expanding far faster than the governance and enforcement structures meant to contain it.

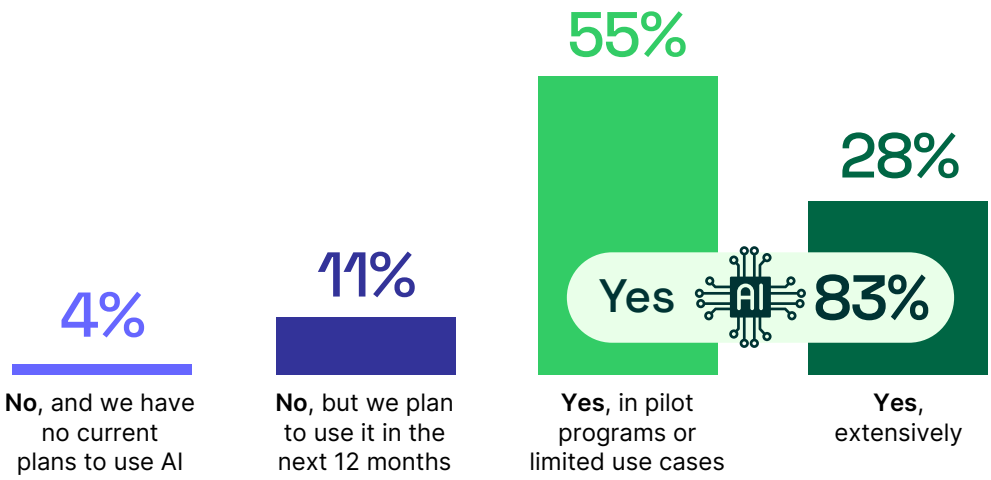
AI Deployments Create Systemic Risk Exposure

While 83% of enterprises already use AI in some capacity, most remain in shallow stages: more than half are limited to pilots (55%) and only 28% report extensive adoption. Maturity lags even further, with the largest share describing themselves as only “emerging” (39%).

The AI model footprint is highly concentrated. Nearly four in five rely on ChatGPT or OpenAI (79%), while Microsoft Copilot (57%) and Google Gemini (41%) follow close behind. The most common uses are content and knowledge generation (75%) and productivity and collaboration (71%), which may appear routine, yet they already touch the very operational data that defines how enterprises run.

Even modest pilots can create systemic exposure: a small team testing an AI tool for drafting reports containing sensitive data may trigger rapid uptake across departments, all funneled through the same external ecosystem and leaving the enterprise dependent on a single vendor’s model without safeguards in place.

Does your organization currently use Artificial Intelligence (AI) in any capacity?



Additional responses include: Not sure 2%

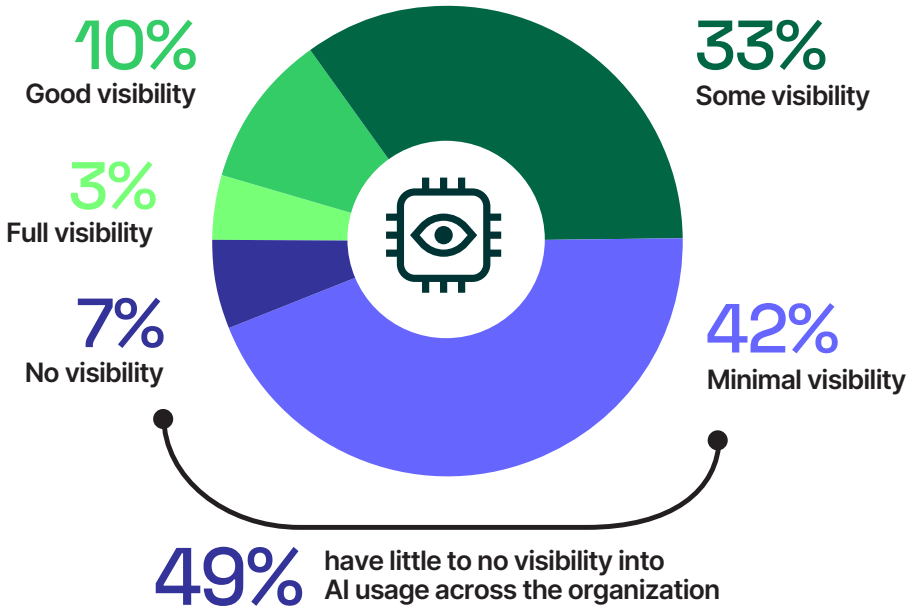
OWASP’s 2025 guidance reinforces this reality: early adoption without oversight maps directly to LLM02 Sensitive Information Disclosure and LLM10 Unbounded Consumption. Continuous discovery of AI tools, classification of the data they touch, and real-time logging of prompts and outputs must begin at the pilot stage. Without these controls, enterprises risk compounding governance debt as adoption spreads.

