The New Rules of Data Management

Creating value in the AI era

splunk> a CISCO company

Contents

Finding value in the data clutter

You know the paradox: drowning in data but starving for insights. The adage was never more true than it is today.

The right data fuels insights that help organizations invent better customer experiences, identify malicious threats, and improve countless other processes to strengthen digital resilience. The plain fact, though, is that cloud services, connected devices, and AI are overwhelming organizations. And instead of thoughtfully arranging their data, they are stockpiling it like a garage cluttered with gardening tools, camping gear, and childhood memorabilia.

We wanted to know how organizations are cleaning out their data garages (so to speak), so we surveyed 1,475 IT, engineering, and cybersecurity professionals across the globe about their data management practices. We've based this report on our findings, revealing the best practices to ensure data is on hand when you need it, while creating more value.

Organizations have long followed the conventional wisdom of centralizing data into one place to unify visibility and better make sense of it. Although this practice offered organizations some control and visibility, data structures became more complex. Consequently, data management became more difficult, requiring strategies that went beyond simply centralizing data into one location. In an attempt to control costs and manage the explosion of data, organizations started expanding their storage locations with a medley of hybrid environments, opening the door for new sets of challenges.

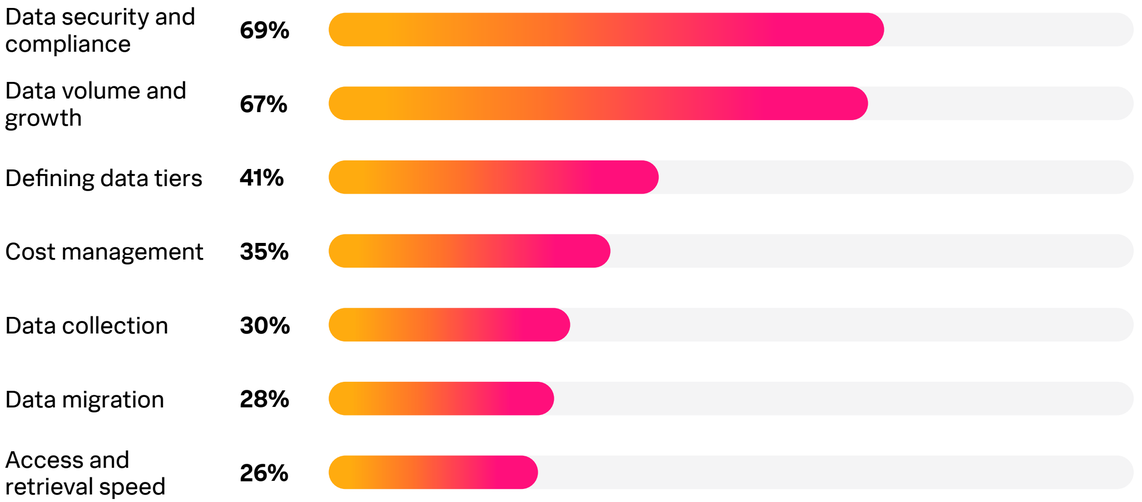
We think there's a better way. The new rules of data management can help you realize your security and observability objectives and advance your mission, while you also optimize costs and compliance. Keep reading to see what data management leaders do differently. Discover how to tamp down data complexity and maximize its value in the AI era.

Data at the crossroads

The survey confirms what many organizations may have suspected for years — the exponential rise of data is giving way to increased complexity that makes it more difficult to access, analyze, and secure data, as well as comply with regulatory mandates. This is why having a sound and comprehensive data management strategy is crucial for digital resilience.

But here again, volume and too many siloed data stores get in the way. In fact, 67% of survey respondents cite data volumes and growth as a challenge when implementing their data strategy, surpassed only by 69% who call maintaining data security and compliance a top data management obstacle. They agree that defining data tiers, cost management, and other activities were also obstacles.

