


2026

# State of Observability

## Restoring Predictability to the Agentic Era

splunk> by CISCO

# Contents


4 Chapter 1: The end of predictability

8 Chapter 2: Agentic AI accelerates — and complicates — observability

12 Chapter 3: Agent observability emerges as a priority and a gap

16 Chapter 4: Leaders restore predictability with agentic AI

22 Chapter 5: AI transforms team dynamics

26 Chapter 6: Observability practices prioritize predictable costs




# Executive foreword

Applications used to be built on a foundation of certainty. We deployed features and updates in predictable, versioned cycles — or at least we tried to. We relied on golden signals to ensure service health and performance. And we knew, with absolute confidence, that the inevitable 2 am alert would be waiting for us exactly when things went sideways.

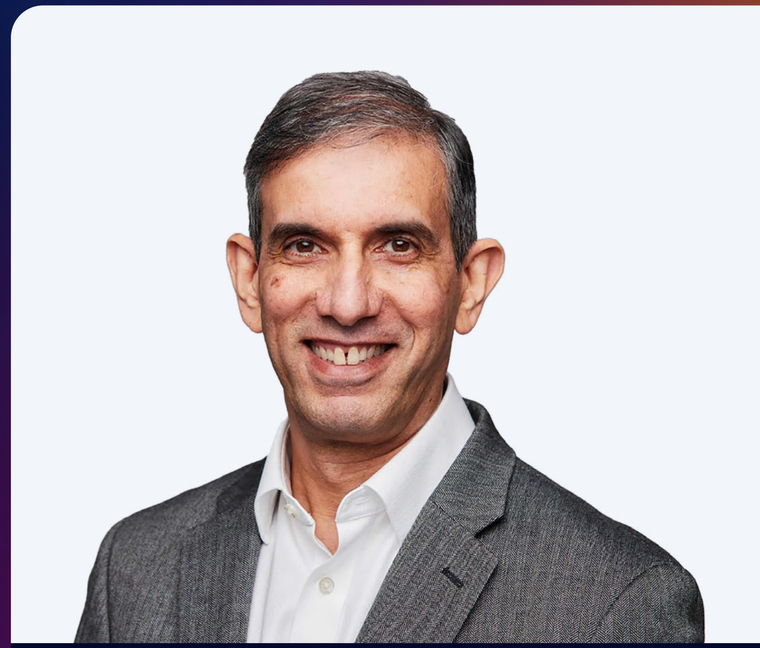
But the nature of the systems we build has been changing beneath our feet. The rise of agentic AI has ushered in a new era where systems don't just follow instructions; they adapt in real time towards a stated goal. The shelf life of playbooks and runbooks continues to decrease. And it's no longer enough to detect and remediate; we must find ways to prevent the unpredictable. This is a sign of progress, but it requires a new level of responsibility.

For the *State of Observability 2026*, we surveyed 2,050 ITOps and engineering professionals to understand how they are navigating this transition. The findings reveal a clear divide: while many organizations are searching for their footing, a standout group of leaders is already finding order in the chaos.

To thrive, these leaders are adopting three principles to form their strategies: leaning on AI agents to manage complexity, widening observability to include agent behavior, and connecting technical signals to business outcomes. Crucially, none of this is possible without a data foundation. You cannot govern what you cannot see, and the quality of your AI outcomes depends entirely on the data that feeds the system.

Leaders embracing this framework are already pulling ahead. They are collaborating more deeply with security teams, achieving higher ROI from their observability platforms, and proactively managing their data pipelines to keep telemetry costs predictable.

Observability now serves as the primary force for restoring predictability to our systems. It turns agent behavior into empirical knowledge, provides the governance and control required to trust AI, and keeps AI spend tethered to business value. Observability is the north star that empowers teams to navigate this new era with confidence.



Kamal Hathi
SVP and GM, Splunk

Chapter 1

# The end of predictability

The arrival of agentic AI marks a stark divide in the course of every observability practice: before and after.

The “before” rules of observability were written for a more predictable world — one in which you could trace the root cause of every performance dip, follow the breadcrumbs of each error to a specific fix, and trust that infrastructure health reliably mirrored customer experience.



Predictable operations have been slipping for a while. Dynamic infrastructure and tangled third-party dependencies chipped away at it long before AI arrived. The difference is that those factors were unpredictable by accident; AI is unpredictable by design.

As AI agents permeate nearly every digital environment, teams face a more nuanced class of telemetry. Agents choose their own paths in real time as an event unfolds, rather than follow a static script. Systems drift, hallucinate, and adapt in ways that seem to defy logic. Their non-deterministic nature is what lets them reason and act autonomously — and it's precisely what makes them so effective. It's also what makes them so challenging for teams that are used to tracking fixed, predictable paths.

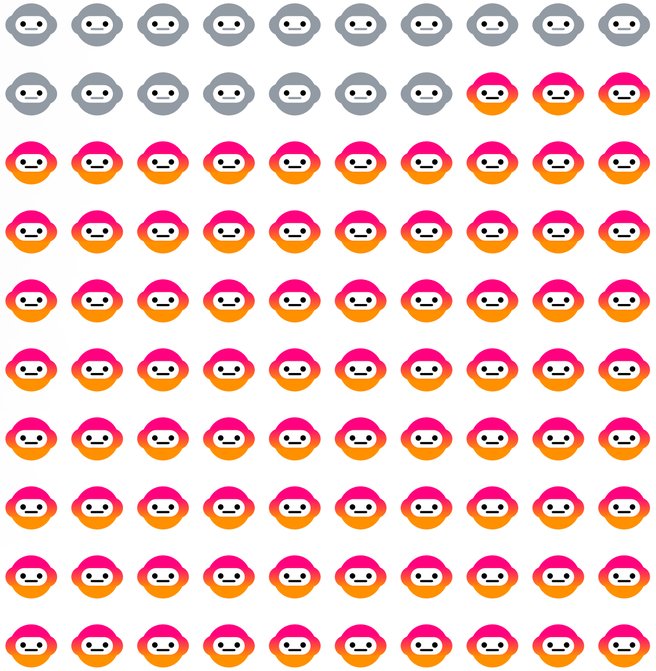
"Nothing is predictable anymore — not your infrastructure, not your apps, and certainly not how they're used," says Greg Leffler, director of developer evangelism at Splunk. "Now anyone can vibe out an app and toss it into production within an hour."

This shift isn't lost on the teams building these systems. There's a growing, industry-wide recognition that AI makes software easier to create and deploy, but agents can behave in ways that are difficult to predict, explain, or contain. A whopping 83% of respondents say they're concerned about the unintended consequences of AI agents.

The most successful teams understand that AI innovation and reliability are not opposing goals — but they must be governed together. Observability is where that work begins.

"Observability is one of the best training grounds enterprises have for AI. It teaches the skills every organization needs to use AI effectively: reading signals, applying context, governing outcomes, and learning continuously," says Mark Woods, field CTO and chief technical advisor for EMEA at Splunk. "If you want your business to use AI with confidence, start with observability."

> **Nothing is predictable anymore — not your infrastructure, not your apps, and certainly not how they're used.**
>
> Greg Leffler, Director of Developer Evangelism, Splunk



**83%** worry about the unintended consequences of AI agents

# Managing the data deluge

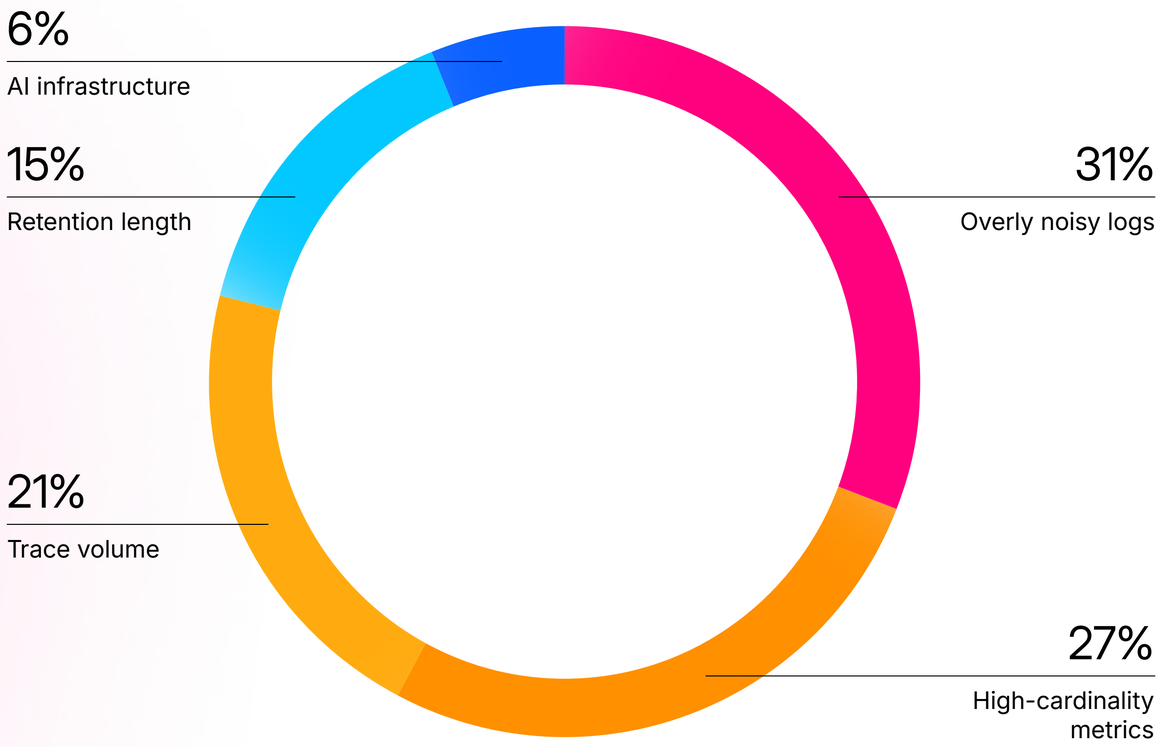
Meanwhile, teams using observability tooling are dealing with a deluge of data — a tale as old as time (and only intensifying). Sixty-seven percent say the number of sources sending data to their observability platform — apps, containers, and servers, for example — has increased, with 26% say the increase is *significant*. This data surge spans the entire stack. It's no longer just the golden signals of classic infrastructure and applications like latency, traffic, and errors, but also prompt-response streams that reveal how AI agents reason and act.

