Machine Data Outlook

Operationalizing AI in the Enterprise

splunk>
a CISCO company

Contents

AI has moved from experimental curiosity to the center of the global economy, driving unprecedented capital investment and reshaping labor markets. This surge is sparking concern by governments, consumers, and legal experts seeking to establish regulatory frameworks, increase transparency, and implement ethical safeguards. For technology leaders, the mandate has shifted from showcasing AI's potential to delivering tangible ROI for shareholders and customers.

Yet, the AI conversation often overlooks the massive machine data explosion accompanying it. IDC projects that by 2028, global data volume will reach 394 zettabytes, with 55% of the growth attributed to machine data.[1] By 2029, IDC also predicts the number of AI agents will exceed 1 billion worldwide, contributing to machine data growth.[2]

Machine data is the lifeblood of AI. It's the operational context AI needs to act reliably — yet managing it is becoming a substantial hurdle. Companies must curb runaway infrastructure costs while harnessing this massive data stream to power operational efficiency and innovation.

We surveyed 4,380 cybersecurity, AI, engineering, and other technology decision-makers to understand how AI is pushing machine data infrastructure to its limit. Seventy-four percent say AI and automation have caused *moderate* or *significant* data management challenges.

The research also suggests organizations acknowledge that machine data is important but are still determining the best way to realize its value for AI.

"Many organizations know what good looks like. The challenge is making it real with AI and data you can trust, at the speed and scale modern enterprises require," says Kamal Hathi, Splunk SVP and GM.

This report explores the importance of machine data to inform AI insights, recommendations, and actions. It's a reciprocal relationship that when done right, is a self-sustaining cycle that continuously improves AI's next output. As organizations seek to operationalize AI, it is critical to ensure the cycle is cost-effective and trustworthy.

74%

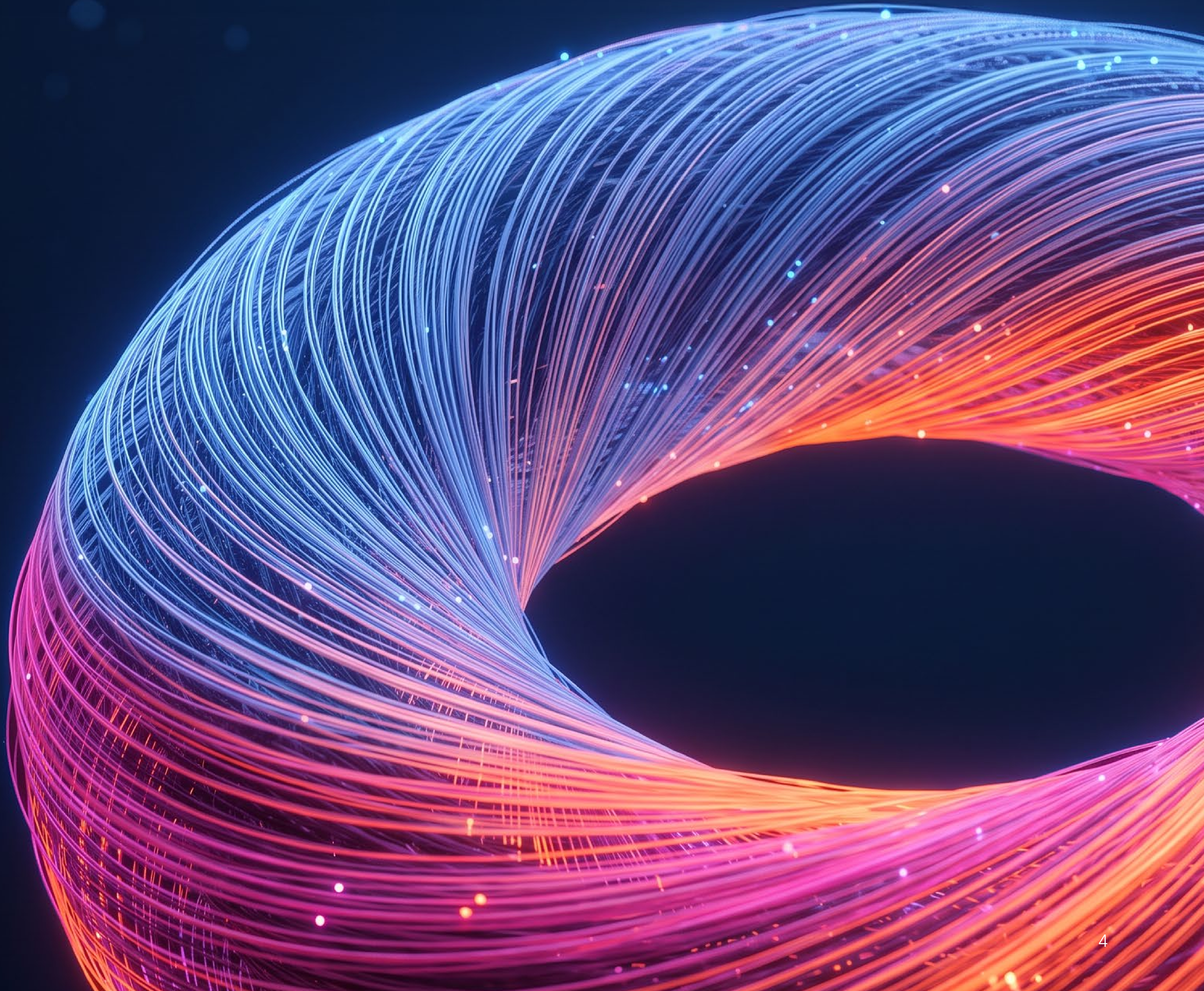
AI and automation is creating *moderate* or *significant* data management challenges

1 Adam Wright, "Worldwide Global DataSphere Structured and Unstructured Data Forecast, 2025–2029," August 2025, IDC.

2 Adam Wright, "Agent Adoption: The IT Industry's Next Great Inflection Point," IDC, December 10, 2025

The continuous AI + machine data cycle

AI and machine data create a self-sustaining loop where each fuels the other's growth. But raw telemetry alone is not enough to drive reliable AI outcomes. Organizations need business and operational context to make data AI-ready.



AI provides a tailwind for machine data growth

Every AI interaction — whether a model inference, automated workflow, or agentic decision — can produce machine data. This machine data comprises the logs, metrics, traces, events that reveal what is happening across applications, infrastructure, networks, endpoints, and AI systems in real time.

Splunk Field CTO Keith McClellan describes machine data as the operational memory of modern digital systems. "Machine data is the digital emissions of systems working together to produce an outcome," he says. "It's where teams can harvest insights to create the self-improvement loop AI needs."