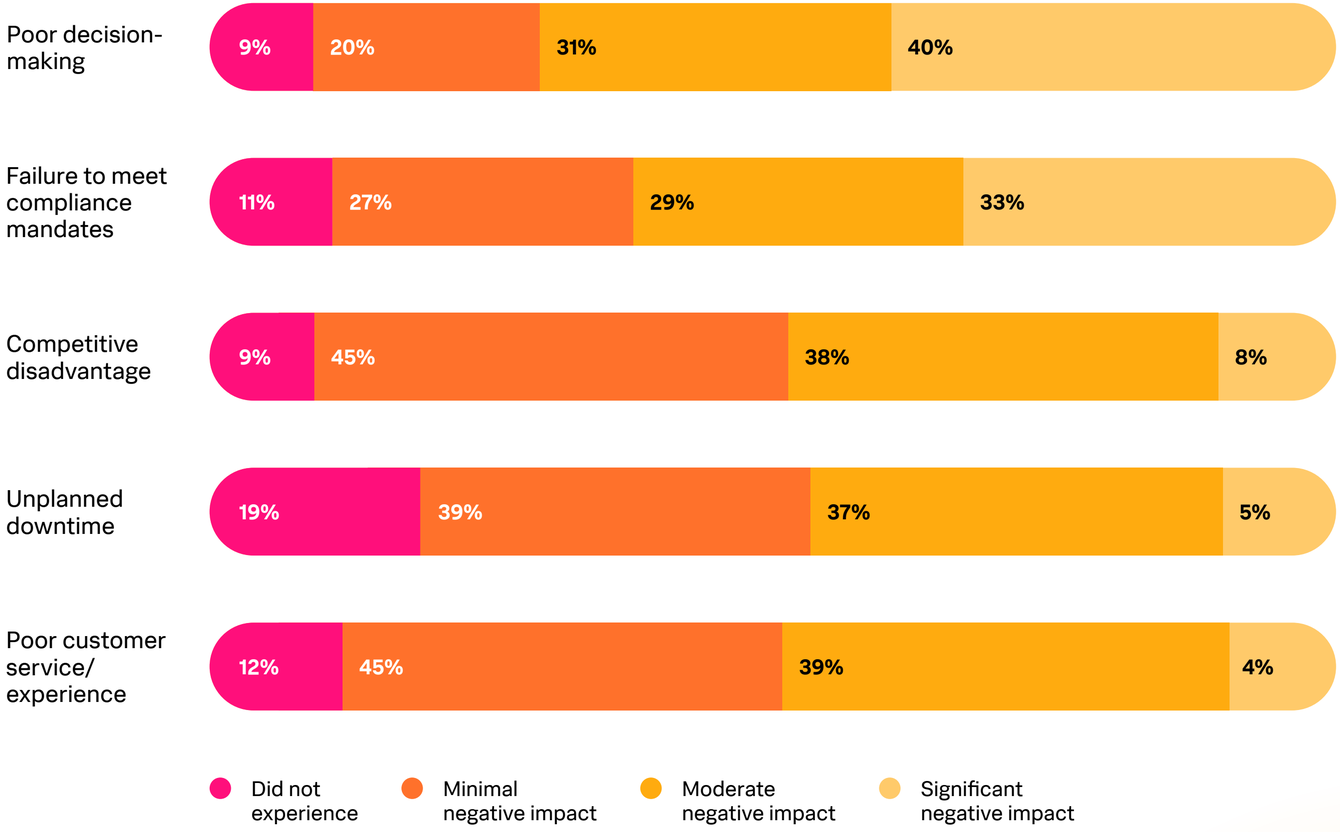
What's standing in the way of your data management strategy?



Organizations wrestling with these data management issues are also feeling far-reaching business impacts. Sixty-two percent of respondents claim that difficulties with data management resulted in compliance failures (33% *significant impact*, 29% *moderate impact*), 71% say they led to poor decision-making (40% *significant impact*, 31% *moderate impact*), and 46% confirm they led to competitive disadvantages (8% *significant impact*, 38% *moderate impact*).

Data redundancy is also a serious dilemma for organizations trying to stay afloat in a tsunami of data. Fifty-nine percent of respondents reveal their current data management strategy has *somewhat* worsened the rate of data duplication, and 20% say the problem is *significantly* worse.

The real world consequences of data management challenges



Breaking down the cost of data management

Data management costs are on the rise for almost everyone. Ninety-one percent of respondents reveal they spent more on data management this year than in the previous year.

Respondents call out volume and compliance again, this time for driving increased costs — nearly three-quarters (73%) label data volume as a primary cause, and shifting compliance regulations came in second at 71%. The latter reflects a groundswell of more expansive and rigorous compliance mandates, requiring organizations to understand exactly where and how data is stored and protected across their ecosystem, and more importantly, who has access to data.

Complying with current regulations such as FedRAMP, ISO27001, PCI, and HIPAA (just to name a few) require more financial investment now because more is at stake. Organizations risk potential fines and importantly, collateral damage to their reputations and customers' trust if they fail to comply. And even long-standing regulations have become more expansive and demanding. The European Union's General Data Protection Regulation (GDPR), for example, requires comprehensive visibility across an organization's entire data and customer environment, and is likely one of the costliest regulations to support from a data management perspective.

When evaluating budget allocation relative to the data lifecycle, respondents report spending 6% less on storage and 7% less on indexing on average. In light of steadily rising data costs, organizations have sought out less expensive options for data storage.

While convenient, without the right overarching strategy and controls, these distributed storage methods spread across multiple clouds, data lakes, and other storage locations risk duplication, redundancy, and governance issues. What's more, adopting multiple storage options can create unintended complexity that may thwart efforts to understand and streamline costs — more than a quarter (26%) of respondents maintain they aren't able to accurately calculate the ROI of their data management investment.

On average, respondents admit they spent more than a quarter (28%) of their data management budget this year on search and analysis, up slightly from 24% spent the previous year. This suggests a keen interest in not only reducing the noise in data, but also mining its value.

91%

report their overall spend on data management has increased compared to the previous year