"When we first broke our monoliths into microservices, ITOps and engineering teams were suddenly tasked with capturing every hop, every handshake, and every latency spike across a fragmented landscape just to understand how services were talking to one another," says Alpesh Sheth, vice president of engineering at Splunk. "Today, AI is pushing that demand for data into a different dimension."

Tracing a system's behavior has always required telemetry, but AI raises the bar for what that telemetry needs to explain. Systems now change faster than before, and parts of them are intentionally non-deterministic. Understanding what happened, why it happened, and whether the outcome can be trusted requires richer context, higher fidelity, and more ways to connect data across the environment.

Yet, when it comes to observability costs, AI infrastructure isn't the pressure point respondents cite most often. Training and running LLMs may still be expensive, but in this context, only 6% identify AI infrastructure as a top observability cost driver. The bigger issue appears to be inefficient data practices. Overly noisy logs — think recording every successful page load when only the failures really matter — emerged as the main culprit, with nearly a third (31%) of respondents saying this was their biggest observability cost over the last year.

## Top drivers of observability costs this year



This reliance on indiscriminate logging is exactly why it's so difficult to define "good" data in a world of autonomous systems.

"When you're faced with the challenge of identifying the needle in the haystack before an incident even begins, sometimes keeping the entire haystack feels like the most rational choice," says Steve Flanders, senior director of engineering at Splunk.

Telemetry only earns strategic value when organizations can tie it to business impact. Imagine an apparel retailer launching its biggest product drop of the season. With observability tied to business impact, the team can see that a 200-millisecond slowdown on the payment page is costing them conversions during the exact hour demand peaks — and act on it before revenue slips away.

Yet as teams fixate on technical logs — an expensive habit in its own right — that critical piece of the puzzle remains out of reach. Sixty-two percent point to inaccessible or difficult-to-parse data as the main obstacle to connecting telemetry with business outcomes.

For years, it was accepted industry practice to limit access to business data to limit risk of insider trading or leaks about company performance. But organizations are realizing the value of ITOps and engineering teams understanding how their work impacts the business.

This disconnect has very real consequences. When teams can't tie technical data to commercial results, telemetry never makes it into the decisions that shape the product, the roadmap or the business. Nearly half (46%) of respondents admit that telemetry data is not a key factor in their organization's strategic planning process.

"Strategic planning without operational telemetry is just guessing," says Patrick Lin, senior vice president and general manager of observability at Splunk. "Integrating technical insights into business decisions ensures that your tech strategy and business goals are actually pulling in the same direction."

> **When you're faced with the challenge of identifying the needle in the haystack before an incident even begins, sometimes keeping the entire haystack feels like the most rational choice.**
>
> Steve Flanders, Senior Director of Engineering, Splunk

**46%**

**say telemetry is not a key factor in their organization's planning process**

Chapter 2

# Agentic AI accelerates — and complicates — observability

Just one year ago, many organizations viewed agentic AI for observability as an experiment. Now, it's quickly becoming a core component of the modern stack.

While many are still exploring how agentic AI fits into their observability practices, 34% say they plan to start using it soon, suggesting we're on the precipice of another steep adoption curve.



# Agents perform the heavy lifting — with guardrails

In last year's *State of Observability* report, we anticipated a sharp rise in agentic AI adoption, including a future where AI agents would take on more of the incident workflow. That prediction has come to fruition today, as "AI SREs" tackle the troubleshooting that used to eat up practitioners' weekends. Most notably, teams are using AI agents *frequently* or *all the time* to determine root causes (62%), continuously improve alert configuration (58%), and write runbooks (57%).

As systems have become more distributed and interdependent, the burden of connecting every signal has grown heavier. Now teams lean on agentic AI to do that heavy lifting. These top use cases suggest that teams are only delegating tasks to agents when a human is still the final authority. Tasks like alert tuning include operational guardrails in which AI suggests or updates the rules, but humans set the thresholds.

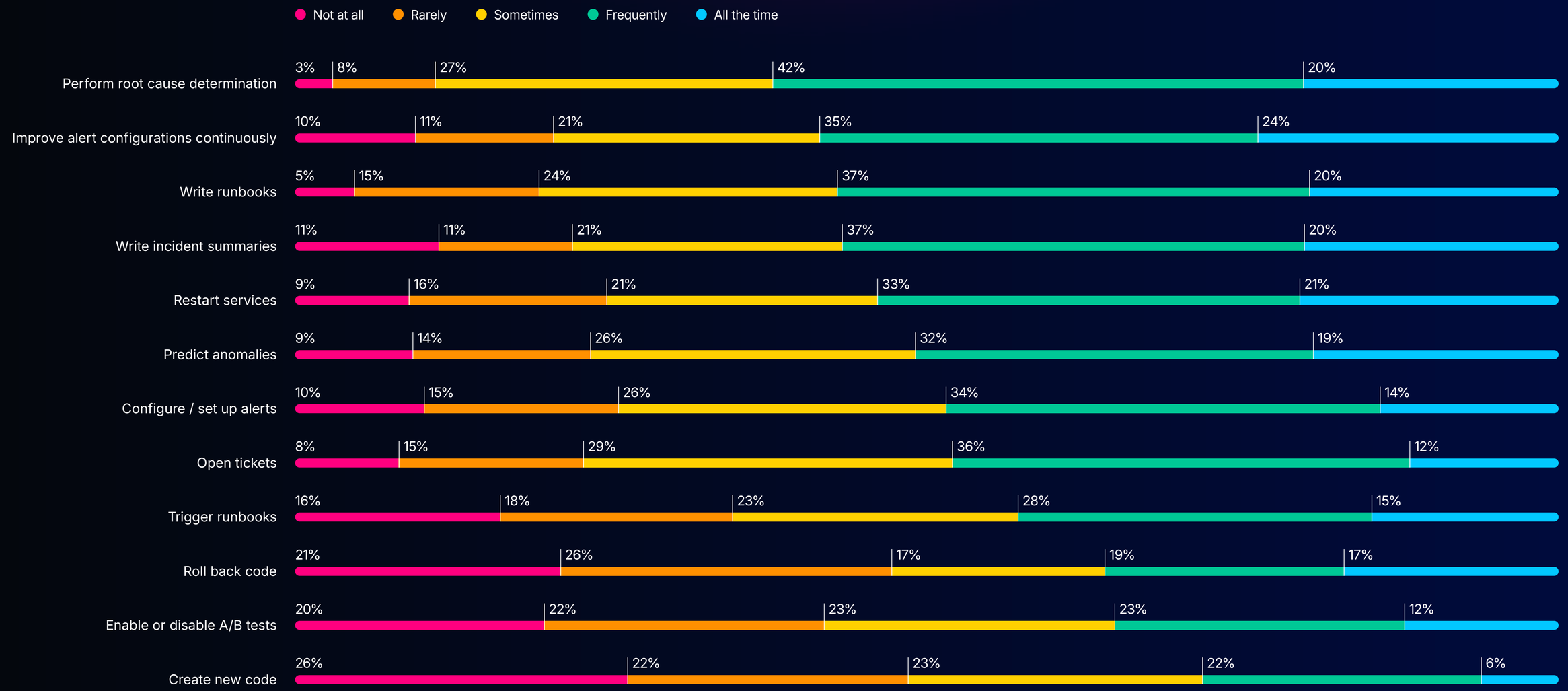
Agentic AI is particularly suited to root cause determination because it can draw on the entire history of every past incident, like an SRE with a photographic memory. That gives teams a significant head start, bringing a new layer of institutional knowledge and surfacing patterns across massive volumes of telemetry far faster than manual investigation would allow.

Teams are more comfortable with AI tuning existing systems than creating new ones. Code creation ranked as the least common use case — only 6% say they use agentic AI to write new code *all the time*, and over a quarter (26%) say they *never* use it.

It may take time for AI agents to earn enough trust to handle changes that could affect production. The clearest signal is a comparative one: respondents are more than twice as likely to say AI is over *five years away* from reliably writing new code (14%, plus 2% who say never) as they are to say the same about improving alert configurations (6%). The takeaway isn't that most teams expect a long wait — it's that they're measurably more cautious about code creation than about tasks they can easily undo. It's a clear hierarchy of trust: teams are willing to give AI more autonomy when the task is reversible. Given how quickly agentic coding capabilities are advancing, this hierarchy of trust may shift faster than the data suggests — as AI models continue to improve at code generation, teams' comfort with production-level autonomy is likely to grow in step.

# AI agents are trusted more to diagnose than to act

How often respondents deploy agents for specific observability use cases




Percentages may not add up to 100 due to rounding

# Agentic AI evolves from searching to solving

Incident response is where agentic AI's strengths — and limits — come into sharper focus. AI agents excel at relentless, round-the-clock detection, with most respondents (86%) reporting improvements in mean time to detect (MTTD) due to their adoption of agentic AI, and 36% classify the improvement as *significant*.

Remediation tells a more layered story. While fewer respondents say agentic AI has improved mean time to recovery (MTTR) (72%), the depth of improvement is greater: 42% say MTTR *significantly* improves with agentic AI. This is a massive win for an SRE. For nearly half of respondents using the technology, they're moving faster during moments that threaten customer experience, revenue, or trust.