A single interaction may produce a small amount of data, but as enterprises extend AI deeper into operations, data will multiply exponentially. IDC estimates that agents will execute 217 billion actions per day by 2029.[3]

Our survey validates this surge in machine data, with AI providing the tailwind: 92% of respondents report an increase in machine data over the *last* 12 months, with 91% expecting continued growth over the *next* 12 months. The primary drivers are autonomous AI agents and AI-assisted software development (63% each).

Top five drivers of machine growth

63%	63%	54%	51%	47%
Autonomous AI agents	AI-assisted software development	Cloud and multicloud adoption	Internal AI model training	Digital infrastructure scaling

Respondents could select all that apply

3 Adam Wright, "Agent Adoption: The IT Industry's Next Great Inflection Point," IDC, December 10, 2025

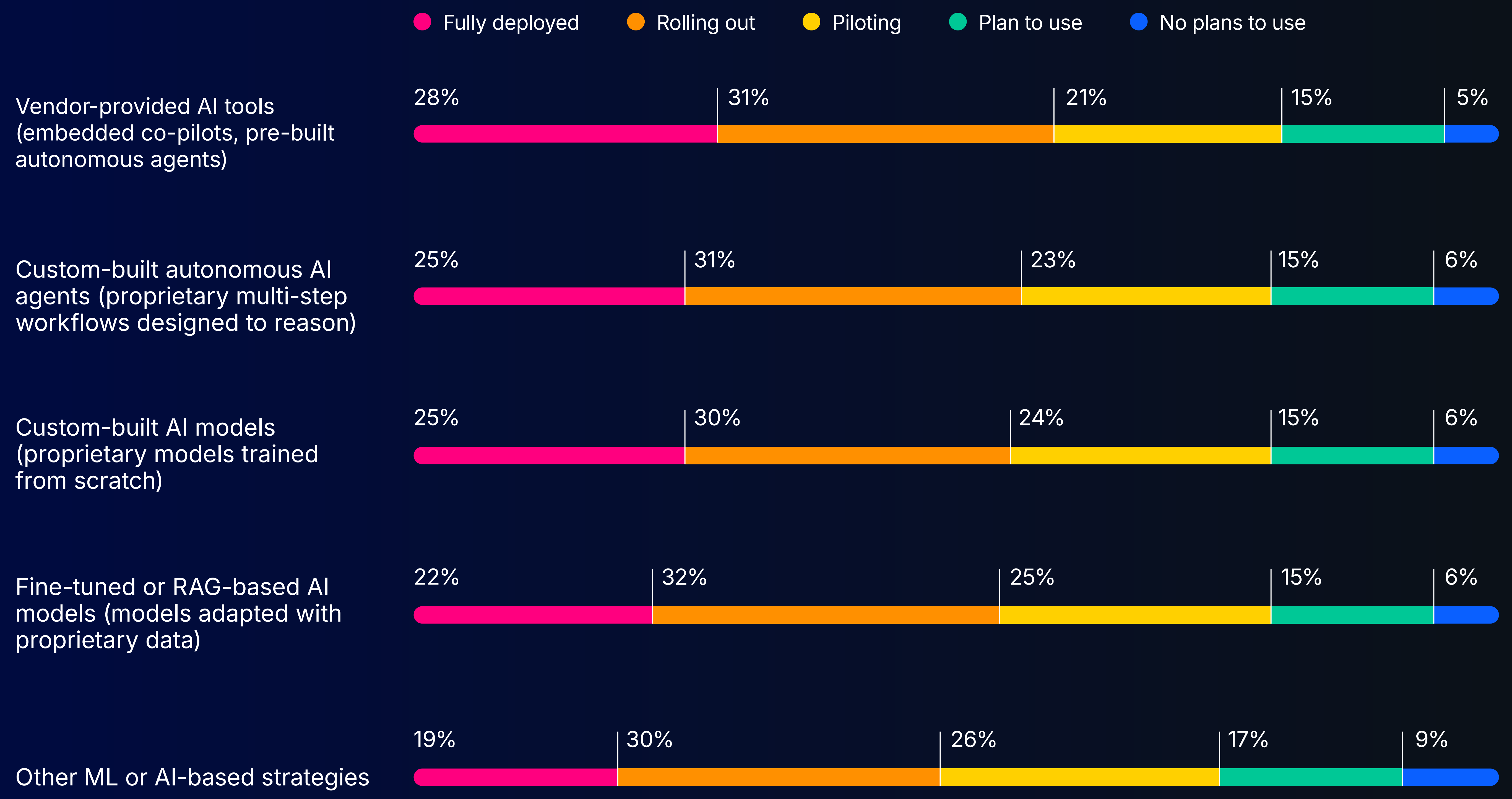
McClellan says the proliferation of machine data from AI also represents a shift in where that data is created. “AI is generating data everywhere, from places teams aren’t used to collecting large data feeds, like employee laptops,” he says. “It’s distributed, expensive to move, and it’s really messy.”

This distributed growth will accelerate as organizations transition AI from pilot to production. Respondents confirm they are *rolling out* or have *fully deployed* a range of AI tools and systems: vendor-provided AI tools (58%), custom-built autonomous agents (56%), custom-built AI models (55%), and fine-tuned or retrieval-augmented generation (RAG) AI models (55%).

Each deployment yields its own operational footprint, adding to the volume and complexity of machine data that organizations must manage. That expanding footprint is also becoming central to how AI performs. The top AI use cases informed or enhanced by machine data are AI development, such as model training (59%), observability for AI (56%), and autonomous AI agent workflows (53%).

> AI is generating data everywhere. It’s distributed, expensive to move, and it’s really messy.
>
> Keith McClellan, Field CTO, Splunk

The transition from AI pilot to full deployment is underway



Percentages may not add up to 100 due to rounding

Machine data fuels AI value

Organizations are using machine data extensively to fuel AI workflows and increase AI's value. Forty-seven percent say machine data will be *extremely* important to adopting and scaling AI over the next two to three years.

Getting a return on AI investment is also top of mind for executives and shareholders. Survey respondents say machine data drives AI value by enabling real-time automated action (54%), improving model accuracy and reliability (53%), and connecting operations to business outcomes (49%); 43% report that machine data helps drive AI value by providing domain-specific context to AI.

Organizations tap all kinds of data to train, fine-tune, or provide context to AI models. Business/contextual data is the most widely used, cited by 65% of respondents, and human-generated/unstructured data follows at 62%. Machine data comes third at 56%, despite its central role in telling AI systems what is happening across applications, networks, and endpoints.

Taking a broader view of AI outcomes, business and operational data play a vital role. Most respondents (72%) call enriching machine data with business and operational context a *high* or *critical* priority for successful AI initiatives.