The top drivers of increased data management costs

73%
Increasing data volumes

71%
Shifting regulations

62%
Data management technology

53%
Data security needs

14%
Sustainability efforts

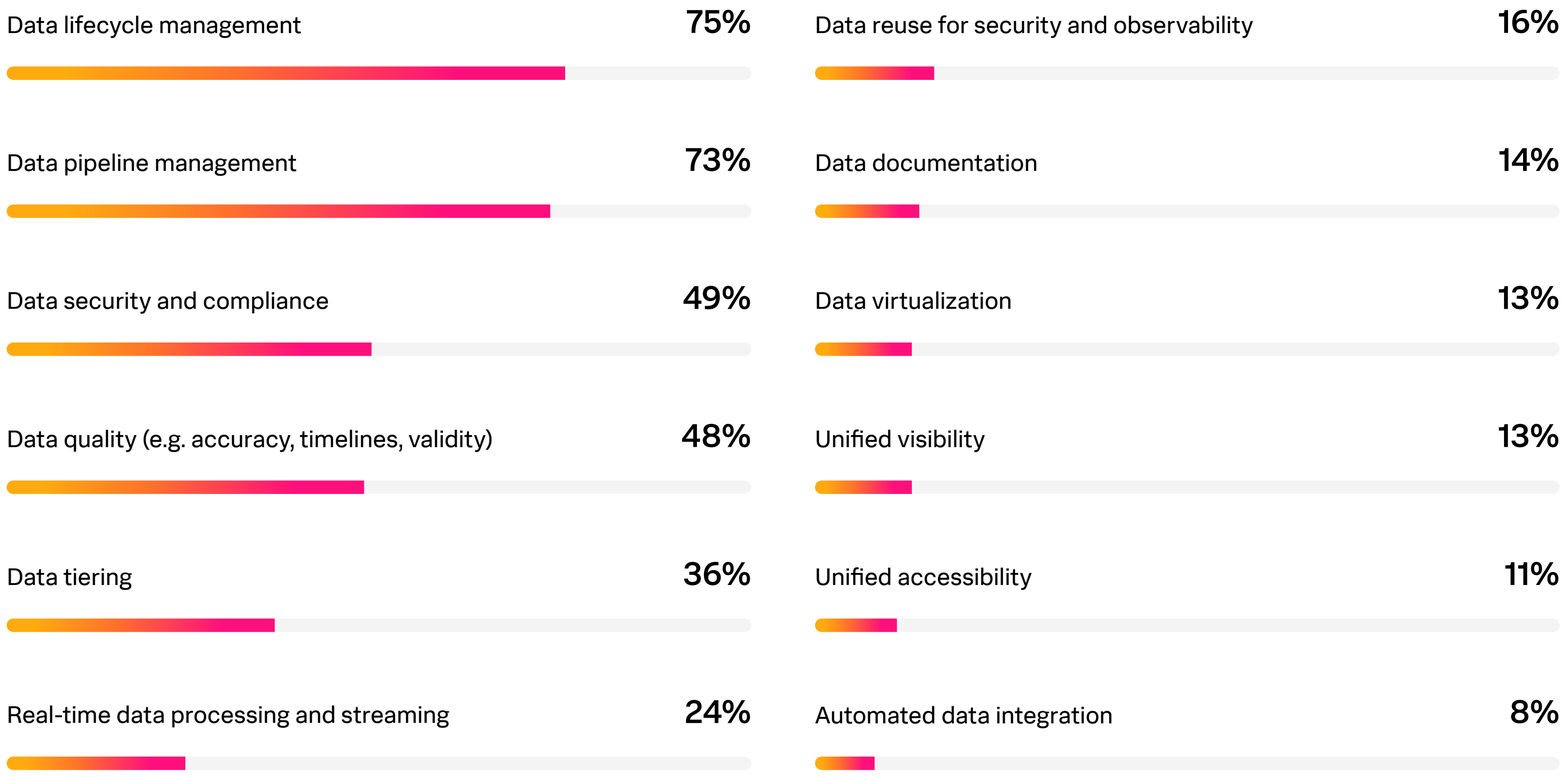
Data management practices have fallen behind

A data management strategy is a set of practices to help organizations tame data complexity and manage its lifecycle. However, many organizations haven't evolved their data management practices in line with data growth and complexity. The old way of data management either requires your data to be consolidated into one location at significant cost, or that you live with data silos and sacrifice visibility. As a result, organizations are compelled to migrate data frequently and struggle with privacy across their environments.

Data access is a prevailing issue. Fifty-three percent admit they have to log into different platforms to access different data sources. And few respondents say their data management strategy includes components such as unified visibility (13%) and unified accessibility (11%). One reason for this disparity could be the need to break down organizational silos across multiple systems and teams that makes achieving unified visibility and unified accessibility difficult to realize.

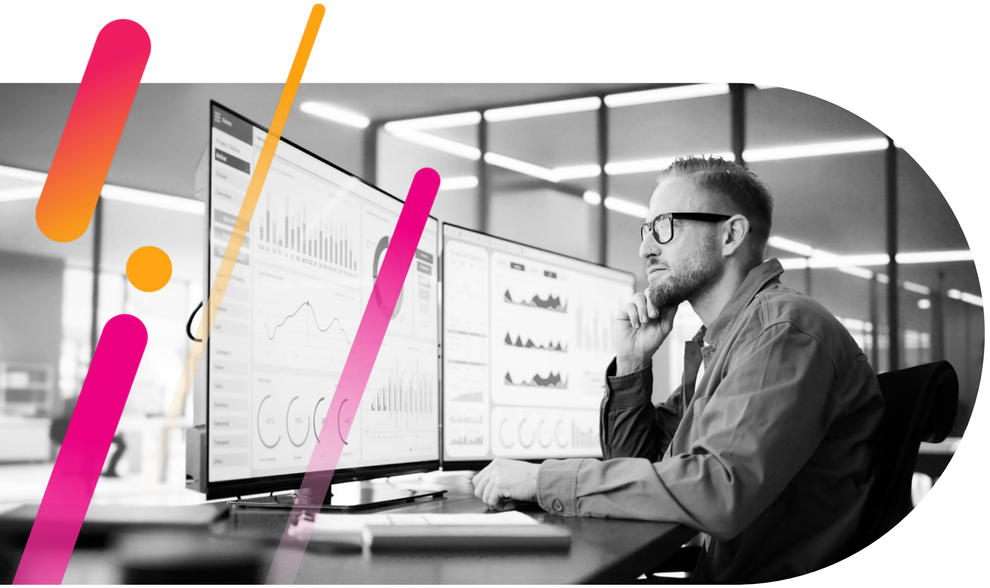
Many organizations are still moving data from disparate sources to consolidate their environments and gain visibility. Forty-seven percent move data monthly, and most say they have been migrating more data to cloud infrastructure (76%) over the last two years. But moving data monthly carries risk, opening the door to potential security breaches, data leaks, and compliance violations if not properly handled. Depending on the size and complexity of the data being moved, costs can quickly add up and take a toll on an organization's bottom line.