Blind Spots Define AI Risk

Over 8 out of 10 enterprises use AI, yet only 13% report robust visibility into how it is being used. Nearly half admit to have no or low visibility. The result is that most CISOs cannot reliably answer where and how AI is operating inside their own organizations. Oversight remains reactive: a third of enterprises review AI activity logs only after incidents, while just 9% monitor in real time. Only 14% detect anomalous or rogue AI behavior as it happens, and more than one in five do not monitor at all.

It's not hard to see the risk: a sales team enables an AI copilot to draft proposals, and weeks later, fragments of sensitive pricing data surface in outputs, discovered only after the fact because no one was watching.

How much visibility do you have into AI usage across your organization, including external AI tools (e.g., ChatGPT, Copilot), embedded AI features in SaaS (e.g., Salesforce Einstein, Notion AI), and homegrown AI applications?



Additional responses include: Not sure 5%

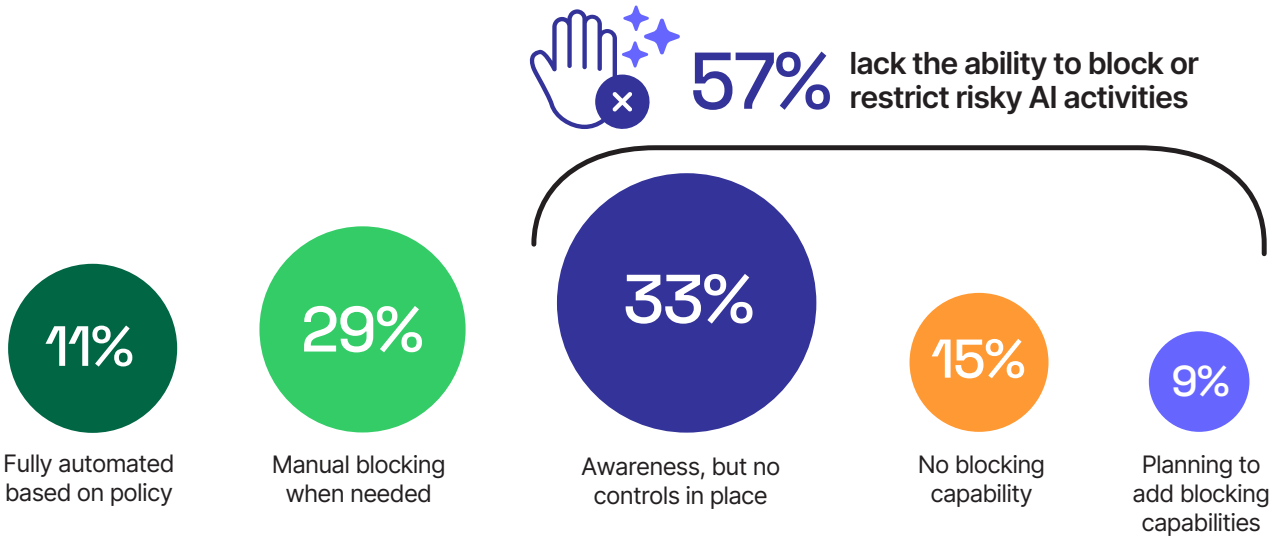
OWASP guidance aligns directly with this risk: insufficient monitoring fuels LLM02 Sensitive Information Disclosure, while post-incident log review leaves organizations vulnerable to LLM08 Vector and Embedding Weaknesses and LLM10 Unbounded Consumption. Continuous discovery, real-time logging, and anomaly detection are essential if visibility is to move from hindsight to defense.

Controls Lag Behind Reality

Even when organizations recognize AI risk, enforcement is weak. Only 11% have automated blocking in place, while 29% rely on manual intervention. A third (33%) admit they have awareness without controls, 9% are planning to add blocking capabilities, and 15% say they cannot block misuse at all. In other words, more than half of enterprises are powerless to stop risky AI activity in real time.

This imbalance leaves teams reactive. Pilots and copilots are embedded into core workflows, yet control systems trail behind. Awareness without enforcement creates an illusion of safety: CISOs may know misuse is occurring but lack the ability to intervene. A recruiting copilot, for example, can be manipulated through a crafted prompt to exfiltrate candidate data; and without automated blocking, the only option is cleanup after the fact.

What's your ability to block or restrict risky AI interactions today?



