White Paper, sponsored by:
Western Digital Corporation

Built for Scale: The Enduring Role of HDDs in the AI Era

Written by:
Ed Burns
Research Director, Semiconductors and Enabling Technologies

Adam Wright
Senior Research Manager, Global DataSphere and Global StorageSphere

Published:
September 2026

Table of contents

In this white paper

AI is rewriting the rules of enterprise data infrastructure, and not just in how data is processed, it is also changing how much data must be stored, for how long, and at what cost. AI is also fundamentally changing the economics of storage. As AI workloads both require and generate unprecedented volumes of data, enterprise infrastructure is increasingly being designed as a long-lived data system, one in which long-term retention, selective rapid retrieval, and cost-effective scale are core architectural requirements. Hard disk drives (HDDs) remain not just relevant but foundational in this environment, serving as the indispensable high-capacity, low-cost storage layer that makes AI infrastructure economics viable at scale. Flash storage, including enterprise TLC and QLC SSDs, will continue to play a critical role where performance-sensitive, latency-dependent workloads demand speed that HDDs cannot match, making the two technologies complementary rather than competitive in a well-architected AI storage stack. Without high-capacity storage as the dominant capacity tier, the data retention, training data set accumulation, and archive economics underlying AI would be prohibitively expensive or technically constrained. This is a reminder that AI has shifted storage from a capacity management problem to an infrastructure economics problem.

"Storage is still one big problem that we need to resolve in order to get AI moving efficiently."

— IT decision-maker, distribution and systems integration

To examine how these dynamics are playing out across the enterprise, WD commissioned IDC to conduct a primary research study encompassing a quantitative survey of 763 IT and business decision-makers at manager level and above across seven countries, all with direct responsibility for AI infrastructure or data storage decisions. IDC also conducted in-depth qualitative interviews with two senior storage channel leaders, providing practitioner perspective on deployment realities and evolving customer priorities. What the data reveals is a compelling and consistent case — across workload types, lifecycle phases, and economic dimensions — for HDDs as the indispensable foundation of AI infrastructure at scale.

Executive summary

AI is driving a structural, long-term expansion of enterprise storage demand that favors capacity, economics, and reliability over raw performance. The economics of storage are being rewritten in the process. Across data growth, retention behavior, workload priorities, and infrastructure deployment, the findings validate the central hypothesis of this paper: that HDDs are not a legacy technology navigating disruption, rather they are the indispensable foundation of AI infrastructure at scale. Key findings are mentioned in the sections that follow.

AI is creating explosive data growth

- 94.7% of organizations report data volume grew as a result of AI/GenAI over the past 12 months, with 60.9% reporting growth of 25%+.

- 73.8% expect data volume growth of 25%+ over the next 3 years — with the share expecting 50%+ growth nearly doubling from 16.1% today to 30.5% forward looking.
- 85.4% report data lake volume has grown over the past 12 months, and 85.8% expect it to grow further over the next 3–5 years.
- AI-generated data — synthetic outputs, model logs, and inference results — is the number 1 cited driver of data lake volume growth at 59.4%, ahead of training data set expansion at 50.1%.

Data is becoming more valuable and retained longer

- 94.7% say the value of their organization's data has increased as a result of AI/GenAI adoption.
- 74.3% say AI/GenAI has caused them to retain data longer.
- 75.9% report increasing the volume of archived cold-tier data being brought back online to support AI workloads.
- 96.0% anticipate needing faster archive retrieval to support AI inference or retrieval-augmented generation (RAG) pipelines.

AI changes storage economics and HDD benefits

- 98.2% rate TCO per terabyte as an important or very important factor in storage decisions.

Introduction

The world is generating and storing data at a scale that would have been unimaginable just a decade ago, and the pace is accelerating. IDC's Global DataSphere projects that 274.4ZB of data will be created worldwide in 2026 alone, rising to more than 718.4ZB annually by 2030. IDC's Global StorageSphere shows that the installed base of total storage capacity worldwide stands at 13.1ZB in 2026 and is expected to more than double to over 26ZB by 2030. The gap between data generated and data stored reflects the reality that not all data warrants retention, but the data that is kept, and the infrastructure required to store it reliably, economically, and at scale, has never mattered more. AI is fundamentally changing the economics of storage, and for enterprise infrastructure leaders, getting that economics right has become a strategic imperative.