The many practices that make up a data management strategy



And while organizations may have a data management strategy in place, many struggle with fundamental governance or enforcement. A considerable portion of respondents reveal the following data management policies are not well-enforced: role-based data access (57%), instructions about where data types should be stored (57%), and defined data retention periods (44%). These loopholes can jeopardize compliance standing, potentially resulting in hefty fines, legal repercussions, loss of brand and reputation, and diminished customer trust, among other ramifications.

Additionally, 79% of respondents don't have a policy governing data destruction (33% have no plans to create one) — further muddying the waters. Organizations are already overwhelmed by a complex and noisy data landscape. Inconsistent and outmoded data management practices only add to their struggles.



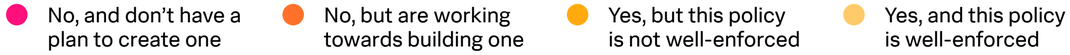
Organizations waffle on enforcing policies

A clear policy that instructs where specific data types should be stored

A clear data access policy based on employee role

A defined retention period for how long data should be stored in each location

A clear policy that governs the destruction of data



The new rules of data management

Organizations with a forward-thinking data management strategy are primed for digital resilience, as they can access and process data faster, and have higher-quality data to surface insights and produce more reliable outcomes. They have adopted practices that provide a clear upside for data management.

According to the survey, the two practices accounting for the lion's share of organizations' data management strategies are data lifecycle management (75%) and data pipeline management (73%). (We'll come back to these foundational practices in the next chapter.)

Looking more closely at the data reveals other practices that, while less utilized today, move respondents a step closer to value creation. Data quality, data reuse, data tiering, and data federation all help organizations access, see, and understand their most critical information. In short, these practices help organizations know what data is being generated in their enterprise and allow them to access it cost effectively, regardless of where it resides.

Data quality

Data quality is a significant part of a data management strategy for 48% of respondents, who claim they've experienced a myriad of improvements compared to those who don't emphasize it. For example, 73% of organizations that make data quality a priority (vs. 51% of all other respondents) say mean time to respond (MTTR) has improved. They're also more likely to successfully neutralize threats (54% vs. 41% all other respondents), identify root causes (45% vs. 34% all other respondents), and improve threat detection capabilities (61% vs 37%).

Data reuse

Each data source can serve multiple purposes, but factors like data accessibility and proprietary formatting often drive data duplication and blind spots.

Data reuse might not be as pervasive as other practices — only about 16% of respondents say data reuse for security and observability comprises their data management strategy. But if anything, this indicates more opportunities to save costs by avoiding redundant data collection, enhancing collaboration, engaging in data stewardship across teams, and generating new insights by combining datasets from different sources.

Organizations that include data reuse in their data management strategy say it has generated notable value. Among other benefits, they are less likely to face hurdles when handling high volumes of data (46% vs. 71% all other respondents). They're also more likely to reduce the impact of incidents (52% vs. 35% all other respondents), and experience fewer data breaches (44% vs. 33% all other respondents). Organizations that reuse their data also see better threat detection performance (62% vs. 47% all other respondents).