Additional responses include: Not applicable 3%

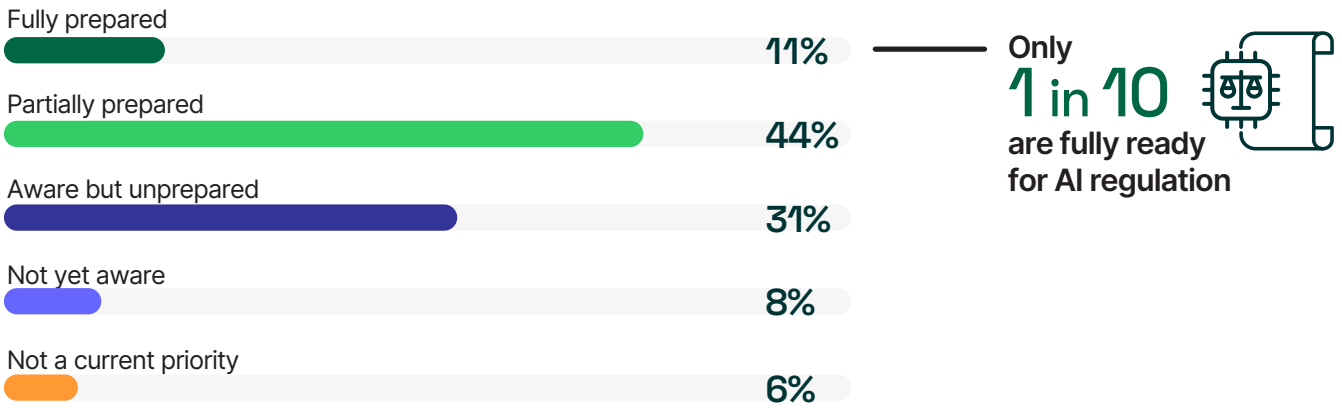
OWASP classifies these exposures as LLM01 Prompt Injection, LLM02 Sensitive Information Disclosure, and LLM06 Excessive Agency—a reminder that what CISOs see as operational problems are already codified as systemic risks. Automated blocking, kill switches, and rate limits must be designed-in from the start, otherwise enterprises remain one crafted prompt away from exposure.

Governance Without Readiness

AI adoption is racing ahead, but the organizational structures to govern it are still catching up. Only 11% of organizations consider themselves fully prepared for regulatory requirements tied to AI data governance. Nearly half (44%) admit they are only partially prepared, while another 31% are aware of obligations but are unprepared. The remainder are either not yet aware (8%) or do not see regulation as a current priority (6%). This leaves the majority of enterprises exposed to compliance failures at the very moment regulators are beginning to enforce AI-specific mandates more robustly.

Ownership of AI governance is equally fragmented. Just 7% report having a dedicated AI governance committee. For most, responsibility is distributed: 34% say ownership is shared, 17% place it with the CIO, and only 12% see it with the CISO. With accountability scattered across IT, security, and risk leaders, governance gaps continue to widen. And while policies may be drafted, robust enforcement often falls between silos.

How prepared is your organization to meet regulatory requirements related to AI data governance (e.g., EU AI Act, NIST AI RMF)?



OWASP guidance makes the stakes clear: lack of clear ownership leaves enterprises exposed to LLM02 Sensitive Information Disclosure and LLM06 Excessive Agency, while shallow readiness invites compliance failure under LLM10 Unbounded Consumption. Governance must anchor in accountable ownership, auditable controls, and data-aware policies applied consistently across the AI surface.

Agents and Prompts: The Exposed Edge

While organizations are more comfortable with embedded AI inside trusted SaaS platforms, their confidence collapses once autonomy or external prompts come into play. Autonomous agents and public LLM prompts now define the riskiest parts of the enterprise AI surface, and most organizations admit they lack the guardrails to keep them in check.

- 76% say autonomous AI agents are the hardest to secure, and 70% name external prompts to public LLMs as equally high risk.
- Even embedded SaaS AI is not trusted by all: 43% still see it as difficult to secure.
- 40% acknowledge shadow AI is already present, operating outside sanctioned oversight.
- 21% grant AI broad access to sensitive data by default, and 66% have already caught AI over-accessing information it didn't need.
- Nearly a quarter (23%) admit to having no prompt or output controls in place, while filtering, monitoring, and redaction remain inconsistently deployed.

These findings highlight a paradox: organizations feel safe with AI embedded in familiar tools, yet at the very points where AI operates autonomously or crosses external boundaries, oversight collapses, creating exactly the kind of risks attackers are most likely to exploit.