> "For every compute component that you implement in a server … you will need probably at least two to three times more storage than what they used to have because data is currently being generated at a very fast pace."
>
> — IT decision-maker, distribution and systems integration

AI has emerged as a major accelerator of this growth. AI workloads are not only consumers of data but generators of it, and while much of what they produce is ephemeral, a strategically important subset is retained for reuse, retraining, compliance, or future analytical value. It is this retained data that is reshaping enterprise storage strategy. While compute investment remains significant, AI deployments are increasingly constrained not only by compute but by storage architecture. Storage infrastructure has emerged as an increasingly critical scaling challenge for AI environments, and study results confirm that this shift is already well underway: more than 93% of organizations report storing more data than before as a direct result of AI adoption, and more than 85% report that storage capacity specifically supporting AI and machine learning workloads has increased over the past two years (see Figures 1 and 2).

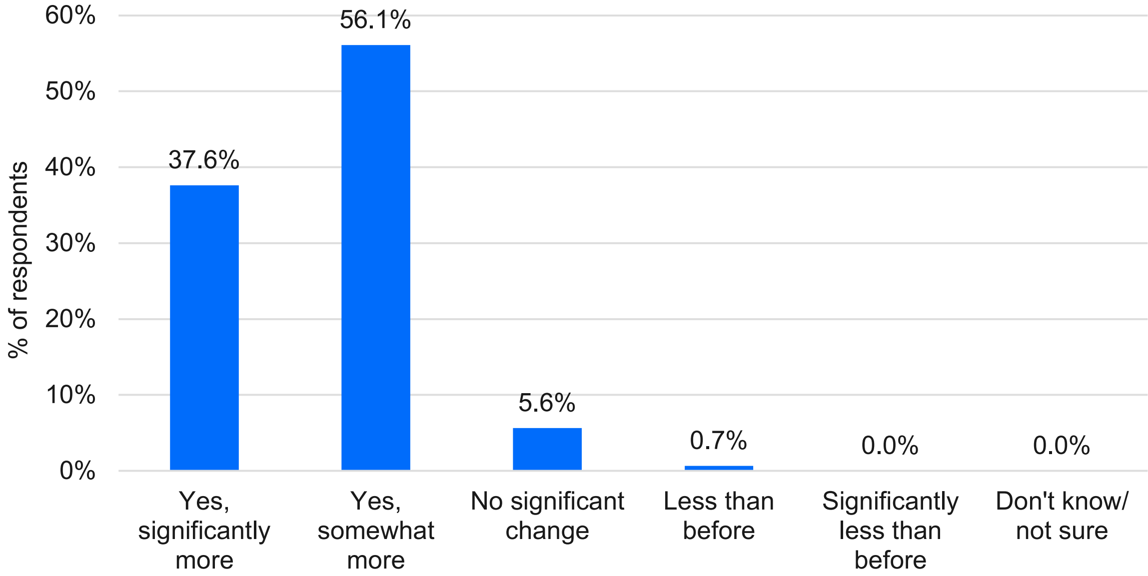
> "I think actually storage is the bigger challenge in a lot of cases versus the CPUs."
>
> — Senior product manager, hyperscale cloud service provider

This dynamic is playing out across industries and geographies. Organizations are rediscovering the value of data they once considered dormant. Surveillance footage is being reprocessed through computer vision models. Telecommunications networks are mining historical records to train customer behavior models. Public sector agencies are digitizing decades of physical archives to feed government AI initiatives. As one study respondent noted, hospitals are actively consolidating patient records and imaging archives into centralized AI-ready repositories, driven in some cases by government mandate. The common thread is not the AI model itself, but the data that powers it and the infrastructure that makes that data accessible and economically viable to retain over time.

FIGURE 1

Organizations are storing more data since AI adoption

Q. As a result of AI/GenAI adoption, is your organization storing more data than before?



n = 763 (unweighted base)