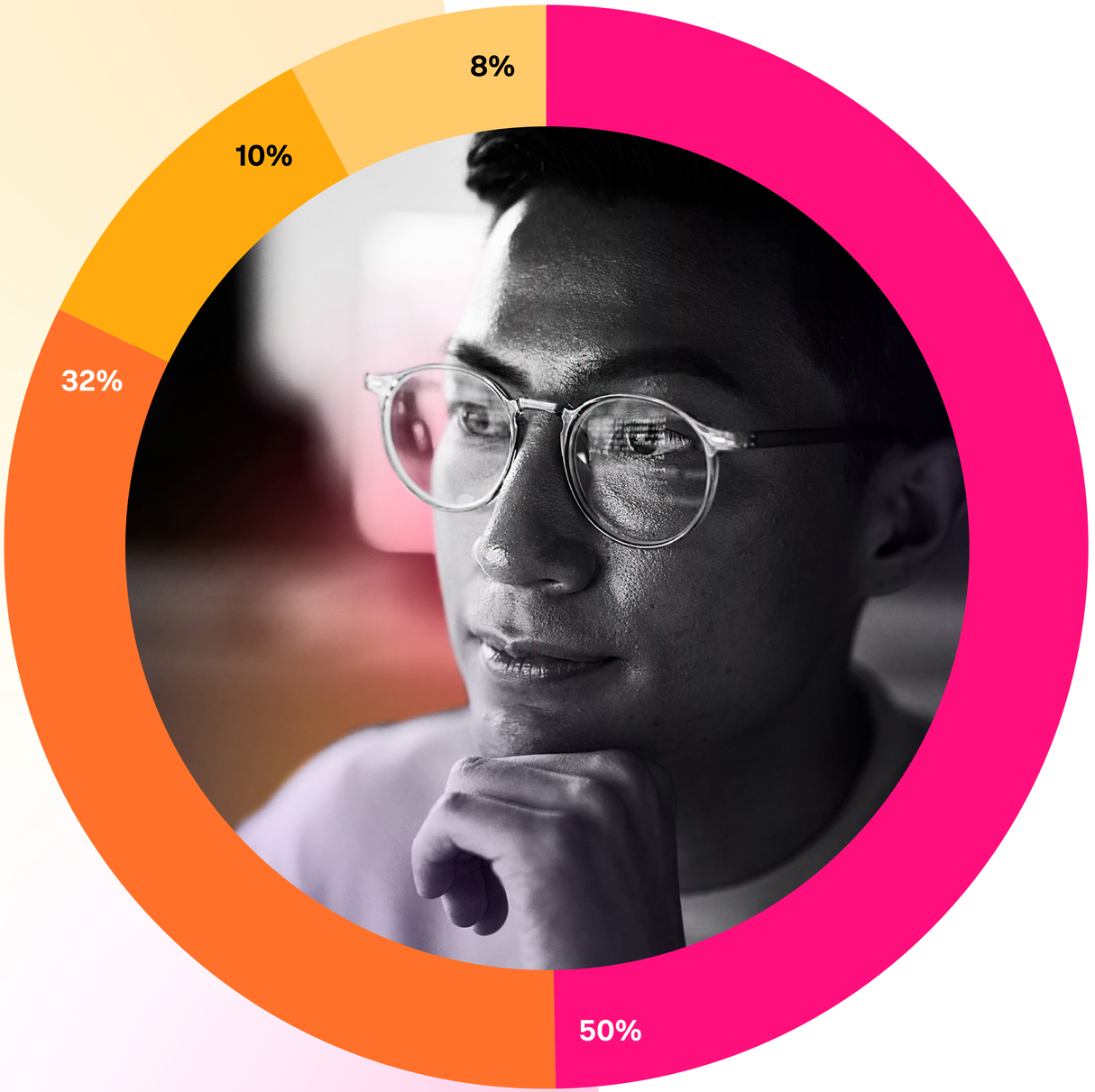
Data tiering

Data tiering prioritizes data based on factors such as access frequency, age of the data, and usage patterns. According to the survey, 36% of organizations employ this practice as a part of their data management strategy, saying they aim to reduce data storage costs and accelerate access times for commonly used data types.

Data tiering enhances access times and security while reducing costs

Benefits ranked number one by respondents

- Reduced storage costs
- Accelerated access times for commonly used data types
- Increased security for older data types
- Increased data analysis productivity



Respondents who have implemented data tiering experience many benefits, with 50% ranking reduced storage costs as the number one positive impact, followed by accelerated access for commonly used data types (32%), increased security for older data types (10%), and increased data analysis productivity (8%).

Organizations that tier data are also less likely to encounter challenges with access and retrieval speed (18% vs. 31% all other respondents), cost management (18% vs. 44% all other respondents), and data migration (18% vs. 34% all other respondents).

Data federation

Organizations that employ federation can access and analyze data from multiple, disparate data sources and locations as though it were a single dataset and without moving the data. However, few have mastered the art. While 92% confirm having some form of a federated practice, only 20% claim it's fully implemented.

With so many storage locations, access methods, analytics platforms, and data workflows to navigate, a federated data management strategy makes a lot of sense. Organizations that have adopted data federation, whether fully or partially, reveal a slew of benefits, including faster data access (67%), improved data governance (54%), and improved compliance posture (47%).

The survey shows that a federated data management strategy provides organizations enormous advantages across security, observability, AI, and other critical areas. Ultimately, if organizations aim to maximize the value of their data, adding federation to complement their current data management strategy will be key.