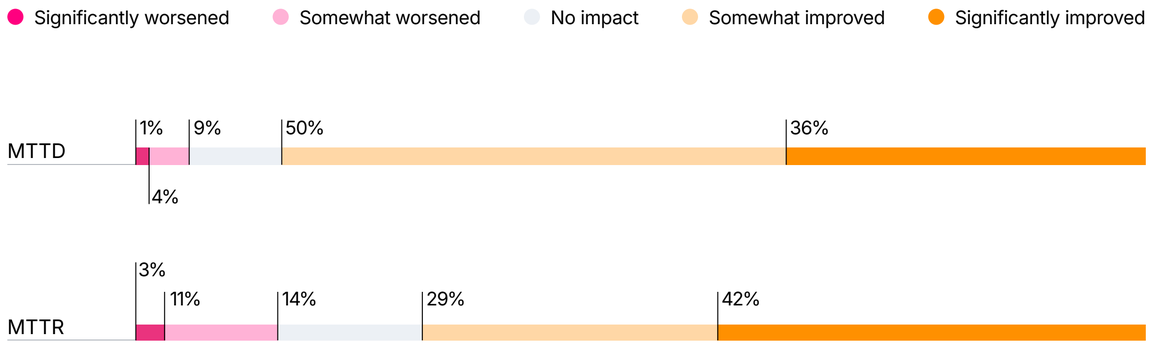
As AI improves detection and resolution, teams are reclaiming time for higher-value work; 84% say AI enabled their teams to focus more on innovation than maintenance, up 6% from last year. It might seem like a modest increase, but it's an indicator of compounding returns; as teams get more comfortable with AI agents, they spend less time in the weeds and more time building the experiences and features that drive value for customers, and ultimately, the business.

But agentic AI isn't a shortcut for speed. While most respondents see improvements in MTTR, 14% admit agentic AI made it *somewhat* or *significantly* worse.

Splunk experts believe this could be a symptom of integration friction. "Introducing agentic AI into your observability stack is essentially like onboarding a new teammate," says Sahinaz Safari, head of artificial intelligence, observability at Splunk. "Like any new hire, they come with incredible potential as well as a period of more oversight as you integrate them in your environment."

This extra time double-checking the agent's work and aligning its "thought process" with the team's standard — often called the verification tax — can slow things down at first. It pays off once teams stabilize their models and establish robust context frameworks that turn individual judgment calls into repeatable, deterministic processes.

## Agentic AI speeds detection and resolution cycles




Percentages may not add up to 100 due to rounding

Chapter 3

# Agent observability emerges as a priority and a gap

Deploying AI agents without observability is a gamble. As organizations ingrain AI agents into workflows that influence everything from software development to customer experience, the risk is not just technical glitches but business failures.



An AI coding assistant accidentally wipes out a company's entire production database. An insurance company indiscriminately denies healthcare claims due to a prompt error. A car dealership's chatbot accepts a customer's offer to purchase a brand-new car for $1.

These snafus represent an important new responsibility for observability practices: prevent AI from going off the rails. Most respondents recognize this, with 87% confirming that monitoring AI is a priority for their organization.

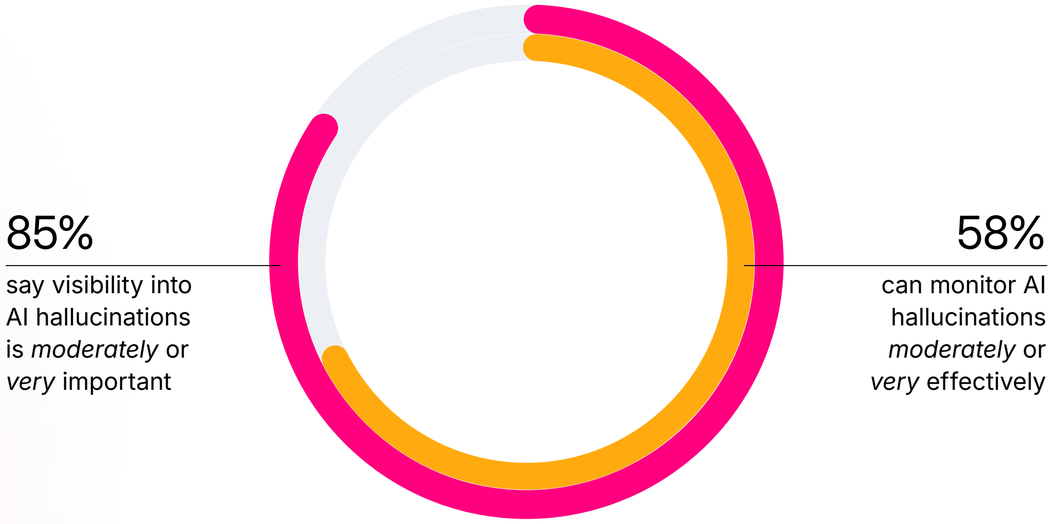
## Bridging the measurement gap

Yet many respondents admit they struggle to turn this priority into action. Over half of respondents (56%) reveal they're unable to measure workload metrics like LLM quality, with 29% ranking it as their number one challenge in agent observability. And, although 85% say it's *moderately* or *very* important to have visibility into AI hallucinations, only 58% claim they're *moderately* or *very* effective at doing so.

Since the dawn of the mainframe when we first monitored blinking lights on a console, observability has measured the quality of our systems through binary, objective facts.

"Historically we've been interested in how many requests, how much latency, and how many errors," says Vikram Chatterji, CEO and co-founder of Galileo and now senior director of product management at Cisco. "But AI agent behavior is subjective and therefore requires a new class of behavioral metrics that are tied to business outcomes and agent decision making."

### Hallucination priorities exceed monitoring capabilities




> **AI agent behavior is subjective and therefore requires a new class of behavioral metrics that are tied to business outcomes and agent decision making.**
>
> Vikram Chatterji, Senior Director of Product Management, Cisco

Measuring these softer signals well means catching problems on two fronts: before AI reaches production, and while it's running. On the front end, that's evaluation —systematically testing an agent's behavior and outputs against expected standards before it ever touches a live environment, so teams can catch flawed reasoning or low-quality responses early. On the back end, it's guardrails — the runtime controls that catch and block inaccurate or harmful outputs in the moment, before they reach a customer.

Until recently, this kind of monitoring was difficult because tooling wasn't built for it. Workarounds, such as using one model to evaluate another, could be costly and difficult to scale. About half (49%) of respondents say that their observability tools lack the capabilities to monitor AI, making it the second highest challenge.

Fortunately, agent observability technology is evolving fast, making it more possible for teams to govern their AI models with the same precision they've always applied to infrastructure.

## Agent observability crosses organizational boundaries

Beyond the technical hurdles, agent observability has yet to find a permanent home in the organizational chart. Thirty-eight percent say their organization does not clearly define ownership and accountability for monitoring AI, with only about a quarter (27%) *strongly* agreeing their responsibilities are well-defined.

Agent observability cuts across every traditional IT boundary. Is it the responsibility of the data science team that built the model? The SRE team that manages the infrastructure? The security team that worries about the risk? This uncertainty tracks with survey findings: 42% say their main challenge in observing AI is that a different team is responsible.

"When agent observability is everyone's job, it becomes nobody's job," says Mark Woods, field CTO and chief technical advisor for EMEA at Splunk. "The fix is structural: accountability through governance, responsibility through defined capabilities, and interoperability through shared standards."

Where should that ownership begin? SRE teams are a logical starting point for agent observability, our experts propose. "They already understand how to instrument and observe complex, highly distributed systems," says Moss Normand, staff developer evangelist at Splunk. "SREs are well-positioned to lead this effort, but ownership has to be bigger than any one team."

In practice, SRE coverage is rarely universal — most teams concentrate on the highest-value services, and some organizations don't have a formal SRE function at all. Organizations should frame the decision around maintaining resilience by making a board-level stakeholder accountable and defining which function leads.

> **When agent observability is everyone's job, it becomes nobody's job. The fix is structural: accountability through governance, responsibility through defined capabilities, and interoperability through shared standards.**
>
> Mark Woods, Field CTO and Chief Technical Advisor for EMEA, Splunk

# AI tokenomics tests financial control

Clear ownership matters just as much on the financial side. Observability is what makes AI spend visible, attributable, and controllable. Without it, costs compound in darkness.

For three years, AI ran like a developer preview with pilot projects, flat-rate plans, and subsidized inference. That era is ending. Spend now tracks directly to tokens, and usage is increasingly hard to predict.

“The instinct is to assume falling token prices mean falling costs. The opposite is true,” says Atindriyo Sanyal, co-founder of Galileo and now senior director of engineering at Cisco. “Consumption is growing faster than prices are dropping, so total spend keeps climbing. We’re heading into the biggest operational-expense shift since cloud FinOps.”

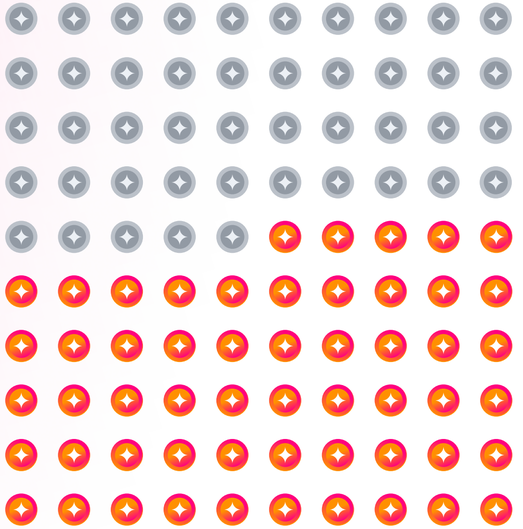
The data bears this out. Over half (55%) say their organization cannot accurately control AI costs.