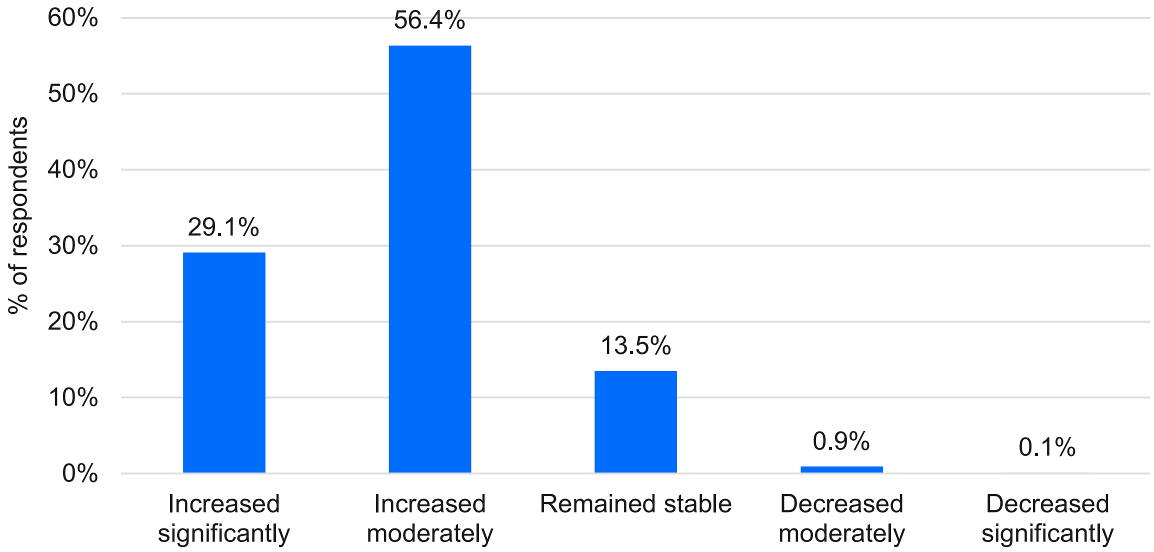
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

No single storage medium serves the full spectrum of AI workload requirements equally well. What the evidence increasingly shows is that hard disk drives occupy an indispensable and expanding role across the majority of the AI data life cycle, not as a legacy technology holding on but as the economic and architectural backbone that makes AI infrastructure viable at scale. Flash storage, including enterprise TLC and QLC SSDs, will continue to play a critical role where performance-sensitive, latency-dependent workloads demand speed that HDDs cannot match. But for the vast majority of AI data, the governing requirements are capacity, cost per terabyte, reliability, and sustained throughput. These are precisely the dimensions on which HDDs excel, and precisely the dimensions that the AI era is amplifying. The sections that follow examine each of these dimensions in turn, drawing on survey data and practitioner insights to build the case for HDDs as the active, load-bearing storage foundation the AI era demands.

FIGURE 2

Storage capacity growth for AI training

Q. Over the past two years, how has your organization's total storage capacity changed specifically to support AI/ML model training workloads?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

Data growth and evolving customer needs

AI-driven data growth is not linear, it is compounding. Organizations are generating more data, retaining more of it, assigning it higher strategic value, and discovering new value in data they previously would have discarded. New data types like synthetic data sets, AI-generated content, model checkpoints, and logs are emerging as significant storage drivers. The result is a structural, sustained expansion of enterprise storage requirements that is insulated from the cyclicality of compute demand, and one that is increasingly governed by infrastructure economics, not just capacity planning.

The scale of this growth is already visible in the data. Nearly all respondents report that their organization's total data volume has grown over the past 12 months as a direct result of AI and GenAI workloads, with more than 60% reporting growth of 25%+ (see Figure 3). What makes this particularly significant is the forward trajectory: looking ahead three years, the share of organizations expecting growth of 50%+ nearly doubles. The growth challenge is not stabilizing. It is accelerating (see Figure 4).

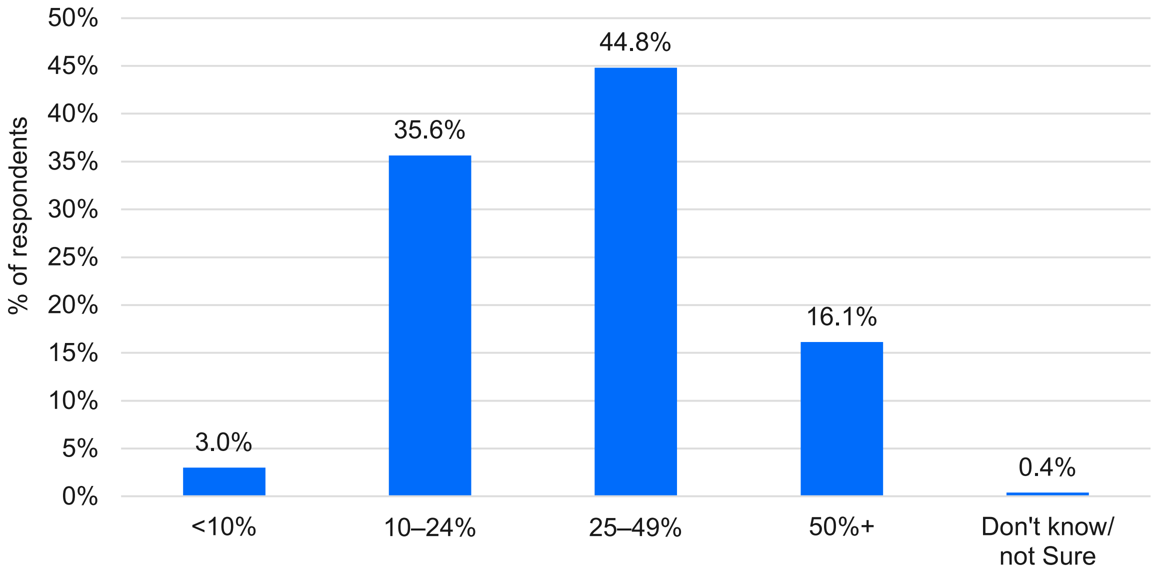
"For inferencing, you want to have fast local access happening on a particular data set and you want to colocate the data with the compute. Let's say you have 100TB of model weights. In one region, that's a very manageable amount of space. But when you want those model weights available for fast loading and fast inferencing across 40 regions, you take that 100TB and multiply it by 40. And now that has just exploded, just on model weights alone."