What data federation can do for you

67%
Faster data access

54%
Improved data governance

47%
Improved compliance posture

37%
Reduced data redundancy

36%
Cost savings

21%
Minimized data movement

Leaders play by the new rules

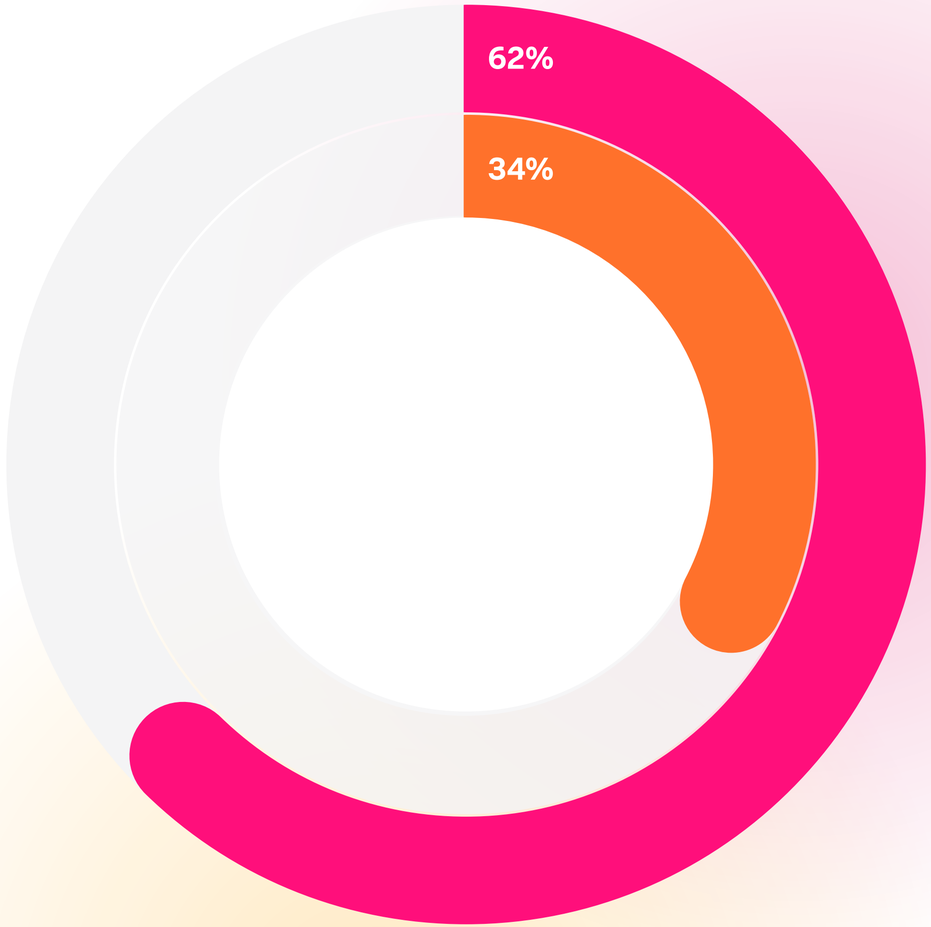
When developing a winning data management strategy, the survey indicates organizations that have adopted a trifecta of practices — fully implemented data federation, data pipeline management, and data lifecycle management — are often ahead of their peers. These data management leaders not only make strategic data management investments, they also realize a host of business benefits.

The leader cohort reports greater business performance improvement over the last two years in several key areas compared to all other respondents, including net operating profit margin (69% vs. 56%), sustainability (58% vs. 38%), and speed of innovation (55% vs. 44%).

Data management leaders are also more likely to state their data management strategy has enhanced other key data-related metrics, such as speed to access (79% vs. 73% of all other respondents), speed of overall data processing (76% vs. 69% of all other respondents) and their amount of computational overhead (62% vs. 45% all other respondents). The leader cohort also sees other valuable benefits from their data management strategy, most significantly cost savings (62% vs 34% all other respondents).

Data management leaders slash costs

Percentage of respondents who reported cost savings



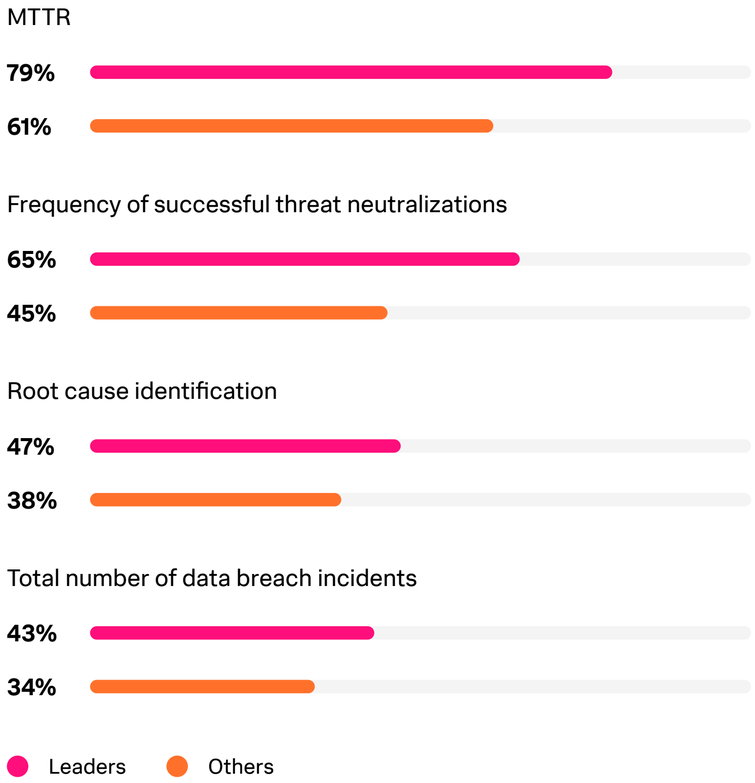
Modern data management strengthens cybersecurity

A modern data management strategy does more than wrangle and organize data; it also has a measurable impact that boosts other security outcomes. Leaders report *significant* improvement in all aspects of TDIR — threat detection (26% vs. 12% all other respondents), investigations (22% vs. 9% all other respondents), and response (33% vs. 20% all other respondents).

Data complexity can expand an organization's attack surface, providing more opportunities for threat actors to engage in nefarious activities. Left unchecked, this complexity can impede business success. However, a winning data management strategy helps organizations align their data with security goals. Data management leaders report faster mean time to respond (MTTR) (79% vs. 61% all other respondents), more successful threat neutralizations (65% vs. 45% all other respondents), quicker root cause identification (47% vs. 38% all other respondents), and fewer breaches (43% vs. 34% all other respondents).