Looking ahead, 58% of respondents identify improving model accuracy and reliability as the greatest opportunity for machine data to strengthen AI outcomes in the next 12 months, followed closely by connecting operations to business outcomes (51%), and enabling real-time automated action (49%).

72%

enriching machine data with business and operational context is a *high* or *critical* priority for successful AI

Making machine data AI-ready

Raw machine data alone is not enough to fuel continuous AI improvement.

"Translating AI into business value requires machine data context that is connected across environments, discoverable, observable, and governed," says Mangesh Pimpalkhare, SVP and GM, Splunk Platform. "These are already difficult tasks, but even more challenging given the scale and velocity of machine data. Most organizations cannot afford to do this over the long term."

Hao Yang, head of AI for Splunk, believes AI is pushing traditional data management towards obsolescence and that AI must be part of a machine data strategy. "AI is the foundational architecture of machine intelligence," he states. "To match the velocity of modern digital infrastructure, organizations need a machine data and AI ecosystem that is self-optimizing."

An important element of machine data context is metadata — where the data came from, where it lives, how it relates to other data, and what policies apply. Metadata, lineage, and relationships form a connective tissue that allow teams to analyze data where it lives — crucial to helping organizations mitigate soaring data management expenses.

"If you can get metadata right in a fragmented, distributed environment," says McClellan, "you can avoid unnecessary, costly data movement and duplication."

Sixty-six percent of respondents call automated metadata and lineage tracking a *high* or *critical* priority for AI success; 79% say this capability would have a *moderate* or *major* impact on their ability to realize value from machine data.

Yet even strong metadata practices can't compensate for machine data that isn't there. Missing or unavailable machine data is a common hurdle that undermines or prevents business-critical activities.

A narrow majority of technology leaders (53%) reveal they *frequently* or *always* lack the critical machine data required to support AI or automated decision-making. This deficit extends to informing strategic or investment plans (54%) and completing compliance or audit reports (50%).

Survey findings also reveal friction in day-to-day operations: Roughly half of respondents *frequently* or *always* struggle to access the machine data necessary for resolving major incidents or disruptions (51%), troubleshooting customer-facing issues (49%), or conducting cross-functional investigations across SecOps, NetOps, and DevOps (49%).

53%

frequently or *always* lack machine data needed for AI

The AI + machine data cost equation

As AI scales, the cost of storage, compute, and AI token consumption can quickly spiral out of control. Failing to move beyond simple “collect-and-centralize” models will be a significant liability.

The rising cost of AI-ready data

As organizations expand AI, they're encountering a new economic reality: AI doesn't only consume data; it consumes infrastructure, compute, and — increasingly — tokens.

In particular, AI agents are prolific producers of machine data. Agentic workflows do not behave like simple queries. They reason, verify results, and generate new telemetry along the way. Every step burns resources. Every action produces more data. Every new use case tests the foundation underneath it.

Technology leaders are feeling the pinch of machine data growth. Fifty-four percent cite escalating storage and infrastructure expenditures as a data management challenge exacerbated by the surge of machine data.

Data management expenses are also limiting AI potential. Forty-nine percent rank combined storage, processing, and AI consumption costs among the top three barriers to realizing the full value of machine data for AI initiatives.

The cost equation extends to model usage. "Tokenomics" — the economics of using and paying for AI at scale — has become central to fiscal planning. As systems reason and retry, token usage balloons, making cost control a prerequisite for sustainable AI scaling.

"Tokenomics is one piece of AI cost," according to Splunk Field CTO Keith McClellan. "It's about making sure that, when you're using a model you don't manage, you're not spending money you don't have."

49%

AI costs are a top barrier to realizing the value of machine data for AI

> Many organizations are still trying to support AI growth with data strategies built for a different era.
>
> Mangesh Pimpalkhare, SVP and GM, Splunk Platform

Data management investment lags data growth

On its own, the cost of AI-ready data is a dilemma. But there is a larger problem looming: Machine data volume is multiplying faster than the investment required to manage it.

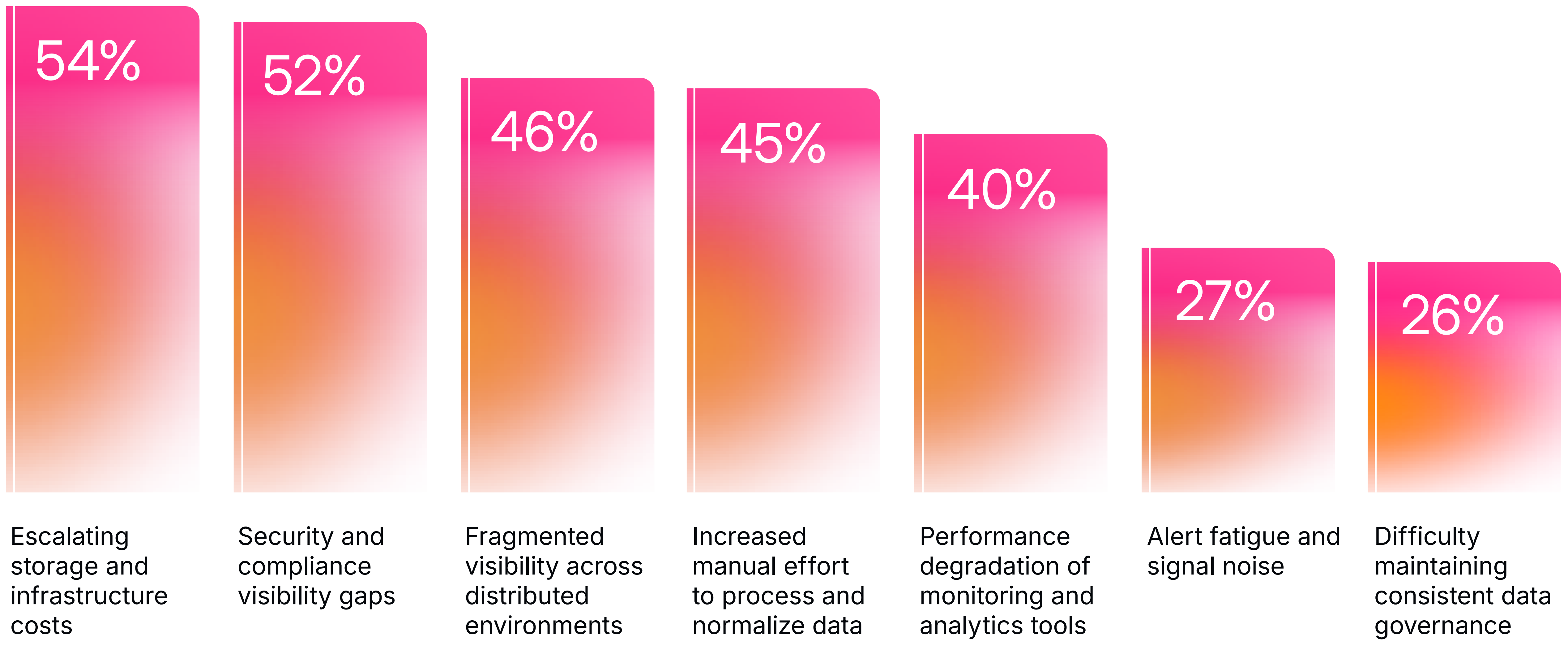
Thirty percent of technology leaders report a *significant* increase in machine data volume over the *past* 12 months, while only 22% report a corresponding rise in data management investment. The trend is expected to persist, with 37% of respondents anticipating *significant* volume growth over the *next* 12 months, compared to only 32% planning for similar escalation in data management investment.