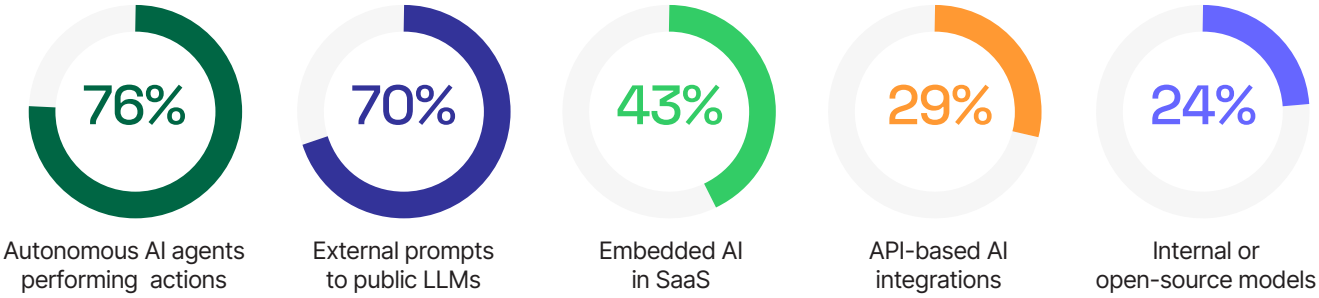
Agents and Prompts Are the Pain Point

Not all AI interactions carry equal risk. Three-quarters (76%) of organizations say autonomous agents are the hardest to secure, followed closely by external prompts to public LLMs at 70%. By contrast, embedded SaaS AI is flagged as difficult by only 43%. Inside the SaaS boundary, AI feels manageable; once autonomy or public prompts are in play, confidence collapses.

The shadow dimension is already here. Four in ten organizations report the presence of unsanctioned or "shadow AI" operating outside approval and oversight (40%). These tools mirror the risks of rogue agents, expanding the attack surface in ways security teams cannot see or control.

The failure mode is easy to imagine. A Slack integration built on an experimental AI agent begins pulling data from private channels. Intended to streamline collaboration, it instead creates a silent data leakage path that no one notices until sensitive conversations appear in unintended outputs.

Which AI interaction types are hardest to secure in your environment?



OWASP maps these exposures directly: LLM01 Prompt Injection, LLM02 Sensitive Information Disclosure, and LLM06 Excessive Agency all converge when unsupervised agents and external prompts are in play. The remedy is narrow agent scopes, explicit approvals, and kill switches—controls that must be in place before autonomy or public LLM access is allowed inside the enterprise.

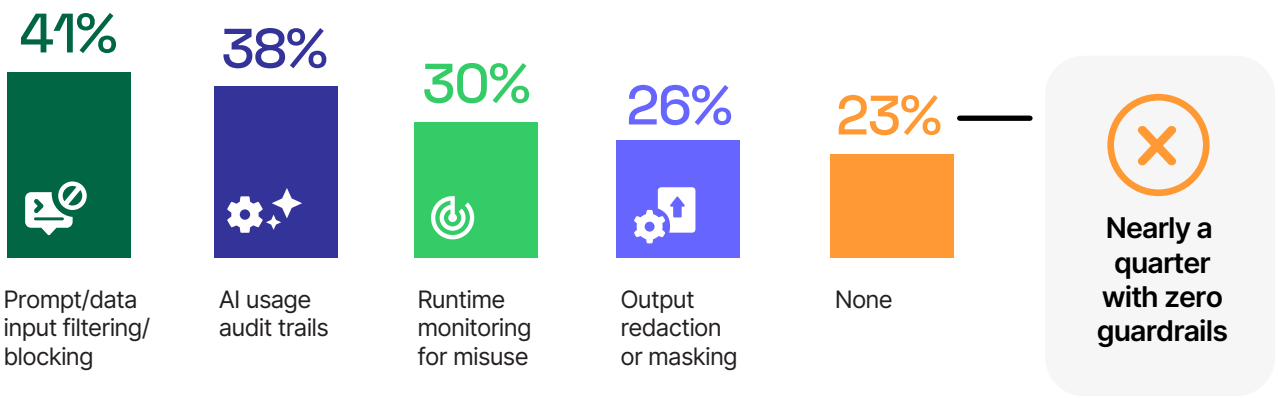
Guardrails Missing at the Prompt Layer

Prompts and outputs are the choke points where sensitive data most often slips into or out of AI systems, yet most organizations admit they lack meaningful safeguards. Nearly a quarter (23%) have no prompt or output controls in place at all. Filtering or blocking of risky inputs is applied by fewer than half (41%), while only 26% redact outputs. Audit trails are present in 38%, and runtime monitoring covers just 30%.

This leaves the majority exposed. Inputs can be manipulated, outputs can leak sensitive information, and misuse can pass unnoticed through everyday workflows. The paradox is that even security teams themselves rely heavily on AI: 77% already use it to enhance SOC operations, often without the same rigor applied to other critical interfaces.