Data management leaders elevate security posture

Areas that have *slightly* or *significantly* improved



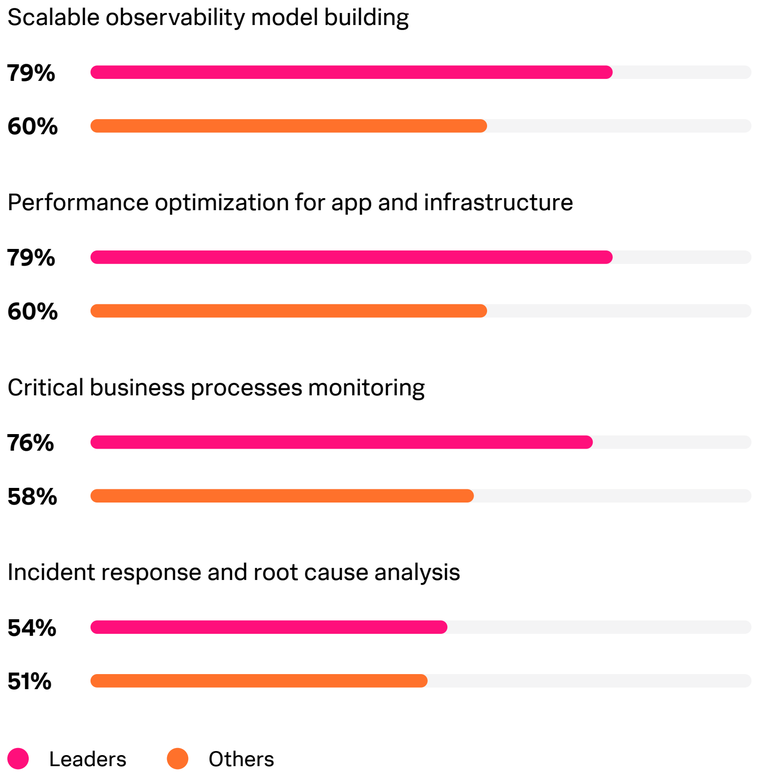
Robust data management boosts observability, ITOps practices

A data management strategy composed of fully-implemented data federation, data pipeline management, and data lifecycle management has similarly rewarding outcomes in ITOps and observability practices. In observability, data management leaders experience substantial gains in scalable observability model building (79% vs. 60% all other respondents). Leaders also confirm their data strategy improved performance optimization for app infrastructure (79% vs. 60% all other respondents), as well as critical business process monitoring (76% vs. 58% all other respondents).

As in security, the winning trifecta of data management practices improves IT metrics for leaders. Leaders see significant gains in KPIs such as mean time to resolve (MTTR) incidents (78% vs 58% all other respondents) and log volume and pattern optimization (56% vs 38% all other respondents).

Data leaders boost observability outcomes

Areas that have *slightly* or *significantly* improved



CHAPTER 5

The symbiotic relationship of data management and AI

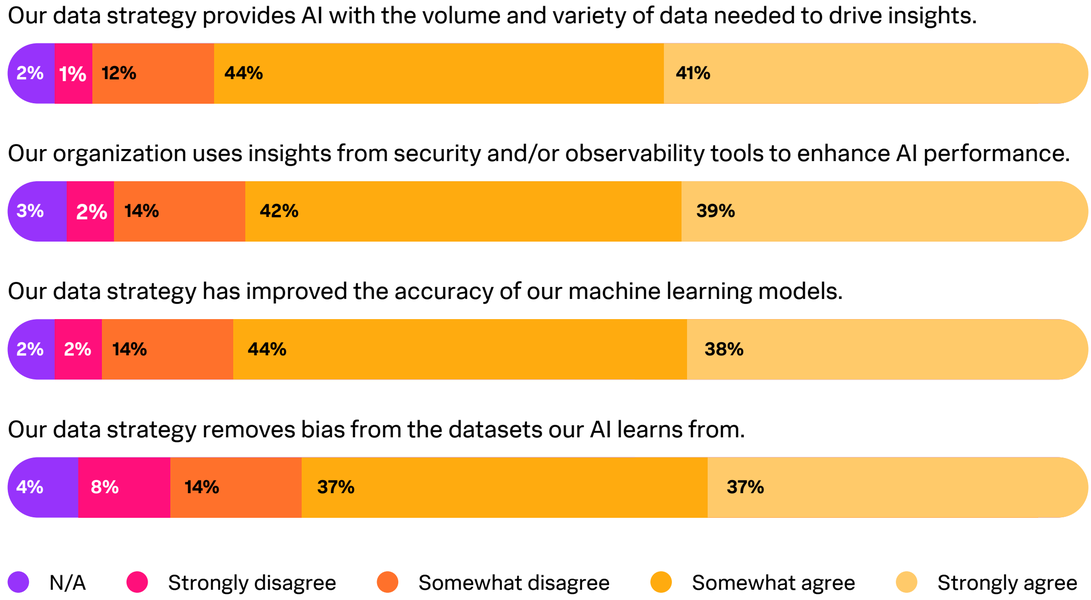
The survey suggests the relationship between data management and AI is mutually beneficial. AI depends on quality data, so a strong data management strategy plays a vital role in how AI models perform. The inverse is also true — AI helps fill in the gaps of organizations' data management practices by boosting productivity and automation when woven into workflows.