Every single prompt, re-run, and token generated carries a price tag. When organizations run dozens of agents across multiple models, the costs stack up in ways that are incredibly difficult to track in real time.

“The question isn’t just what your AI costs — it’s whether that cost is producing anything worthwhile,” says Patrick Lin, senior vice president and general manager of observability at Splunk. “Without visibility into agent behavior, the performance of the entire AI stack, and token consumption, teams can’t connect spend to outcomes. And if you can’t make that connection, every dollar looks the same — whether it’s a good investment or pure waste.”

Additionally, most respondents haven’t defined what success actually looks like. Fifty-four percent say their organization doesn’t have clear business outcomes for monitoring AI. This suggests that for many, AI initiatives haven’t progressed to enterprise-grade maturity.

The path forward is to tie every AI signal to a business decision. What does monitoring model quality let you do — ship faster, avoid a customer-facing issue, control runaway costs? Define that link so that monitoring AI becomes a measurable business asset.

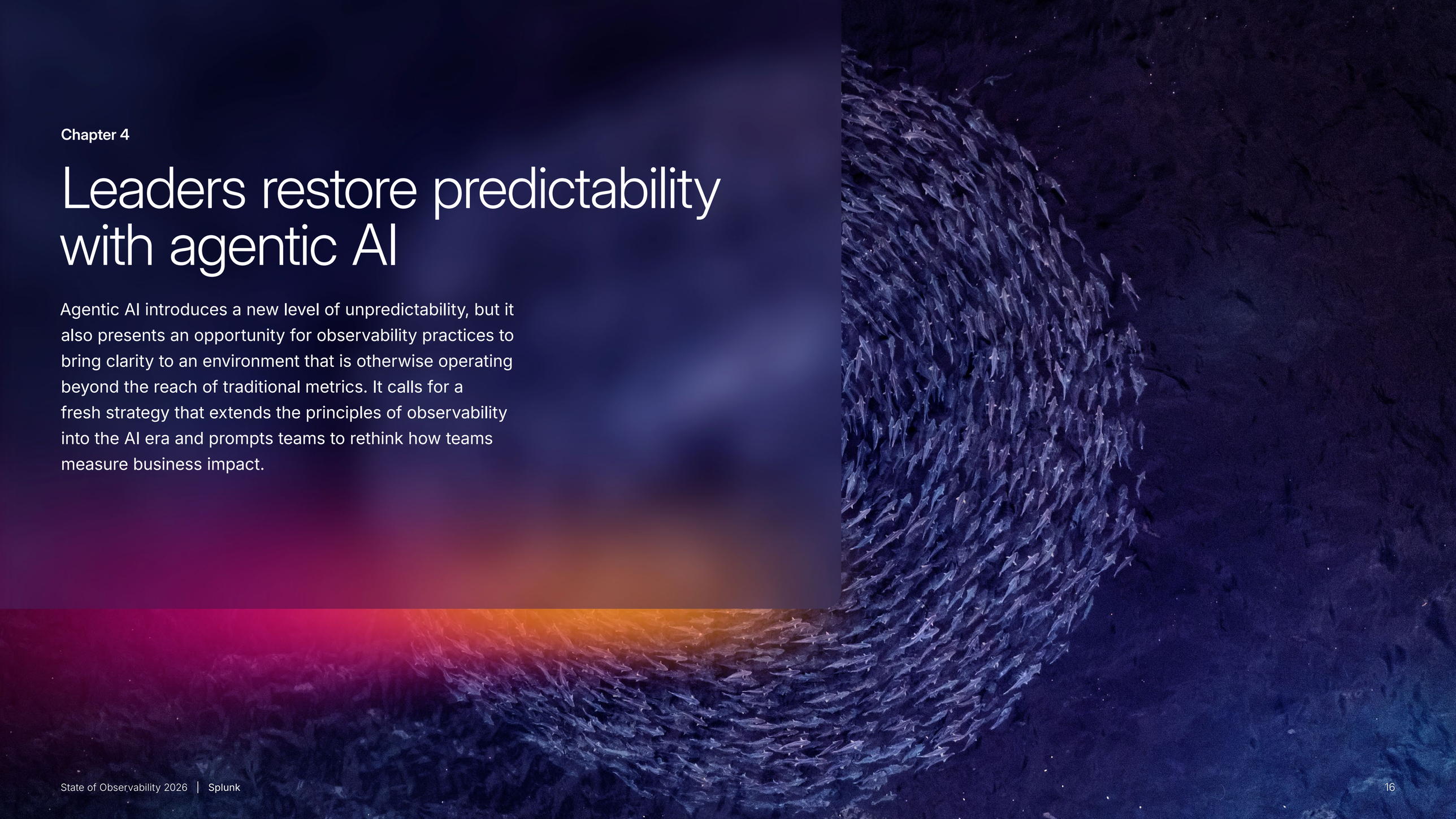


**55%** say their organization cannot accurately control AI costs

Chapter 4

# Leaders restore predictability with agentic AI

Agentic AI introduces a new level of unpredictability, but it also presents an opportunity for observability practices to bring clarity to an environment that is otherwise operating beyond the reach of traditional metrics. It calls for a fresh strategy that extends the principles of observability into the AI era and prompts teams to rethink how teams measure business impact.



A distinct group of organizations is already getting this right. Throughout this chapter, we refer to them as leaders: teams that follow three specific practices, defined below. The data shows this cohort achieves better outcomes on nearly every dimension we measured.

"This is a cultural declaration as much as it is a necessary evolution of how we work," says Lin. "Advanced organizations recognize that the key to scaling AI is building a foundation of trust that allows us to map the unpredictable and ship with confidence."



## What defines a leader

We define leaders by three practices that, put together, set them up for success in the agentic AI era.

1

They rely on AI agents to fix and prevent issues — and they do so frequently, not just as a one-off. These observability leaders use agents to perform remediation tasks like writing and triggering runbooks, rolling back code, determining root causes, and restarting services.

2

They have widened their scope to include observability of AI infrastructure and services. Leaders monitor the factors that keep AI in check — model quality, cost, usage, security, and privacy — and they do this *very* effectively.

3

Leaders connect signals to business impact. Observability leaders strongly agree that telemetry is a key factor in their organization's planning process, reflecting a more mature observability practice: one that supports day-to-day reliability while also shaping strategic decisions.

# Leaders escape the grind

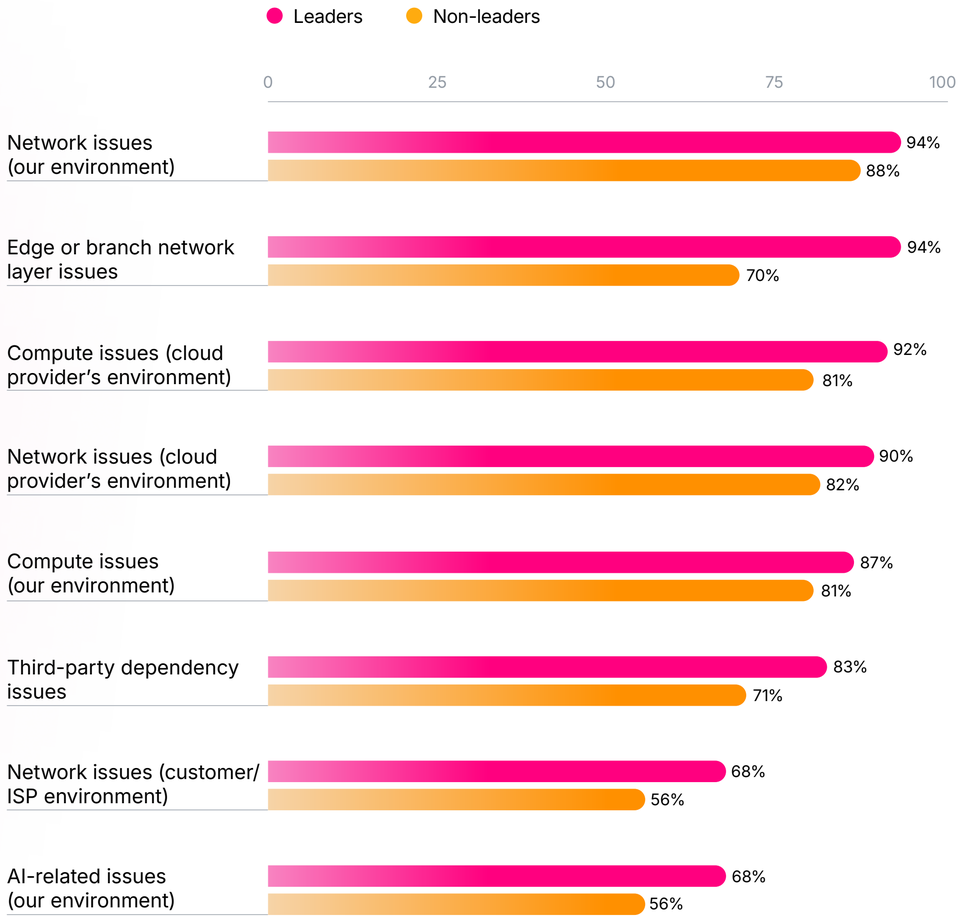
The payoff materializes first in the daily rhythm of incident response, where leaders are fundamentally changing how their teams solve problems.

The impact is dramatic. While 40% of non-leaders report *significant* MTTR improvements with agentic AI, that figure jumps to 63% among the leaders — suggesting that leaders have learned to trust agents with diagnostic heavy lifting, moving from manual investigation to automated resolution with confidence.