The risks are straightforward. A customer support bot with minimal filters is fed a crafted prompt, which returns confidential billing records to the requester. The model did not "break"—the failure was the absence of guardrails at the input-output layer.

Which controls are in place to govern AI prompts and outputs?



Additional responses include: Not sure 8%

OWASP identifies this gap as central to LLM01 Prompt Injection and LLM02 Sensitive Information Disclosure, compounded by LLM05 Improper Output Handling. Effective governance requires filtering, monitoring, redaction, and auditability to operate together as the default policy. Without this stack in place, prompts and outputs remain open channels for exploitation.

The Identity Gap in AI Governance

Identity and access management remains the cornerstone of enterprise security, but most organizations are still applying human-centric models to AI. The result is over-permissioned systems and disconnected controls, and a new identity class that is growing without governance.

- Only 16% treat AI as a distinct identity class, while 77% either blur AI with humans or apply inconsistent rules.
- 21% grant AI broad access to sensitive data by default, and 66% have already caught AI accessing more information than necessary.
- Only 9% say data security and identity controls are fully integrated for AI, while 36% admit classification exists but is not linked to AI enforcement.
- Nearly a quarter (23%) of organizations have no formal governance for AI data access, relying on legacy role-based access or manual approvals.

Taken together, these findings show that identity governance built for people cannot simply be stretched to cover AI. Unless AI is defined and managed as its own identity class, enterprises will keep discovering that their most sensitive data has been exposed by systems treated as “just another user.”

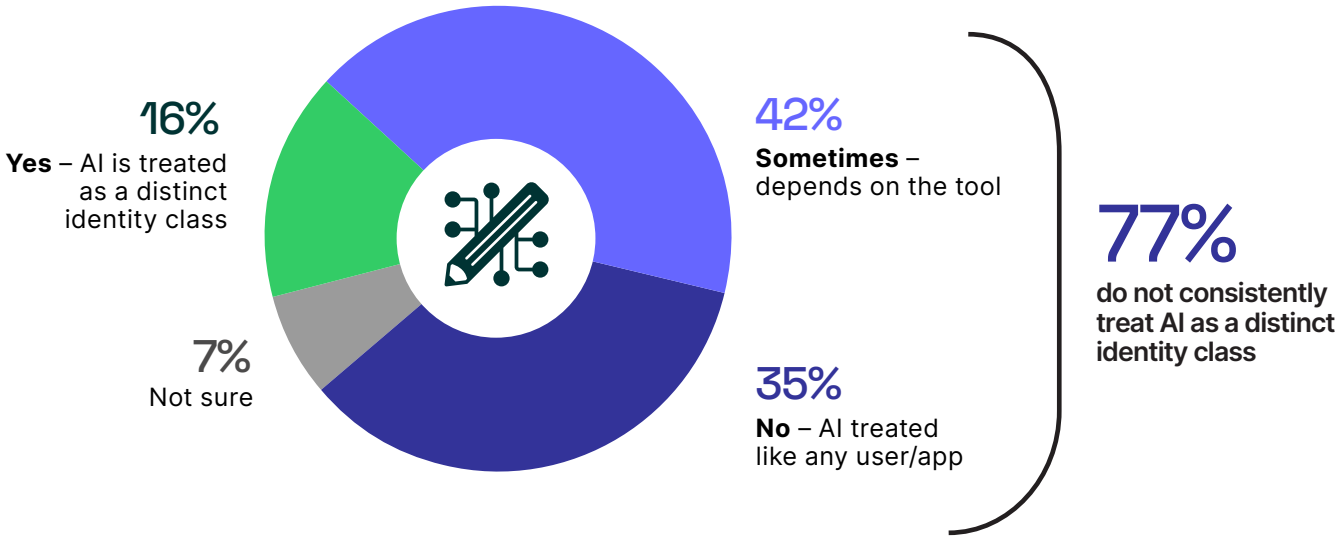
AI Still Lacks Its Own Identity

Identity is the foundation of modern security, yet most organizations still govern AI through human-centric models that were never designed for machine-scale activity.

Only 16% of respondents treat AI as a distinct identity class in their access control and monitoring systems. The majority blur the lines, either treating AI like any user (35%) or applying inconsistent rules (42%). It's a governance contradiction: enterprises know AI behaves differently, yet persist in managing it as though it were a human user.

Compounding this, nearly a quarter (23%) admit they have no formal governance for AI data access at all. In many enterprises, AI isn't just missing a unique identity class, it isn't governed systematically at all. This leaves a structural blind spot, where AI operates with permissions modeled after people, but without the accountability or constraints designed for machines. The problem is that AI identities behave fundamentally differently from people: they can generate requests at machine speed, act autonomously across domains, and operate without intent or accountability.