This disparity is an obstacle for AI. As Mangesh Pimpalkhare, SVP and GM, Splunk Platform notes, "AI depends on operational context — the telemetry that explains what happened, where it happened, and what to do next. Many organizations are still trying to support AI growth with data strategies built for a different era."

That era often assumed data should be centralized. That assumption no longer holds.

The growth of machine data makes everything harder

Respondents reveal top data management challenges most exacerbated by machine data growth



Respondents could select all that apply

Centralized storage is cost-prohibitive

The uptick in AI-powered work is pushing machine data generation closer to where that work happens — applications, devices, networks, and operational technology employees use.

Moving every byte into a central repository is increasingly impractical, slow, and expensive. Seth Brickman, VP of global product management at Splunk, notes the solution is to bring intelligence to the source: "You can't always afford to move the data. You have to get intelligence as close to the point of generation as possible."

In fact, 54% of respondents rank escalating storage and infrastructure costs as a top data management challenge exacerbated by the explosion of machine data. But when organizations optimize how they manage and use machine data, they report cost savings — 41% saw reduced organizational cost and effort through improved data use practices.

VP of Products, Cisco Data Fabric, Hanlin Fang believes a stronger approach is to bring analytics to the data. "By analyzing machine data where it lives, teams reduce unnecessary movement and preserve the operational context AI requires to reason across environments." Fang continues, "The future lies in architectural unification where AI effectiveness hinges on our ability to map relationships across distributed environments, rather than forcing data into a single repository."

You can't always afford to move the data. You have to get intelligence as close to the point of generation as possible.

Seth Brickman, VP of Global Product Management, Splunk

Cost control measures erode context

Organizations are making hard — but practical — choices about what data to store, retain, and process.

Respondents admit they *frequently* or *always* adopt the following measures to get a tighter grip on machine data management costs: 68%[4] shorten retention windows to lower storage expenses; 68% also prioritize compliance over operational utility, keeping only what's legally required; and 61%[5] discard or sample data to avoid ingestion/processing fees.

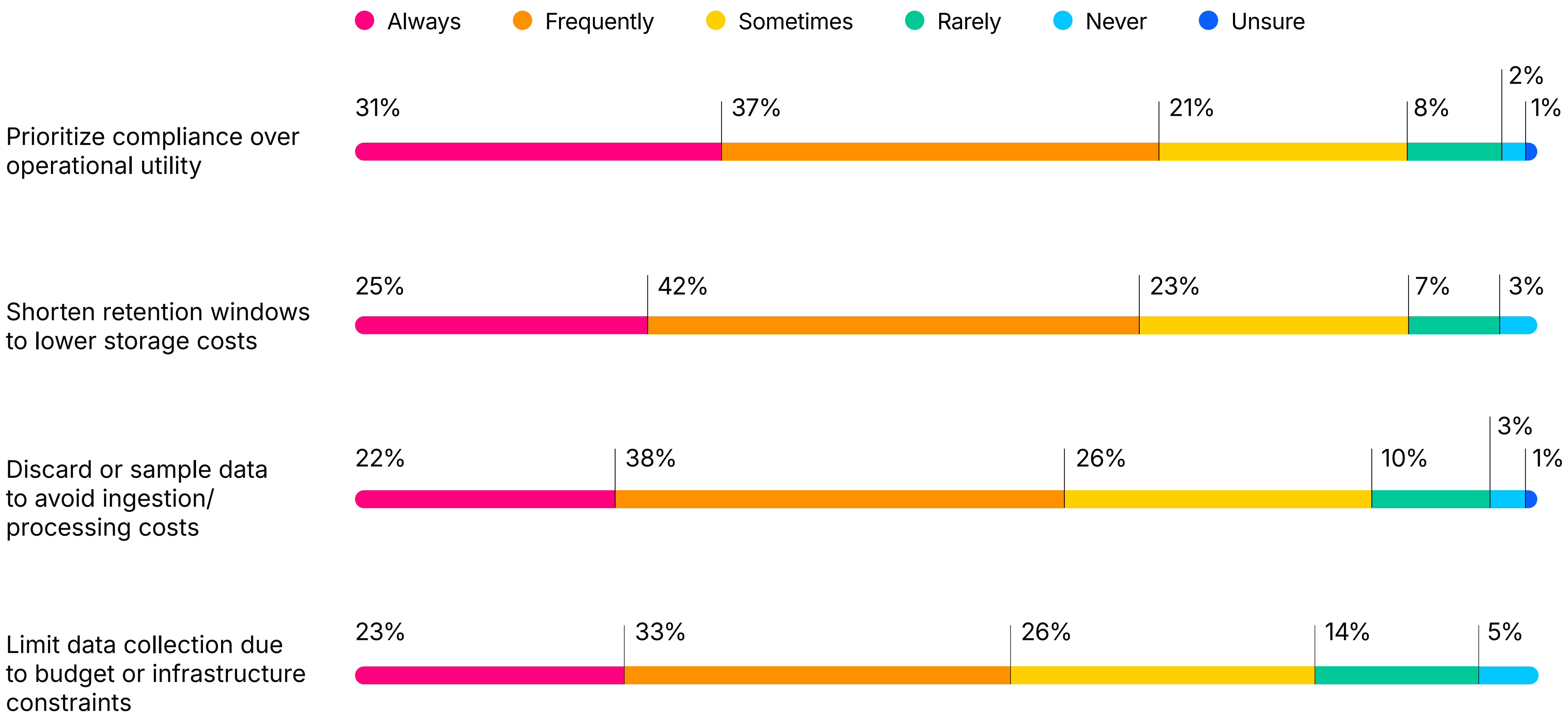
Applied thoughtfully, these practices can help teams manage the high price tag of machine data at scale. The risk comes when practices become blanket defaults instead of use case-specific.

"Sampling data is good for some problems and bad for others," says Brickman. "It's good if you're trying to determine normal behavior, but it's bad for outlier detection because the outliers have the potential to naturally go away when you sample."