This confidence is reflected in how these teams approach the most difficult technical challenges. Leaders consistently report being *moderately* or *very* confident in their ability to isolate the root cause of infrastructure and application issues, outpacing their peers. While most teams have developed the necessary expertise to keep the lights on in their core environments — the difference is imperceptible between leaders and all others for compute issues (87% vs. 81%) and core network (94% vs. 88%) — leaders differentiate themselves at the edge and within the black box of AI logic. Nearly every leader (94%) is *moderately* or *very* confident in their ability to isolate edge or branch network layer issues, compared to just 70% of their peers. And 68% say the same for AI-related issues in their environment (versus 56% of other respondents). Leaders recognize that the most fragile points in the customer experience now sit close to the edge and inside AI-driven decision making, so they are investing accordingly.

## Leaders more confidently isolate root causes

Respondents who say they're *moderately* or *very* confident isolating the root cause



The impact extends to team culture, as leaders spend less time fighting fires and gain more bandwidth for high-value work. Leaders are much more likely to say that AI allows them to prioritize innovation over routine maintenance, with 37% *strongly* agreeing compared to just 22% of their peers. Almost all leaders (94%) agree that AI has reduced the on-call burden for their team, compared to 79% of other respondents.

## Leaders make observability a team sport

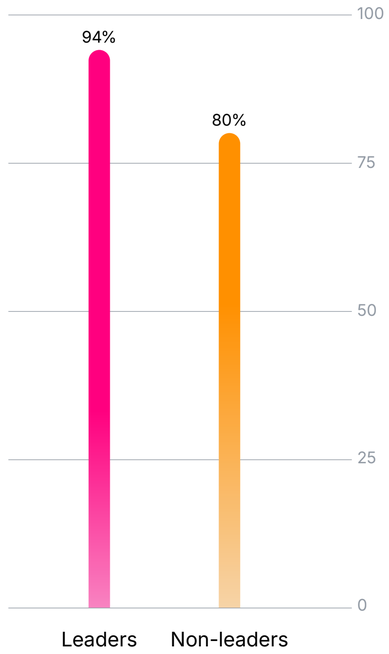
Agentic AI fuels a cultural shift that extends beyond the network operations center (NOC) — and these changes are playing out more prominently for leaders. Monitoring AI workloads has pushed leaders to work more closely with their security counterparts — nearly all (94%) say it has increased collaboration with security teams, compared to 80% of non-leaders.

There's a reason the two functions are converging (and we'll talk about this more in the next chapter). In the agentic era, security and observability have become two sides of the same coin. Picture an AI agent that spirals into an infinite loop, spinning up compute and hammering an internal API thousands of times per minute. To the engineering team, it's an observability failure. To security, it looks like a denial-of-service event. But it's one incident, not two. As agents gain autonomy, the gap between a performance problem and a security threat narrows to almost nothing.

That's why this collaboration is more than goodwill; it changes how problems get solved. Leaders are more likely to accurately trace performance issues back to their security root causes (86% vs. 71%), connecting dots across domains that siloed teams would miss. For leaders, observing AI is a shared mission that breaks down walls between ITOps, engineering, and security.

### Leaders collaborate more due to agent observability

Respondents who agree that monitoring AI has increased collaboration with security teams



# Leaders turn data discipline into business value

For some organizations, AI is a shiny add-on bolted onto existing workflows. But for leaders, AI is a strategic mandate incorporated into the very fabric of their business strategy. They've moved past the novelty phase and into a reality where AI drives operational and financial performance.

The results of this mindset are clear. Leaders are 2.4 times more likely than other respondents to declare the ROI of their observability platforms *significantly* exceeds expectations. Over half of leaders (54%) say their observability investment yielded at least $5 million in annual business value, versus 28% of non-leaders.

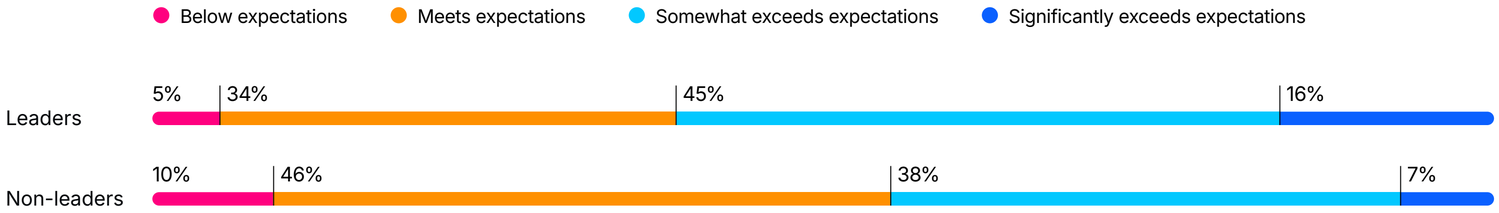
So how do they get there? One possibility is that leaders are more disciplined about what data they keep. Rather than collecting and storing everything — an expensive habit that buries useful signals in noise — leaders are deliberate. They keep high-value data that drives decisions and trim the rest, using techniques like sampling (keeping a representative slice of data instead of every record) and routing (sending data to the right storage tier based on its business value, urgency, and retention needs). Leaders are far more likely to *moderately* and *significantly* use these methods, including sampling (91% vs. 75%) and routing/load balancing (82% vs. 65%).

These techniques actually create more visibility, not less. By being selective about what they store, leaders ensure the data they do keep is high quality. Leaders are over twice as likely as their peers to say all their observability data is readily accessible and immediately queryable (30% vs. 13%).

This discipline translates straight to the bottom line; leaders are *significantly* less likely to see their observability costs exceed their operational budgets (45% vs. 58%), and to be surprised by telemetry bills (45% vs 51%). Leaders' deliberate approach connects to results at the highest level: 64% rate their organization's revenue growth as *very* strong, compared to just 40% of their peers.

## Leaders experience better observability ROI

Respondents rate the annual ROI of their observability investments



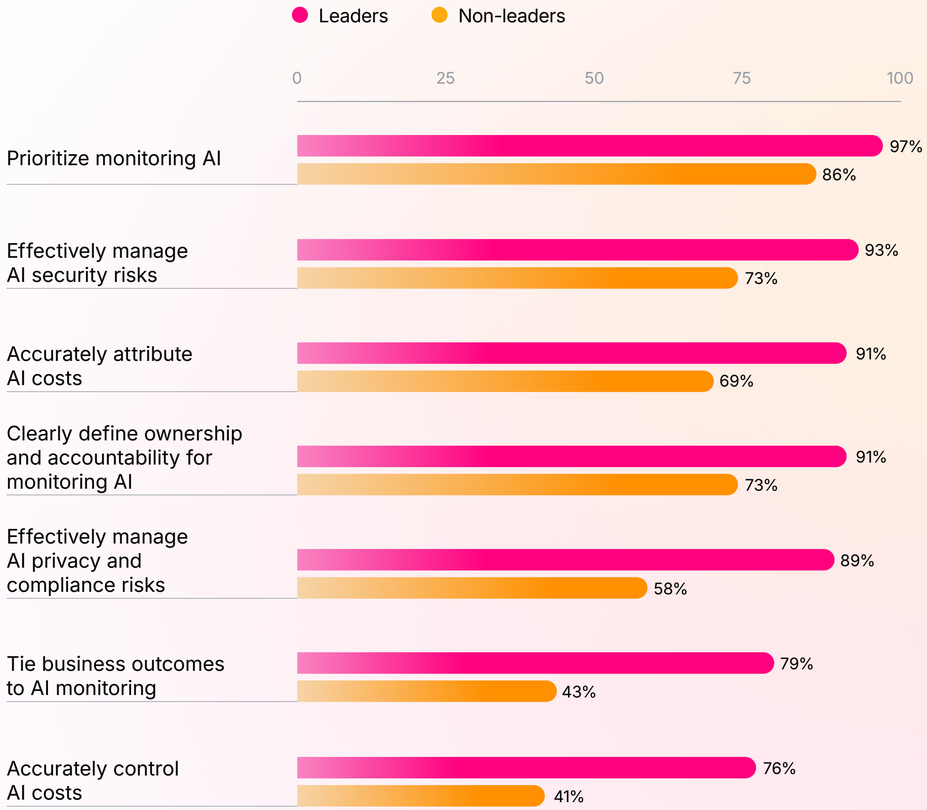
# Leaders govern AI to accelerate

For leaders, guardrails are key to reaping the benefits of AI. While monitoring AI workloads is more challenging for leaders — 57% say it makes their jobs more complex, compared to 38% of their peers — these leading teams lean into that complexity, recognizing that the only way to safely accelerate with AI is to prioritize governance and accountability.

Leaders turn AI from a liability into a competitive advantage. A full 91% clearly define ownership and accountability for monitoring AI, compared to 59% of their peers. They also maintain a firm grip on AI costs, with 76% of leaders saying they can accurately control them, compared to just 41% of their peers. The source of that control is visibility; leaders are far more likely to say they can accurately attribute AI costs (91% vs. 69% of others), pinpointing which agents, models, and workloads drive spend.

This approach pays off in a few ways. For one, trust. Leaders are more likely to trust AI with mission-critical tasks (86% vs. 75%). They also treat monitoring as a means to an end, not an end in itself — 79% say monitoring AI is tied to clear business outcomes, versus 43% of others. And because they can actually measure whether it's working, the benefits build up: 77% of leaders say AI improves their observability practice overall, versus 57% of their peers.

## Leaders have tighter control of AI risks and costs



Chapter 5

# AI transforms team dynamics

We tend to picture the AI-driven workplace as cold and clinical, where the whirr of machines replaces the chatter of a team solving a problem together in real time. In reality, AI is fostering a more collaborative culture. Because agents can operate within every layer of the stack, monitoring them naturally leads to more interaction between ITOps, engineering, and security teams, eroding the walls that have historically kept them apart.

In fact, 81% say that monitoring AI workloads has increased collaboration with security teams, with 51% *strongly* agreeing.

"Agent observability is the ultimate team sport," says Cory Minton, global field CTO at Splunk. "ITOps, engineering, and security continue to own their own specialized missions, but they're sharing the same data to solve big, complex puzzles together."