Do you differentiate AI tools from human users in your access control and monitoring systems?



OWASP guidance ties this gap to LLM06 Excessive Agency and LLM02 Sensitive Information Disclosure. Treating AI as a first-class identity, with its own policies, reviews, and least-privilege rules, is essential. Without this shift, organizations will continue to misapply legacy controls to a new identity they do not fully understand.

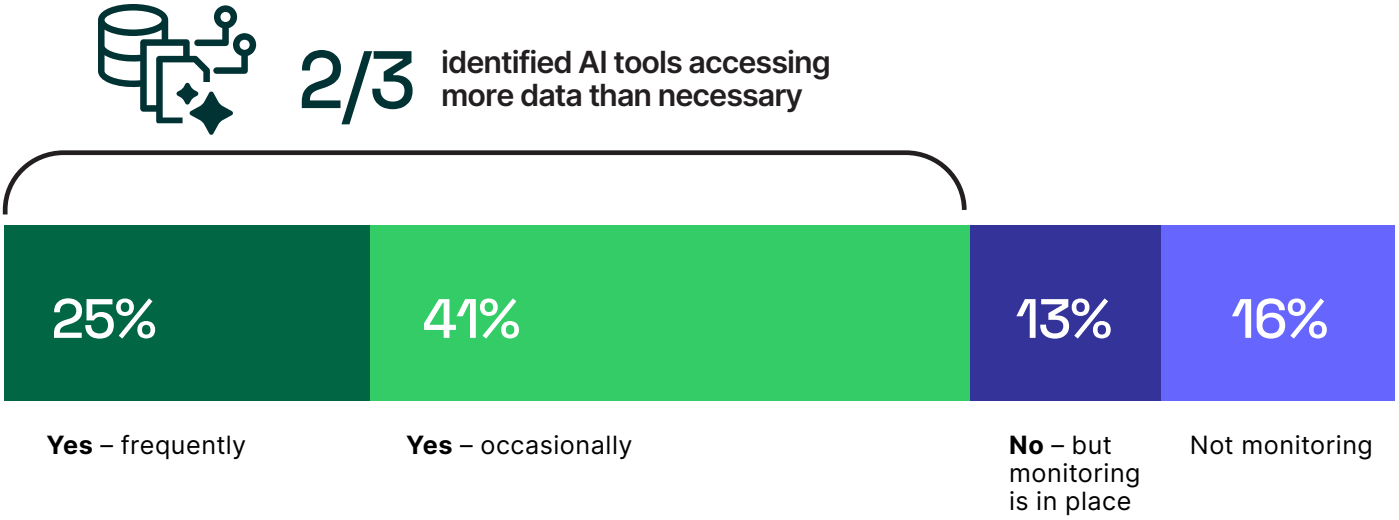
Over-Access Is Built Into AI

AI is amplifying one of security's oldest problems: excessive access by default. One in five organizations (21%) admit their AI systems are granted broad access to sensitive data from the start. While 42% scope access by project or team and 26% limit AI to anonymized or non-sensitive data, permissive defaults remain common.

The results are predictable: two-thirds of organizations have already caught AI tools accessing more data than necessary—25% frequently and 41% occasionally. Only 13% report no incidents where monitoring is in place, while 16% admit they are not monitoring at all.

The exposure is familiar: an enterprise enables a copilot to assist with document searches. Within days, users discover the system is surfacing confidential files well beyond intended scopes—not because of a breach, but because the defaults were too broad and no real-time controls were in place.

Have you identified any AI tools accessing more data than necessary?