This is the danger of context erosion: A rare event may be the signal that explains an outage, and a filtered dataset may remove the evidence AI needs to validate a recommendation or learn from a previous action.

Tanya Faddoul, VP of product and technology, strategy planning at Splunk notes the goal isn't to store everything indefinitely. "We must shift from reactive pruning to a model that preserves the context essential for high-fidelity decision-making, as the real risk is the irreversible loss of intelligence required to navigate complex environments."

Organizations make trade-offs to curtail machine data costs



Percentages may not add up to 100 due to rounding

4 Percentages are rounded to the nearest whole number: (25.3% + 42.2% = 67.5%)
5 Percentages are rounded to the nearest whole number: (22.4% + 38.2% = 60.6%)

The consequences of fragmentation

Data silos in distributed environments create blind spots that prevent teams and AI models from seeing the entire digital environment. Correcting this fragmentation helps AI agents make decisions with complete, accurate context.

Modern organizations are trending toward a sophisticated blend of distributed and centralized data management, according to their needs for scale, compliance, performance, and cost control.

When allowed to select all that apply, respondents confirm they employ multiple locations to store and manage their machine data. Nearly two-thirds (63%) opt for hybrid cloud and on-premises environments, while 40% rely on a centralized data lake or data warehouse, and another 40% use on-premises data centers. Twenty-three percent use edge or near-source storage.

Every storage environment has pros and cons. When it comes to discovering, querying, and accessing machine data, some environments get relatively high marks.

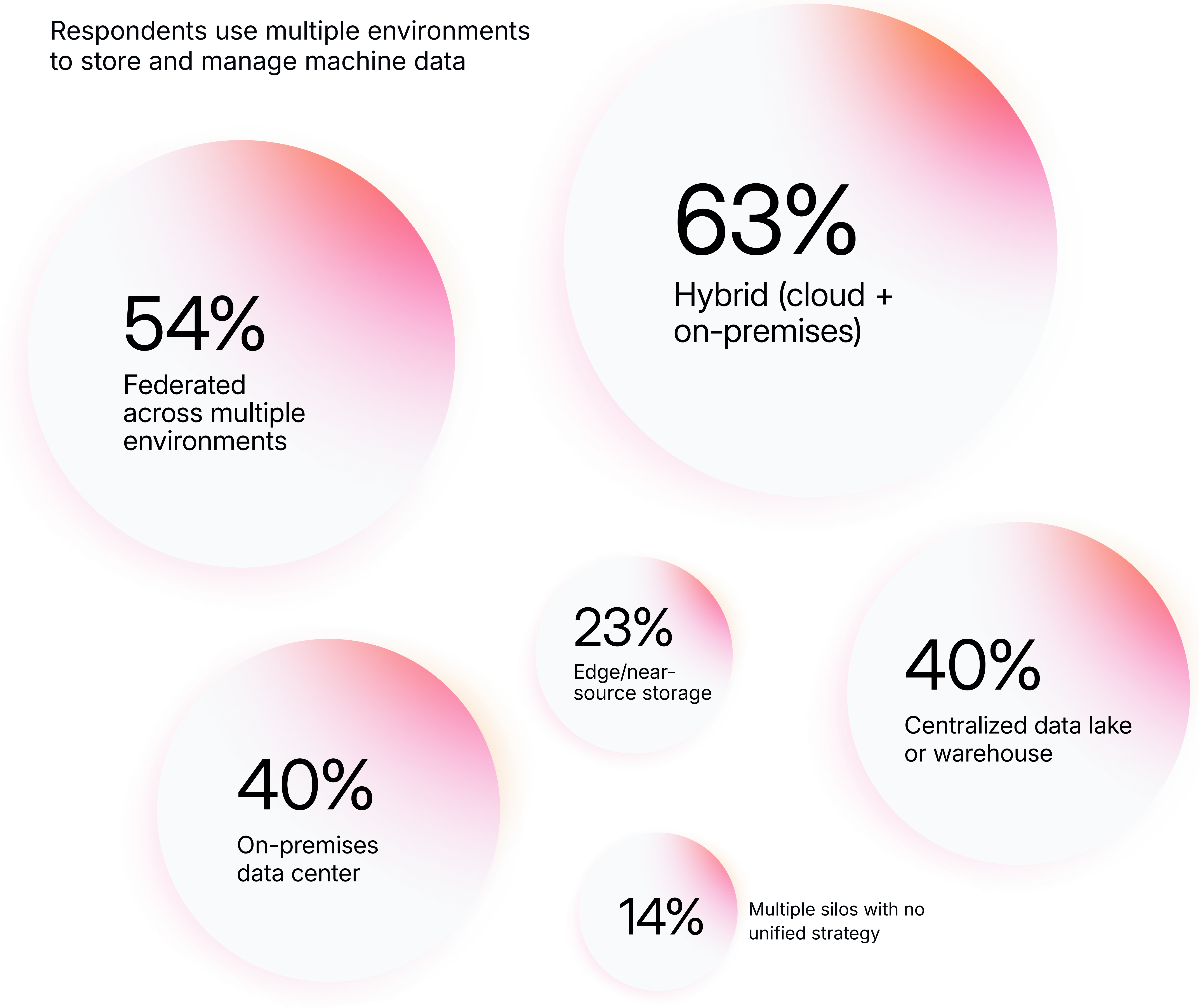
Fifty-four percent of respondents say their organization is *very* effective at managing machine data in the cloud; and 40% claim the same level of proficiency with on-premises data centers. The picture is slightly mixed for endpoint storage: while 40% report they are *very* effective at leveraging endpoint data, 15% admitted they are *not very* or *not at all* effective.

The edge didn't fare as well. Just 35% say their organization is *very* effective at utilizing machine data in edge environments, and 17% report they are *not very* or *not at all* effective.

This finding is noteworthy as AI expands beyond the cloud. Applications, devices, and operational technology will generate more machine data at the edge and endpoints, precisely where organizations are least effective at leveraging it.

Machine data is distributed

Respondents use multiple environments to store and manage machine data



Respondents could check all that apply

Fragmentation compromises context

When organizations use several locations to store and manage data, they feel the effects of fragmentation, defined in the survey as silos that prevent unified access and sharing across teams and environments.

Fourteen percent of respondents report having multiple, siloed systems with no unified strategy. And data fragmentation is ranked as one of the top three biggest barriers to fully realizing the value of machine data for AI (48%).

Fragmentation makes it difficult to preserve machine data context and connect signals across disparate systems, so AI agents and humans can interpret them. Seventy-one percent of respondents express *moderate* or *major* concern that fragmented data will lead to poor AI decisions.