Our research shows this cultural shift taking hold. More ITOps and engineering teams are troubleshooting and solving issues with security teams according to 77% of respondents, versus only 62% in 2025. Data sharing and reuse between teams has also climbed from 74% to 84%.



> **Agent observability is the ultimate team sport. ITOps, engineering, and security continue to own their own specialized missions, but they're sharing the same data to solve big, complex puzzles together.**
>
> Cory Minton, Global Field CTO, Splunk

**81%**

agree that monitoring AI workloads has increased collaboration with security

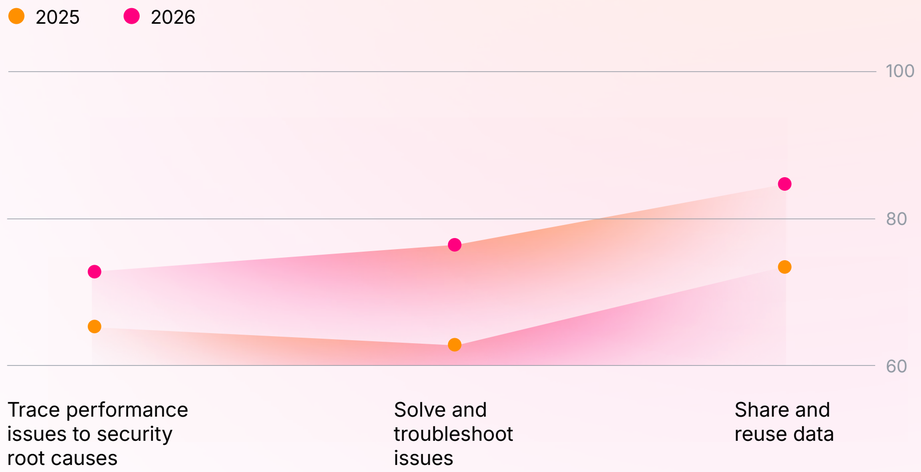
Consider a customer-facing AI agent that suddenly leaks sensitive user data. Security spots the anomalous input pattern, engineering traces the agent's unexpected behavior, and ITOps flags a sudden spike in token usage and GPU compute costs. In silos, these are isolated alerts. Without shared telemetry, those alerts can look like three separate issues. With a shared view, teams can connect them earlier, understand how the incident is unfolding, and respond as one coordinated function rather than three separate teams. Even if teams cannot apply that context in the moment, they can use it after the incident to reconstruct what happened, quickly pinpoint where the breakdown occurred, and prevent the issue from recurring.

But that upside isn't automatic. "Collaboration without structure just spreads the pain around," Minton adds. "Define clear ownership, set explicit handoffs, and — especially where sensitive or regulated data is involved — make sure the right people have the right access, not everyone. Otherwise, you trade silos for compliance risk and turn collaboration into a permanent, all-hands fire drill."

For teams that put governance in place, the results speak for themselves. Shared context becomes the key to connecting the dots, allowing teams to trace performance issues back to security root causes with newfound precision. In fact, 73% of respondents can now accurately trace performance issues to security root causes — a significant jump from 65% last year.

## Collaboration surges in the agentic era

How observability and security work together



# Teams reclaim balance through trusted AI

These positive changes are made thanks to a foundation of growing trust in AI. Over three-quarters (76%) of respondents now trust AI to handle mission-critical tasks, up from 71% last year.

This is a major milestone in AI maturity. But we should be careful not to mistake enthusiasm for rock-solid conviction. The level of trust likely reflects a mix of genuine confidence and the intense pressure to keep pace with the AI gold rush.

"Even on the most mature teams, 'trust' usually means 'trust but verify,' with engineers keeping a tight grip on the guardrails," says Craig Robin, field CTO at Splunk. "Teams are extending that trust faster than ever, but we're still early in proving it through consistent, measurable performance — not just optimism."

Trust in AI and trust between teams reinforces each other, and the result is a more unified front. That's a big win for team morale, enabling organizations to chip away at the "always-on" culture that has defined the industry for decades. Eighty-one percent of teams report that AI has already reduced their on-call burden.

"When the pager stops screaming at 2 am because an agent successfully remediated a known issue, everybody wins," says Christine Low, head of observability, APJC at Splunk.

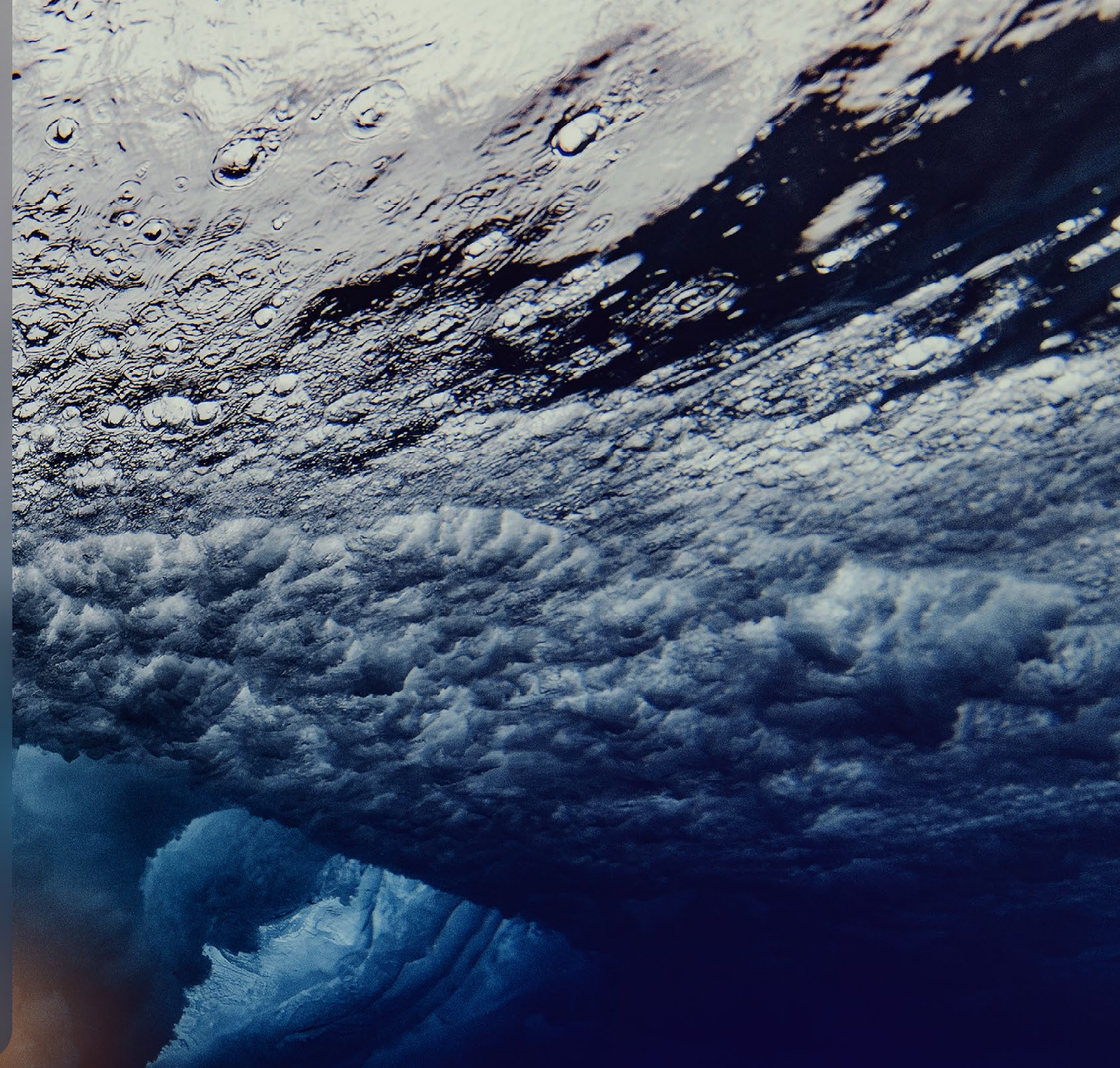


> **When the pager stops screaming at 2 am because an agent successfully remediated a known issue, everybody wins.**
>
> Christine Low, Head of Observability, APJC at Splunk

Chapter 6

# Observability practices prioritize predictable costs

For years, the approach to telemetry was straightforward: collect everything, store it forever, and ask questions later. It felt like the ultimate safety net, but the “collect everything” strategy has reached a breaking point as AI exacerbates a volume problem that makes sifting through mountains of data impossible.



Sixty-seven percent say the number of sources sending data to their observability platform increased over the past year — a number that will almost certainly climb as organizations deploy more autonomous agents. Every agent is both a source and a consumer of machine data because it emits telemetry about what it's doing and depends on a steady stream of context to do it well. That two-way demand multiplies data volume fast — and with it, the pressure to keep costs predictable.

## Unpredictable costs strain operational budgets

Telemetry costs are on an inevitable upward trajectory, making financial predictability more critical than ever. When observability bills eclipse the cost of the actual infrastructure being monitored — as they did for 57% of organizations this year — it's worth asking why. Sometimes the answer is sound: complex or heavily regulated systems can warrant monitoring costs that exceed operational ones. But just as often, it's a sign the practice is paying for noise rather than visibility.

Pricing structures only add to the financial strain, with half of respondents (50%) experiencing surprise costs with their telemetry. Budgets are inherently based on predictability, but utility-based pricing models introduce a level of volatility that makes it difficult for organizations to commit to a long-term plan.

For 60% of respondents, the primary operational strain comes from unexpected billing spikes rather than high baseline costs.

"It's like shopping for groceries where the price of a single tomato could jump from two dollars to ninety dollars overnight… and you don't get the bill until the end of the month," says Leffler. "In a store, you'd see the price and skip making pasta that week. But you can't skip observability, which is why pricing predictability is so important."