Additional responses include: Not applicable/not sure 5%

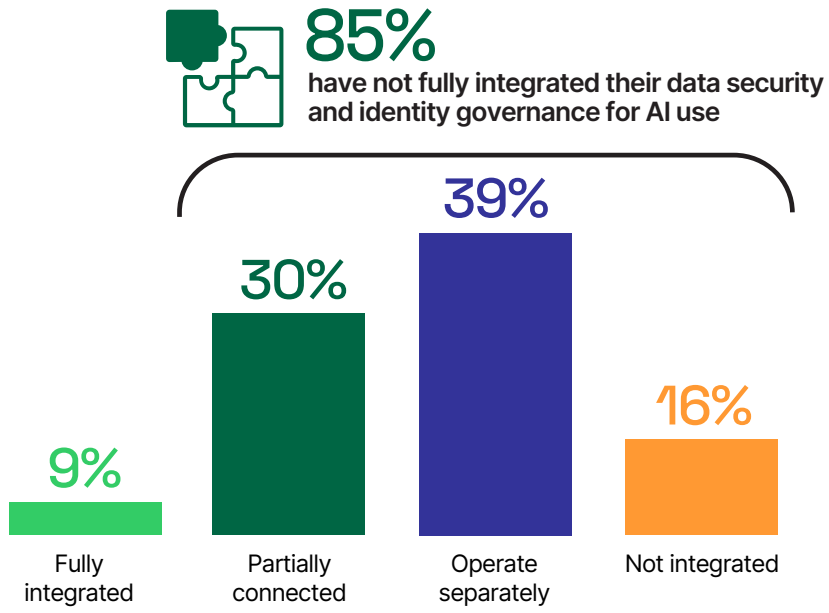
OWASP guidance is explicit. Excessive access links to LLM02 Sensitive Information Disclosure and LLM06 Excessive Agency. Data-aware least privilege, enforced continuously, is the only reliable guardrail. Without it, AI identities will keep reaching further than they should and enterprises will only notice after the fact.

Data and Identity Are Siloed

In the AI era, security depends on linking who (identity) with what (data) in real time—yet most organizations still govern these domains separately. Only 9% report that their data security and identity governance are fully integrated for AI use. Another 30% say they are partially connected, but 39% admit the two operate separately and 16% say they are not integrated at all.

Data classification shows the same weakness: just 15% apply it to real-time access control, while most use it only for static policies (31%) or keep it disconnected from AI controls entirely (36%). The rest lack classification altogether. This fragmented approach means that AI systems can request or retrieve sensitive data without identity checks informed by context—and identity systems can grant permissions without any awareness of the data being touched.

How integrated are your data security and identity governance controls for AI use?



Additional responses include: Not sure 6%

OWASP guidance aligns precisely with this risk: fragmented controls leave enterprises open to LLM02 Sensitive Information Disclosure, LLM06 Excessive Agency, and LLM08 Vector and Embedding Weaknesses. Effective governance requires convergence—unified, real-time enforcement that maps identity to the sensitivity of data being requested. Without it, AI will keep stepping across silos unchallenged.

Best Practices for Securing Enterprise AI

AI has become mainstream, but most enterprises admit their security controls lag adoption. The survey data shows where the gaps are—OWASP's Top 10 for LLM Applications shows how to close them.

1 **Build Visibility from the First Pilot:**
Treat every pilot as production. Enterprises are already embedding AI across workflows (83%), but only 13% report strong visibility, with a third reviewing logs only after incidents. OWASP flags these blind spots as LLM02 Sensitive Information Disclosure, LLM08 Vector and Embedding Weaknesses, and LLM10 Unbounded Consumption. The fix is continuous discovery, real-time logging, and anomaly detection from day one.

2 **Contain Agents and Prompts:**
Narrow agent scopes, require approvals, and default to prompt/output filtering. Three-quarters of security leaders cite agents as hardest to secure, and 70% point to prompts, yet nearly a quarter have no controls in place. These exposures align to OWASP LLM01 Prompt Injection, LLM02 Sensitive Information Disclosure, LLM05 Improper Output Handling, and LLM06 Excessive Agency.

3 **Redefine Identity for AI:**
Treat AI as a first-class identity with least-privilege and classification-driven access. Only 16% of enterprises do so today, even as two-thirds have already caught AI over-accessing sensitive data. This maps directly to OWASP LLM02, LLM06, and LLM08. AI requires least-privilege by default, classification-driven access, and unified identity-data enforcement—controls that legacy IAM models cannot provide.

Conclusion

The evidence is clear: AI adoption has outpaced data governance, and AI risks are scaling faster than data defenses. The readiness gap is real and will continue to widen. CISOs who act now can use AI as an advantage to the business, and even shift into the mindset of a Chief Data Officer (CDO). Those who wait will inherit AI as an unmanaged liability rather than a tool for business success.

Methodology & Demographics

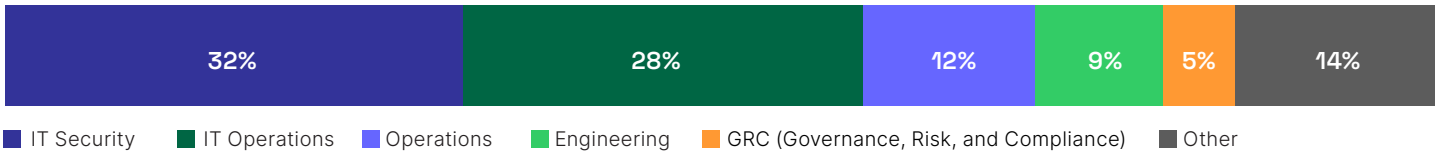
This 2025 State of AI Data Security Report is based on a survey of 921 IT and cybersecurity professionals. Respondents represented a balanced cross-section of industries, company sizes, and roles—including CISOs, IT security executives, architects, SOC leaders, and data governance professionals.