"An agentic workflow may need to connect an application trace in the cloud, a security event from an endpoint, a network signal from a branch office, and telemetry from an operational system," says Cory Minton, Splunk global field CTO. "If those signals remain disconnected, AI may see pieces of the problem without understanding the whole."

Respondents also reveal the consequences of not having a unified view of their machine data: 60% said AI and automation act on incomplete context, leading to unreliable outputs. Meanwhile, 62% say they must make decisions without the full picture.

This finding aligns with the importance respondents place on unified machine data as an input for AI, with 76% declaring it a *high* or *critical* priority for the success of their organization's AI strategy.

As a potential solution, respondents acknowledge that federated search and analytics across distributed environments will be key: 74% indicate it would have a *moderate* or *major* impact on realizing value from machine data; while 64% call it a *high* or *critical* priority to a successful AI strategy.

"Agentic enterprises need a data foundation that preserves operational context across distributed environments without having to move all the data first," according to Hanlin Fang, VP of products, Cisco Data Fabric. "It's much more sustainable and scalable to make machine data available and searchable where it already lives."

But making data accessible across environments raises a further question: Who — or what — is allowed to access it, and under what conditions?

76%

unified machine data as an AI input is a *high* or *critical* priority for AI success

Governance boosts trust in AI

AI is unpredictable, which can render autonomous actions unsafe. Without strong guardrails, organizations can lose visibility into what AI is doing, or if it's following security and compliance protocols.

Trust — and transparency — are paramount as AI advances past providing passive recommendations to executing autonomous actions. Despite this leap, most respondents remain skeptical of the starting point: Two-thirds (66%) are *somewhat* or *very* concerned about poor or misleading AI recommendations.

"AI is inherently non-deterministic," notes Splunk Field CTO Keith McClellan. "If you tell AI it's allowed to change things, it may change things in ways you didn't expect and certainly won't do it the same way every time."

Without governance, it's easy to question AI output and actions.

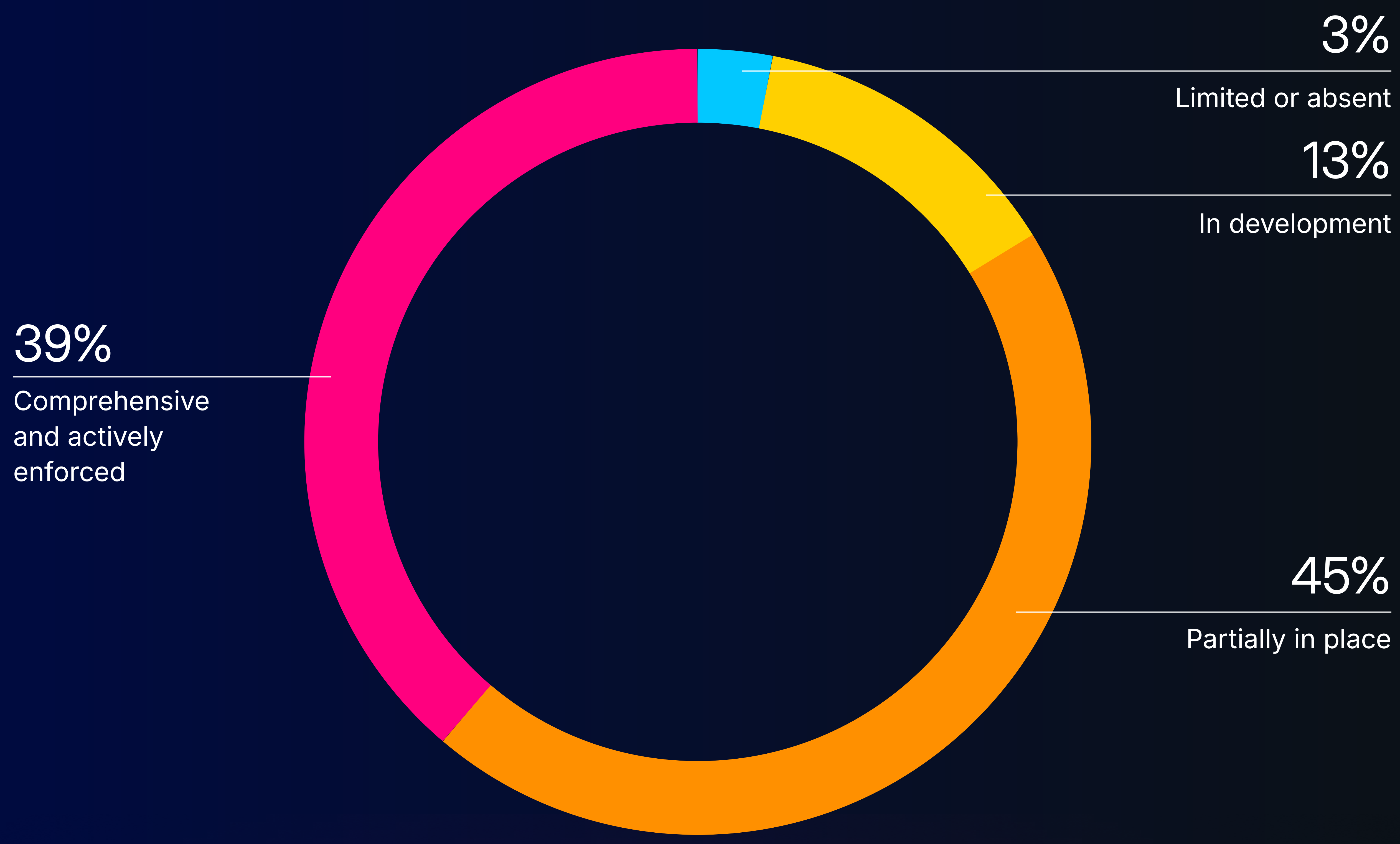
"A team may be able to see that an action occurred," states Splunk Global Field CTO Cory Minton. "But without proper controls, they cannot verify if that action was based on the right context or if it violated policies around data access, security, or compliance."

Relative to other factors, fewer respondents (26%) say strengthening AI governance and auditability is currently helping them drive value from AI. But they don't see it as a major hurdle, either: just 27% call governance and compliance constraints a top three barrier to realizing the value of machine data for AI initiatives.

There is also room to strengthen governance. While 39% describe their governance as *comprehensive* and *actively enforced*, another 45% say controls are only *partially in place*, with inconsistent coverage.