— Senior technology executive, hyperscale cloud service provider

FIGURE 3

AI-driven data growth over the past 12 months

Q. Over the past 12 months, by approximately what percentage did your organization's total data volume grow as a result of AI or GenAI workloads and content generation?



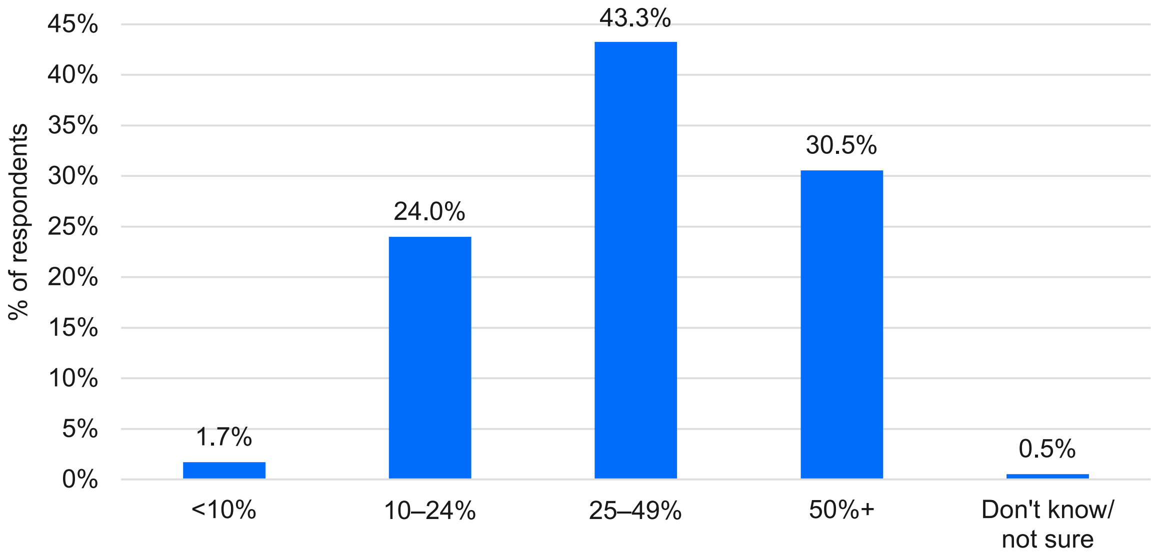
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

FIGURE 4

AI-driven data growth over the next three years

Q. Over the next three years, by approximately what percentage do you expect your organization's total data volume to grow as a result of AI/GenAI workloads and content generation?



n = 763 (unweighted base)