A strong data management strategy will be a force multiplier for AI implementation. Across the board, survey respondents hail the benefits of their data management strategy on AI, with 85% saying it provides AI with enough data volume and variety to generate valuable insights (41% *strongly agree*, 44% *somewhat agree*). Additionally, 74% report their data management strategy removes bias from the datasets from which AI models learn (37% *strongly agree*, 37% *somewhat agree*). And 82% say their organization's data strategy has improved the accuracy of their machine learning models (38% *strongly agree*, 44% *somewhat agree*) — all of which lay the groundwork for competitive advantage as they build out their AI implementations in a crowded market.

What's more, 81% of organizations also say they leverage insights from security and observability tools to enhance AI model training and performance (39% *strongly agree*, 42% *somewhat agree*).

While AI offers many advantages for data management — and vice versa — AI also introduces new obstacles that continue to be a source of frustration. Survey respondents cite that AI has made data integration harder and contributed to the existing challenge of high data volumes.

AI success starts with data management



Yet, despite these obstacles, AI has a largely positive effect on data management. Virtually all respondents (98%) agree that AI made their data management strategy easier (33% say *significantly* easier). AI delivered the most value by performing routine, administrative functions — 73% of respondents believe AI enhanced data quality by automating repetitive tasks. AI also opened up new opportunities. Fifty-nine percent state it helped with data discovery, including scanning large datasets to identify patterns, trends, and anomalies.

The value of data management and AI are interwoven, with each enhancing the effectiveness of the other. The quality and accuracy of your AI is directly related to the data it has access to and the quality of that data. Conversely, AI enhances data management by automating processes, improving security, and optimizing storage, creating a cycle of continuous enhancement. Organizations that leverage both effectively realize a significant competitive advantage in data-driven decision-making.

How AI is transforming data management for the better

73%

Enhance data quality via automating repetitive tasks

59%

Easier data discovery to identify patterns, trends, and anomalies

Getting your data house in order

Like any spring cleaning project to reorganize your drawers, closets, and garage, restructuring your approach to managing data is an opportunity to reset. It helps you not only declutter, but also make room for new possibilities. But first, you'll need to start with the basics: Know what data your organization generates and prioritize business goals and use cases.

Here are a few recommendations to help you maximize your data's value from the ground up.

1. Know your data and classify it

To lay the foundation of data management, you must first understand the data being generated in your organization. Then, define your target use cases according to how data will be used (real-time detection vs. historical investigation), relative to the business constraints (retention requirements, for example). From there, you can then identify which data management practices can help you meet those needs. Classifying your data will also require a strong data governance policy, along with data retention and role-based access. So make sure you're providing regular policy training to your teams so they understand where the data lives and how it can and should be used.

2. Keep your data clean

Quality matters. That holds especially true for your data. However, only half of survey respondents prioritize data quality as a core component of their data management strategy. Even if your data is federated, accessible, and indexed, your data management strategy can still fail. Why? Because you don't have the right data powering your systems, processes, platforms, and applications. Having the right data is an iterative process that starts from the moment it is generated. It should be fundamentally accurate, complete, and formatted to meet needs as they arise, ensuring it's optimized to create value. Prioritizing quality data will be especially important when you start implementing AI. (Remember the old adage, "Garbage in, garbage out?") Good, clean data will help your AI models perform more accurately — and give you better outcomes.

3. Access your data without moving it

We get it, you need to have a single source of truth, and that means ensuring all your data is accessible. That's where a data federation practice provides the most value, offering unimpeded access to all of your data, regardless of where it lives — and without costly migrations. The ability to access your data at rest is critical. It's especially important when accessing data stored in diverse locations, necessary for making informed business decisions. For example, when a user requires additional information during the threat hunting process, they need the ability to run ad-hoc searches against the external data store where that data resides to gather insights and make the right decisions. Data federation enables you to easily reach for specific data related to an incident, allowing you to make accurate, and better informed decisions about your current environment, and how to keep your systems protected in the future.

4. Take a platform approach to your data

While you might be able to query your data from a number of separate tools, you will still need to unify your data so you can clearly see the entire picture. Implementing a unified data platform — one fully equipped with federation capabilities that deliver unified accessibility without having to move data at all or log into different platforms — will not only bring your data into full view, but also make it easy to locate and use for any use case, without breaking the bank. In addition to a holistic view of your data, a unified data platform will also help pare down multiple or redundant tools, streamline workflows, reduce integration headaches from multiple vendors, and ease "swivel chair syndrome" and other issues. Whether you're leveraging data for security or observability, or both (think data reuse), or using it to drive AI, a data platform that enables pipeline management, analytics, and federated access helps you serve the right data to the right teams.

Redefine your data management strategy with Splunk

The CISO Report

Discover how CISOs and their boards are bridging critical gaps on top priorities, including collaboration, data reuse, compliance approaches, and success metrics.

Download the report

Perspectives by Splunk — by leaders, for leaders

Find out how executives and business leaders are rethinking their data management strategy to address industry challenges and realize new opportunities across security, observability, and AI.

Get executive insights

Methodology

Oxford Economics researchers surveyed 1,475 IT, engineering, and cybersecurity professionals from November 2024 through January 2025. Respondents were in Australia, France, Germany, India, Japan, New Zealand, Singapore, United Kingdom, and United States. They also 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. Respondents defined as “data management leaders” consist of organizations that have applied fully implemented data federation, data pipeline management, and data lifecycle management.

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.

Keep the conversation going with Splunk.

splunk> a CISCO company

25_CMP_report_the-new-rules-of-data-management_v11