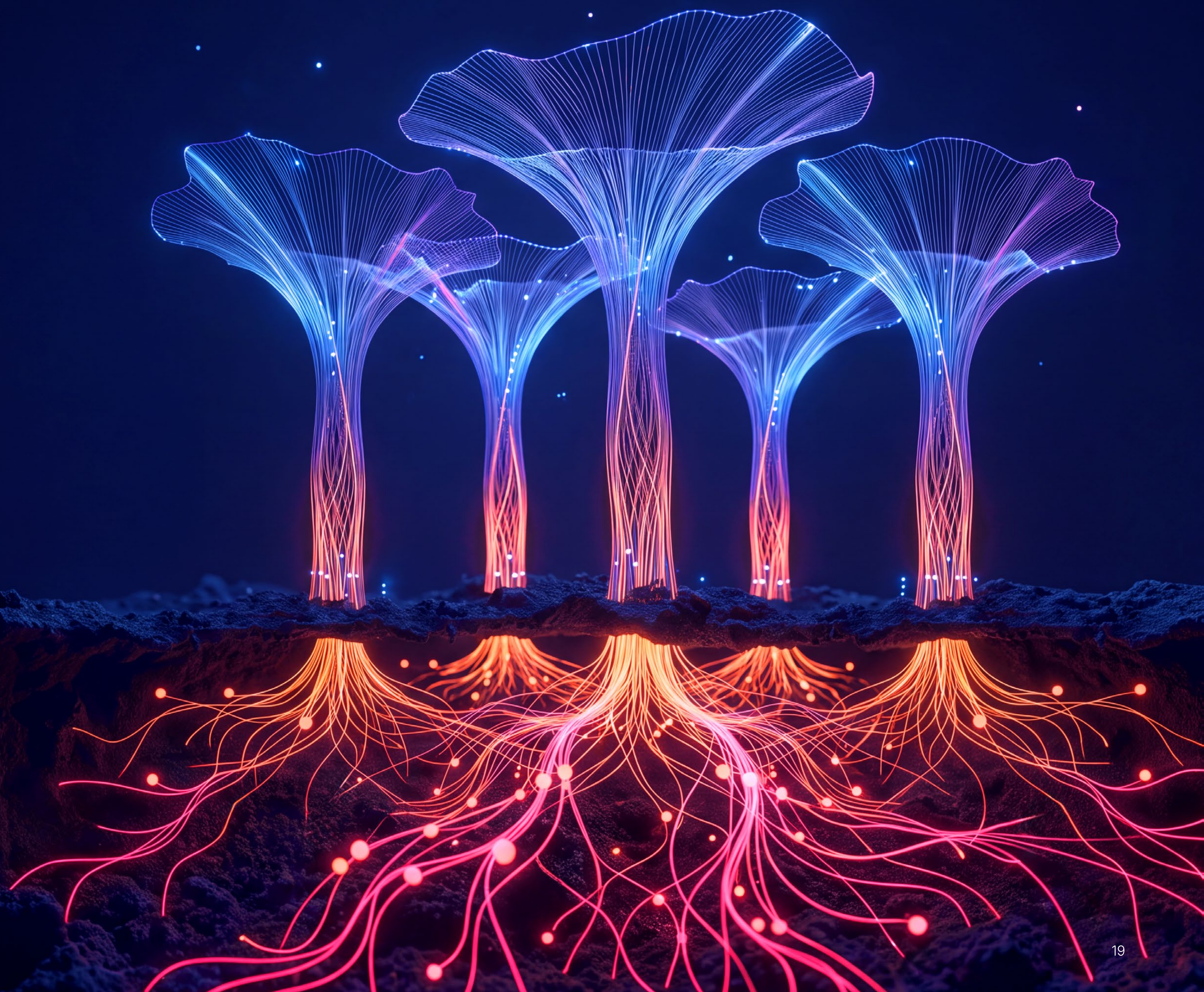
That 39% stands in stark contrast to how far AI deployment has progressed. The data reveals that 56% of organizations have already *deployed* or are *rolling out* custom-built autonomous agents, and 58% have done the same with vendor-provided AI tools. This means most have moved from experimentation to production, but fewer than 4 in 10 can claim their AI governance is comprehensive and enforced.

For most organizations, machine data governance is a work in progress



Five strategies to operationalize AI

AI is moving deeper into operational workflows, but not all organizations have an adequate machine data foundation. These strategies offer a path to operationalize AI so organizations can scale initiatives with confidence.



Machine data is the fuel that powers the AI engine. AI depends on machine data to act, and those actions generate more machine data, that in return, help improve AI's next decision. But managing this loop presents new obstacles.

For technology leaders, the goal is not simply to collect more data. It's to make distributed machine data discoverable, contextual, governed, and ready for AI. Here are five key strategies to help you move from data sprawl to realizing value from AI-ready data.

1 Prioritize context preservation over reactive pruning

Organizations need to trim costs; but pruning data without a plan is dicey. When retention, sampling, and filtering decisions become blanket defaults rather than deliberate choices, organizations may lose the exact signals teams and AI agents need.

- **Build a map before you pare things down:** Capture the data source and structure the moment it enters the environment. This turns raw telemetry into ordered, discoverable information — so when immediate access is no longer required, teams and AI agents can still find exactly what they need.
- **Decouple retention from accessibility:** Not every dataset needs to live in expensive, high-performance storage to remain useful. Give unprocessed data an affordable, long-term home, so teams aren't forced to choose between keeping context and controlling cost.
- **Match pruning strategy to use cases:** Sampling is a reasonable way to understand normal behavior, but outliers or anomalies can (and do) get missed. Instead of applying the same rule to all incoming data, use AI to filter data as it arrives based on use case requirements.

2 Elevate data to AI-ready status

Raw telemetry alone isn't enough to power reliable AI decisions. Without business and operational context, models reason with noise instead of signal, and outputs are difficult to trust.

- **Don't wait to enrich machine data:** Attach business context — user IDs, transaction types, service levels — to machine data as it's collected. This enrichment turns unprocessed logs into structured, context-rich datasets before they reach storage. And, grounds training and fine-tuning in operational reality.
- **Don't enrich data just to lose it again:** Added context is wasted if nobody — human or AI — can find it later. Use metadata to map your entire data landscape so it can automatically adapt to environment changes.
- **Match the model to the data:** Since many AI models are trained on human-generated text, not the raw telemetry your systems produce, they might be forced to rely on guesswork. Models trained on contextualized machine data close that gap.

3 Unify context through federation, not centralization

AI operates across edge, multicloud, and on-premises environments, with each deployment adding to fragmentation. Siloed data keeps AI from seeing the big picture, and the traditional fix — moving everything into a central repository — balloons storage costs and duplicates data. The preferred alternative to centralizing data is unifying its context.

- **Connect distributed sources:** A federated architecture maps relationships between datasets wherever they reside. This gives AI agents and humans visibility into your entire data landscape, no matter where the data is physically stored.
- **Search where data already lives:** With a federated architecture, teams can query and analyze machine data across every silo simultaneously — from cloud and on-prem to edge — without copying or moving it first. This can curb the hidden costs of data movement and duplication.
- **Process data on demand:** It's expensive to structure all incoming data immediately, so wait to index it until you need to search it. This approach keeps your data in its original, complete form and eliminates the waste of preparing information that may never be used.