The survey explored enterprise readiness for AI adoption, focusing on visibility gaps, control maturity, identity and access governance, and priorities for aligning with OWASP's Top 10 for LLM Applications. Responses were self-reported and collected via structured multiple-choice questions.

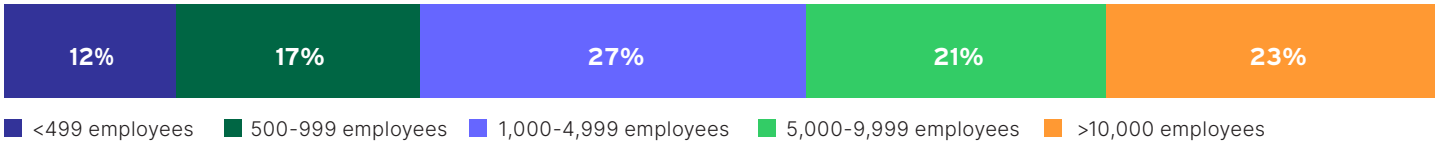
The survey has a margin of error of ±3.2% at a 95% confidence level, providing a statistically robust snapshot of how enterprises are adopting AI, where governance is falling short, and what controls CISOs view as most urgent.

JOB TITLE / LEVEL OF SENIORITY

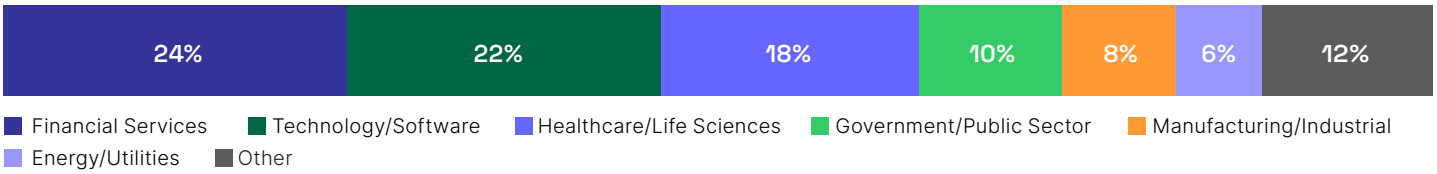
DEPARTMENT



COMPANY SIZE



INDUSTRY



CYERA Research Labs

Cyera Research Labs, the research organization within Cyera, aims to advance security through data-driven research. Our team of researchers, scientists, and cloud experts analyzes how data is created, accessed, and shared to uncover emerging threats and AI-driven risks. Every insight we publish blends rigorous analysis with practical guidance, empowering organizations to protect and govern their data with confidence.

Cyera is the world's leading AI-native data security platform. It gives organizations a complete view of where their data lives, how it's used, and how to keep it safe, so they can reduce risk and unlock the full value of their data, wherever it is. Backed by more than $1.3 billion in funding from top-tier investors including Accel, Coatue, Cyberstarts, Georgian, Lightspeed, and Sequoia, Cyera's unified data security platform helps businesses discover, secure, and leverage while eliminating blind spots, cutting alert noise, and protecting sensitive information across the cloud, SaaS, databases, AI ecosystems, and on-premises environments. Recent innovations like Cyera's Omni DLP extend this platform with adaptive, AI-native data loss protection, bringing real-time intelligence and contextual understanding to how data moves and is used across the enterprise.

www.cyera.io

Cybersecurity INSIDERS

STRATEGIC INSIGHT FOR CYBERSECURITY LEADERS

Cybersecurity Insiders delivers evidence-backed insights that empower security leaders to make informed, strategic decisions. Backed by over a decade of research and a global network of 600,000+ cybersecurity professionals, we provide actionable intelligence to help leaders navigate emerging threats, evaluate new technologies, and shape forward-looking strategies with confidence.

For cybersecurity vendors, we turn research into results — delivering credibility, visibility, and demand through high-impact formats such as:

- Data-powered market reports that establish thought leadership,
- Webinars that build trust with buyers through credible, expert-led narratives,
- CISO guides that showcase best practices,
- Product reviews that independently validate solutions,
- Thought leadership articles that educate buyers, and
- Award programs that elevate brand reputation.

By combining this content with built-in distribution, we help brands earn trust, amplify awareness, and drive demand in a crowded cybersecurity market.

For more information, visit

cybersecurity-insiders.com

Sept 2025 // Version 1.5