These issues have significant consequences. Due to rising telemetry costs, 38% say their organization had to reduce its data coverage and/or retention. And 11% say their organization is at greater risk of downtime due to cost-cutting initiatives — a small percentage, but one that exposes a critical vulnerability: when observability becomes a casualty of budget management, the systems that businesses rely on become significantly harder to protect.

## Precision improves data predictability

Teams are regaining control through more surgical approaches. Sampling — keeping a high-value subset of data instead of retaining every data point — is the most popular choice, with over half (51%) of respondents reporting that data sampling *significantly* helps control telemetry costs. This discipline is easier to maintain when data is consolidated; currently 67% of respondents use a single observability platform, providing a unified view that makes it easier to pinpoint exactly where costs are coming from.

However, sampling is just one piece of the puzzle. To truly harness the economics of observability, teams must shift the conversation with leadership from simple cost management to the delivery of business value.

"Every spending decision comes back to ROI," says Leffler. "If you can show that observability isn't just monitoring infrastructure — it's telling you how the whole business is doing — you've made the case for the investment.'"

## Surprise costs versus high costs: Which factor strains operations more?



**40%**

paying a high baseline cost for ingesting and storing data

**60%**

receiving bills that were significantly higher than expected

# Navigating the agentic future

Observability is how teams make unpredictable AI systems understandable enough to find and fix issues faster, fuel innovation, and move the business forward — even when logic takes an unexpected turn. The findings in this report are clear: the most successful organizations are bringing order to autonomous systems and turning the complexity of AI into a competitive advantage. And for those still waiting, the message is just as clear: AI adoption won’t slow down while observability practices catch up.

We’ve distilled the findings into four key strategies.



## 1. Treat AI quality as a business-critical metric

When agent observability stays exclusively within engineering, risks like bias, hallucinations, and poor-quality outputs often remain hidden until they affect customers, eroding customer trust, damaging reputation, and creating compliance concerns. Instead, treat AI quality as a business-critical metric by forming an AI “center of excellence” and bringing together stakeholders from engineering, security, legal, and compliance to define what “good” looks like. This center of excellence should set shared standards for model quality, cost, and safety. Just as important is where you measure it. Keeping AI quality signals in the same view as the rest of your telemetry lets teams connect model behavior to real-world outcomes, instead of chasing answers across disconnected tools.

## 2. Build trust through a phased agentic strategy

Treat an AI agent like a teammate in training. First, get clear on what it’s doing. If it’s taking action, build trust step by step. Start with diagnosis — finding root causes, drafting runbooks — while engineers stay in charge. Once it proves itself, let it handle bigger tasks like rolling back code. But if it’s interpreting data to connect the dots to business impact, the rules change. That job calls for strong context and careful validation, not more autonomy. One trust model won’t fit both.

## 3. Integrate telemetry into the strategic planning cycle

High-performing teams have moved beyond monitoring infrastructure to using telemetry as a primary input for product roadmaps and customer experience strategies. To replicate this, organizations should first identify “bridge metrics” — shared KPIs that translate technical health into business impact, such as the revenue cost of a checkout delay or the conversion impact of an AI agent’s hallucination. Then, integrate them into the standard business planning cycle by establishing a recurring review between engineering leaders and product stakeholders. This ensures that technical reality is always at the table when business goals are set.

## 4. Shift from cost management to strategic investment

To break the cycle of unpredictable billing, engineering leaders must stop managing observability as a utility expense and start governing it as a strategic investment in business intelligence. Implement a tiered data strategy that prioritizes high-value business signals over raw, unrefined volume. This shift requires partnering with vendors that offer transparent, predictable pricing models, leadership to set hard budgets rather than reacting to volatile usage spikes. By reframing the conversation from “how much are we spending on logs” to “how much business value are we generating from our telemetry,” engineering leaders can secure the funding needed for a robust practice while simultaneously eliminating the noise that inflates their bills.

# Industry highlights

## We identified key insights across four select industries worldwide.

### Financial services

Financial services organizations show some surprising gaps between expectations and reality. The most striking finding is how few have clear business outcomes tied to monitoring AI — just 24%, roughly half the 46% average — and 68% say they are unable to measure AI workload metrics (vs. 56% average). Fifty-eight percent say telemetry is not a key factor in their organization's strategic planning process, above the 46% average, and when correlating telemetry with business outcomes, 35% name data quality issues as the single most challenging obstacle (vs. 27% average). Perhaps most counterintuitive given the sector's regulatory environment, only 32% rate visibility into AI regulatory compliance as *moderately* or *very* important, well below the 52% average — suggesting that for financial services respondents, data quality hurdles may overshadow the compliance dimension one might expect to dominate.

### Telecom, media, and communications

Telecom, media, and communications organizations distinguish themselves most clearly through the speed of incident response. They are dramatically faster at spotting revenue impact: 48% can detect a meaningful drop in revenue within minutes (vs. 18% average) and 35% can define its cause in minutes (vs. 14% average) — a substantial edge that likely reflects the real-time, always-on nature of the sector's operations. It's possible that this speed is partly thanks to agentic AI, as 46% report that agentic AI has *significantly* improved their MTTR (vs. 42% average). That said, the sector is not without friction. More than half (55%) cite a lack of tooling capable of monitoring AI as their main agent observability challenge, above the 49% average, and 56% experienced surprise telemetry costs this year (vs. 50% average) — indicating that even as telecom organizations excel at rapid detection and response, they continue to grapple with AI monitoring capability gaps and cost unpredictability.

## Manufacturing

Manufacturing organizations stand out as one of the more mature sectors when it comes to operationalizing AI within their observability practice. Sixty-three percent report clear business outcomes tied to monitoring AI (vs. 46% average) and 55% say they can accurately control AI costs (vs. 45% average) — signaling that they're managing AI investments with measurable purpose. This maturity carries into agent observability, where 42% rate their practice as *very* effective at monitoring AI hallucinations (vs. 34% average), and into hands-on agent use, with 29% using AI agents to restart services *all the time* (vs. 21% average), reflecting greater trust in agents for execution-oriented tasks. Manufacturing also leads on cross-functional collaboration, as 59% *strongly* agree that monitoring AI workloads has driven closer work with security teams (vs. 51% average) — suggesting that AI adoption in this sector is not only well-governed but is actively strengthening the bridge between observability and security functions.

## Public sector

Public sector organizations face a distinctive set of challenges that set them apart from the broader cohort. Confidence in returns is more muted, with just 4% saying their observability ROI *significantly* exceeds expectations (vs. 8% average), and cost pressures are concentrated in data discipline — 44% name excessively verbose logs as the biggest driver of observability costs over the past year, well above the 31% average. Organizational barriers loom especially large: 31% cite organizational misalignment as the single most challenging obstacle to correlating telemetry with business outcomes, nearly double the 17% average, which may reflect the sector's complex governance structures. These frictions appear to slow incident response, as 24% say it takes weeks or longer to detect a meaningful drop in revenue (vs. 11% average). Notably, though, the public sector is more willing to lean on agents for high-stakes execution, with 24% using AI agents to roll back code *all the time* (vs. 17% average) — a pocket of advanced adoption within an otherwise constraint-heavy picture.

# Country highlights

## Snapshots from ten countries across the globe.

### Australia

Australian organizations present a more confident yet more exposed profile than the global cohort. They are more likely to report observability investments that *significantly* exceed ROI expectations (12% vs. 8%), pointing to strong satisfaction with returns. Yet this confidence sits alongside a lighter focus on AI governance: Australian respondents are less likely to call monitoring AI a priority (79% vs. 87%) and less concerned about the unintended consequences of AI agents (77% vs. 83%). Barriers to strategic value are also more pronounced, with 68% citing inaccessible or difficult-to-parse business data as a challenge (vs. 62% globally). Australian organizations are more likely to report a greater risk of downtime from cost-cutting initiatives (17% vs. 11%), suggesting budget pressures are translating into tangible operational exposure at a higher rate than the global average.

### Saudi Arabia

Saudi Arabian organizations stand out as notably more advanced and AI-forward than the global cohort. Nearly all respondents (98%) say AI has enabled them to focus more on innovation than maintenance, well above the 84% reported elsewhere. This maturity extends to visibility priorities: 90% rate having insight into model quality and drift as moderately or very important (vs. 76% globally), and 94% consider their observability practice *moderately* or *very* effective at monitoring AI cost (vs. 87%).

The most striking divergence is in cost sensitivity: 75% agree that unexpectedly high bills strain operations more than a high baseline cost, more than the 60% seen globally — yet they experienced surprise costs at roughly the same rate this year (52% vs. 50%).

### France

French organizations reveal a more strained relationship with AI cost governance and data management than the global cohort. They are more likely to say their organization cannot accurately control AI costs (62% vs. 55% globally) and to report a lack of clear business outcomes tied to monitoring AI (64% vs. 54%). France also diverges on data strategy: just 33% collect and store data from all infrastructure and applications centrally in one platform, below the 41% global figure, suggesting a more fragmented approach. These pressures appear to carry operational consequences, with 45% reporting they had to reduce the frequency of their data sampling due to rising telemetry costs, well above the 38% seen globally.

## Germany

German organizations show strong intent around AI monitoring but a persistent gap in connecting it to business strategy. While 91% say monitoring AI is a priority — slightly above the 87% global figure — only 31% have clear business outcomes tied to that monitoring, far below the 46% reported elsewhere. This gap extends to telemetry: just 44% say telemetry data is a key factor in their organization's strategic planning process, compared to 54% globally, suggesting operational data is not yet fully feeding higher-level decision-making. Where Germany leads is in privacy sensitivity, with 74% rating visibility into AI privacy as *very* important — their single most important factor and well above the 67% global average.