4 Observe everything agents do

AI agents won't always produce the same action twice, and they are known to hallucinate, act without accurate context, and produce harmful outputs. Because a single agent interaction can consume thousands of tokens, organizations literally pay for these mistakes. This unpredictability can also lead to distrust of agents.

- **Intervene when agents drift:** Systematically test how well an agent reasons, uses tools, and completes multi-step tasks. An evaluation tool can catch harmful outputs, like hallucinations or data leaks, before they reach an end user. Evaluation uses smaller, faster models which makes it economical to monitor agents continuously.
- **Centralize investigations for transparency and control:** Create a shared workspace where teams can validate AI recommendations and confirm that outputs reflect the business need as intended. This facilitates a clear, transparent history of AI-driven decisions while keeping humans firmly in control.

5 Implement strong governance for agentic actions

As autonomous agents move from experimentation to production, the risk of unintended consequences escalates. A governance strategy built around manual checks can't keep pace with agents acting continuously, at machine speed. Governance must be built into agents from the start.

- **Give agents consistent, secure access to what they need:** Adopt an open standard for connecting AI agents to your data and tools, so every agent interacts with your stack the same way, under the same controls, rather than through a patchwork of one-off integrations. This also means teams don't have to choose between agents that can act freely and agents they can trust.
- **Build in guardrails from the start:** Safely turn institutional knowledge into governed, agentic workflows by defining permissions and boundaries as the agent is built. This also reduces the risk of harmful outcomes when agents adapt to the data in front of them, rather than adhering to a rigid, predefined script.

Regional highlights

We identified insights across three key regions across the globe.

AMER

Organizations across the AMER region appear more capable of navigating the fragmented, distributed reality of machine data management. AMER respondents report greater effectiveness at discovering, querying, and accessing machine data across every environment — 66% are very effective in the cloud (versus 54% globally), 45% at endpoints (versus 40%), and 41% at the edge (versus 35%).

This operational maturity extends directly into how the region puts machine data to work. AMER respondents are more likely to use machine data for AI development, with 63% currently applying it to model training and fine-tuning, compared to 59% globally. Confidence follows suit: 58% are very confident in their ability to apply machine data to detecting and responding to failures before business impact (versus 48% globally), and 53% are very confident in ensuring governed access for all authorized teams (versus 46% globally).

The region also stands out in its ability to tie machine data back to the business. Thirty-six percent of AMER respondents report excellent linkage between machine data and business outcomes, compared to 30% globally, and 50% say they are very confident when making decisions based on correlated machine and business data, versus 40% globally.

Yet this progress is tempered by a sharper awareness of risk. Thirty-six percent of AMER respondents cite the increase in security threat exposure due to incomplete visibility as a major concern of managing machine data, compared to 30% globally.

APJC

Organizations across the APJC region largely mirror the global average — a pattern that reflects the report's broader narrative of enterprises collectively working to operationalize AI on a shared, still-maturing machine data foundation. Where APJC does diverge, the differences are modest but directionally consistent, pointing to a region that is neither leading nor lagging, but steadily aligning its machine data practices with the demands of AI at scale.

APJC respondents show a slight edge in how they structure their machine data environments. Fifty-nine percent say their organization's machine data is federated across multiple data environments, compared to 54% globally, and 60% report using machine data to train, fine-tune, or provide context for AI models, versus 56% globally. By making machine data searchable where it already lives, these organizations can preserve the operational context AI needs to reason across silos without incurring the cost and duplication of moving everything into a central repository.

This foundation appears to be paying dividends in how the region turns machine data into better outcomes. Fifty-seven percent of APJC respondents say they experience higher quality decision making as a result of improving machine data management, compared to 54% globally, and 32% agree that unified machine data as an AI input is a critical priority for the success of their AI strategy, versus 28% globally. This suggests that APJC organizations recognize the connection between coherent, well-managed data and the reliability of the decisions that depend on it.

EMEA

Organizations across EMEA trail global peers in unifying and accessing their machine data — a prerequisite to reliable AI. Just 48% say their machine data is federated across multiple data environments, compared to 54% globally, and effectiveness at discovering, querying, and accessing that data lags as well: only 45% are very effective in the cloud (versus 54% globally) and just 29% at the edge (versus 35% globally). As AI is increasingly generating telemetry at the edge — where organizations are least equipped to manage it — this gap is consequential.

These structural gaps carry through to how confidently the region says it can act on its data. Thirty-nine percent of EMEA respondents say they are very confident using machine data to detect and respond to failures before business impact, compared to 48% globally. AI depends on connected, contextual machine data to act reliably, and when that data remains siloed or hard to access, both teams and AI agents are left making decisions without the full picture.

EMEA organizations are also applying machine data to AI less intensively than their global counterparts. Forty-seven percent use machine data to train, fine-tune, or provide context for AI models, versus 56% globally, and just 39% say it is extremely important to harness machine data for their organization's ability to adopt and scale AI over the next two to three years, compared to 47% globally. As AI moves deeper into operational workflows, elevating machine data to AI-ready status will be an important step for EMEA respondents seeking to translate their AI investments into dependable, measurable outcomes.

Methodology

This report is based on findings from a double-blind online survey of nearly 4,380 senior technology decision-makers, conducted by Satori Experience Pte. Ltd. between May and June 2026.

The study spans 30 markets across Asia-Pacific, Europe, the Middle East, Africa, North America, and Latin America. It covers a broad range of industries including financial services, healthcare and life sciences, information services and technology, public sector, retail and consumer goods, media and communications, and education.

Respondents included C-suite executives, vice presidents, and directors responsible for technology, cybersecurity, cloud, infrastructure, data, AI, engineering, and risk functions.

Strengthen your digital resilience with Splunk

The CISO Report 2026: From Risk to Resilience in the AI Era

From adopting agentic AI to measuring security ROI and navigating an expanding role, discover how 650 global CISOs are maintaining resilience and empowering their organizations to thrive in the AI era.

Download the report

Perspectives by Splunk

Perspectives by Splunk — by leaders, for leaders

Looking for more insights on executive strategy, cybersecurity, and observability trends? Learn how leaders tackle today's most pressing challenges including AI and the changing compliance landscape.

Get executive insights

About Splunk

Splunk, a Cisco company, helps make organizations more digitally resilient. Leading organizations use our unified security and observability platform to keep their digital systems secure and reliable. Organizations trust Splunk to prevent infrastructure, application, and security incidents from becoming major issues, recover faster from shocks to digital systems, and adapt quickly to new opportunities.

All images in this report were AI generated.

26_CMP_report_Data-and-AI-report-2027_v17