## India

Indian organizations stand out as advanced in their agent observability capabilities. They are markedly less likely to report being unable to measure AI workloads (46% vs. 56% globally). Their AI visibility priorities also run broader: 74% rate having insight into AI safety as moderately or very important (vs. 57% globally) and 70% say the same for AI bias (vs. 57%), signaling greater attention to dimensions that many organizations elsewhere deprioritize. This ambition is backed by demonstrated effectiveness, with 51% calling their practice *very* effective at monitoring AI model quality, well above the 35% global figure. Perhaps most telling, 73% say telemetry is a key factor in their organization's strategic planning process, compared to 54% globally — indicating that Indian organizations can elevate operational data into high-level decision-making.

## Japan

Japanese organizations present a more cost-pressured and capability-constrained profile than the global cohort, particularly where AI is concerned. Sixty-four percent say their total observability costs exceeded infrastructure and operational costs this year (vs. 57% globally), and they are nearly twice as likely to name AI infrastructure as the biggest driver of those costs (11% vs. 6%), suggesting AI workloads weigh more heavily on Japanese budgets. This pressure is exacerbated by lower monitoring effectiveness on certain AI issues — just 25% rate their practice as *very* effective at catching hallucinations (vs. 34% globally) and only 18% say the same for bias (vs. 22%). The operational payoff from AI also appears more muted, with 33% reporting that agentic AI significantly improved their MTTR, below the 42% seen globally — indicating that Japanese organizations are seeing comparatively less response-time benefit from agentic adoption even as AI-related costs and monitoring gaps grow more pronounced.

## New Zealand

New Zealand organizations reveal a distinctive pattern where confidence in handling AI coexists with notable capability and cost gaps. Their cost drivers diverge from the norm: trace volume is the single biggest driver of observability costs (30% vs. 21% globally), unlike the global cohort, where excessively verbose logs top the list. On the AI front, New Zealand respondents report strong detection gains, with 58% saying agentic AI *significantly* improved their MTTD, well above the 36% global figure. Yet this confidence is tempered by clear constraints. They are more likely to say they cannot accurately control AI costs (68% vs. 55%), and 70% cite a lack of tooling capable of monitoring AI as their top agent observability challenge, far above the 49% global average.

## Singapore

Singapore organizations show a nuanced profile where strong agentic AI gains sit alongside more tempered confidence in turning that adoption into clear outcomes. The clearest bright spot is response time: 51% say they have *significantly* improved MTTR with agentic AI, above the 42% global figure — though detection gains lag slightly, with 30% reporting significant MTTD improvement (vs. 36% globally). Elsewhere, the picture is more cautious. Just 39% say they have clear business outcomes tied to monitoring AI (vs. 46% globally), and only 47% are *moderately* or *very* confident in isolating the root cause of AI-related issues in their own environment, well below the 57% seen globally — pointing to a gap between AI adoption and the diagnostic assurance to support it.

## UK

UK organizations combine notable engagement with agentic AI for diagnostics with an acute awareness of its risks and costs. One area stands out clearly: 32% use AI agents *all the time* to diagnose root cause, well above the 20% global figure, pointing to meaningful operational reliance on agents for this specific task. Alongside this, concern about AI runs higher than average — 55% *strongly* worry about the unintended consequences of AI agents (vs. 44% globally) and just 18% *strongly* agree their organization can accurately attribute AI costs (vs. 23%).

Cost pressure is more pronounced overall, as 65% say their total observability costs exceeded infrastructure and operational costs this year (vs. 57% globally). Sixty percent *strongly* agree that correlating telemetry with business value metrics is a priority (vs. 53%), even as 68% cite inaccessible or difficult-to-parse business data as the biggest barrier to that correlation (vs. 62% globally) — underscoring a market that values telemetry's strategic potential but is still wrestling with the data accessibility needed to unlock it.

## United States

U.S. organizations are remarkable mainly for how closely they mirror the global cohort, with few meaningful outliers across the survey. The clearest point of distinction is in cost drivers: high-cardinality metrics are the top driver of observability costs in the U.S. (31%), whereas the global cohort points to excessively verbose logs as the leading driver. Elsewhere, U.S. figures sit just above the global baseline without diverging sharply: 60% say they are unable to measure AI workload metrics (vs. 56% globally), 27% use agentic AI *all the time* to improve alert configurations continuously (vs. 24%), and 60% say AI is improving their observability practice overall (vs. 59%). Taken together, the U.S. reads as a bellwether market — broadly representative of global trends.

# Methodology

Oxford Economics researchers surveyed 2,050 ITOps and engineering professionals from practitioners to executives (including developers, SREs, systems engineers, infrastructure operations professionals, CTOs, and CIOs) from February through March 2026. Respondents resided in Australia, France, Germany, India, Japan, New Zealand, Singapore, Saudia Arabia, the United Kingdom, and the United States.

Respondents represented 16 industries: business services, construction and engineering, consumer packaged goods, education, financial services, government (federal/national, state, and local), healthcare, life sciences, manufacturing, technology, media, oil/gas, retail/wholesale, telecom, transportation/logistics, and utilities.

The report also covers a select group of 206 respondents referred to as leaders who meet the following criteria. To qualify for this elite club, respondents must:

- Self-describe their organization as effective in fixing and preventing observability issues using agentic AI capabilities
- Self-describe their observability practice is "*very* effective" at monitoring model quality and drift
- *Somewhat* or *strongly* agree with the statement "My organization is able to link telemetry data to business value metrics."

# Strengthen your digital resilience in the agentic era




**The Hidden Costs of Downtime 2026: A $600 Billion Wake-Up Call**

Up 50% since 2024, the cost of downtime has reached a breaking point. Learn what's driving it and how to protect your business.

**Download the report**

**Perspectives by Splunk**

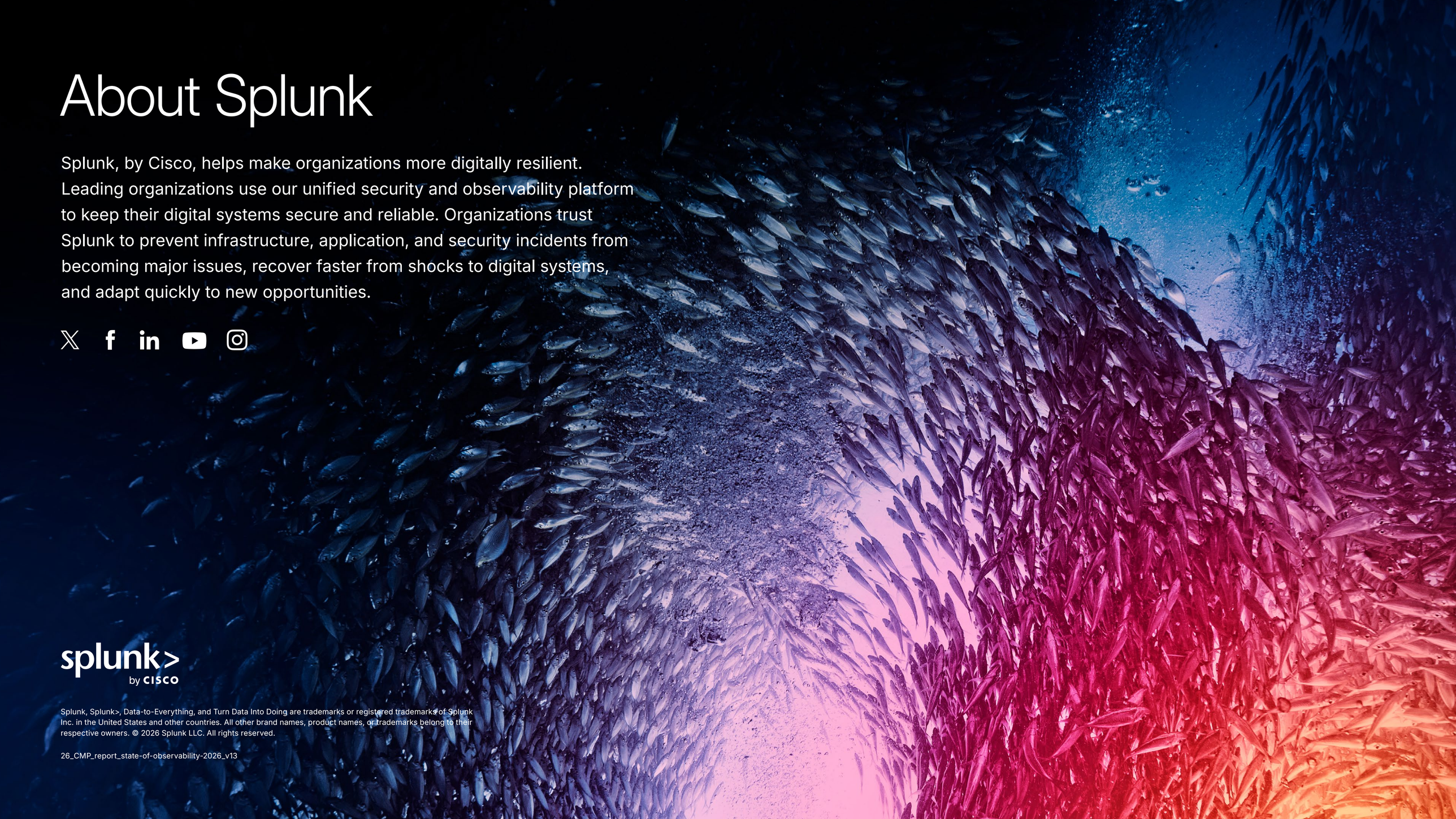
**— by leaders, for leaders**

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

**Get executive insights**

# About Splunk

Splunk, by Cisco, 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.



splunk> by cisco

Splunk, Splunk>, Data-to-Everything, and Turn Data Into Doing are trademarks or registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners. © 2026 Splunk LLC. All rights reserved.

26_CMP_report_state-of-observability-2026_v13