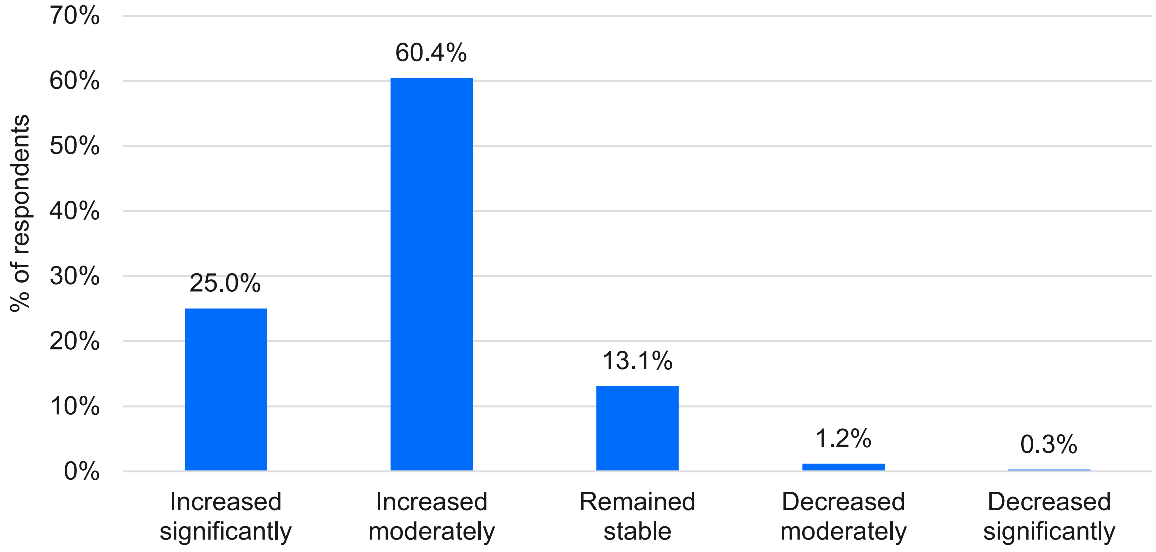
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

Data lakes tell the same story in sharper relief. More than 85% of organizations report that their data lake volume has grown over the past 12 months, and an almost identical proportion expect further growth over the next three to five years, with 40% anticipating significant expansion (see Figures 5 and 6). The primary driver of that growth is AI-generated data, including synthetic outputs, model logs, and inference results, cited by nearly 60% of respondents, followed closely by AI and machine learning training data set expansion at just over 50%. Increased data retention requirements, application modernization, and regulatory compliance round out the top drivers, underscoring that while AI is the dominant force, the growth imperative is broad based and structural (see Figure 7).

FIGURE 5

Data lake growth over the past 12 months

Q. How has the total volume of data in your organization's data lake changed over the past 12 months?



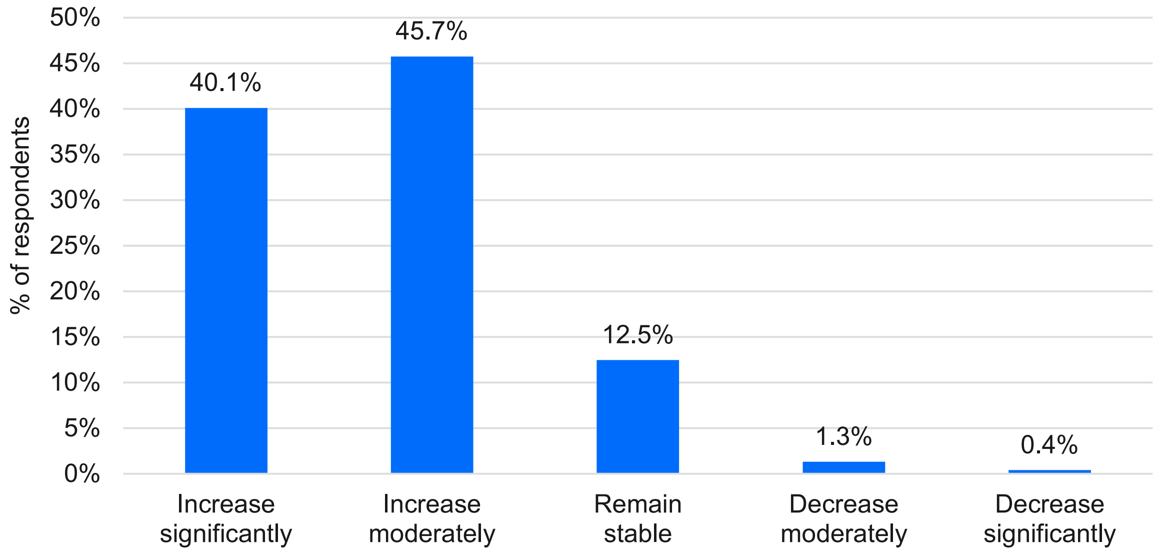
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

FIGURE 6

Data lake growth: Next three to five years

Q. How do you expect your data lake volume to change over the next three to five years?



n =763 (unweighted base)

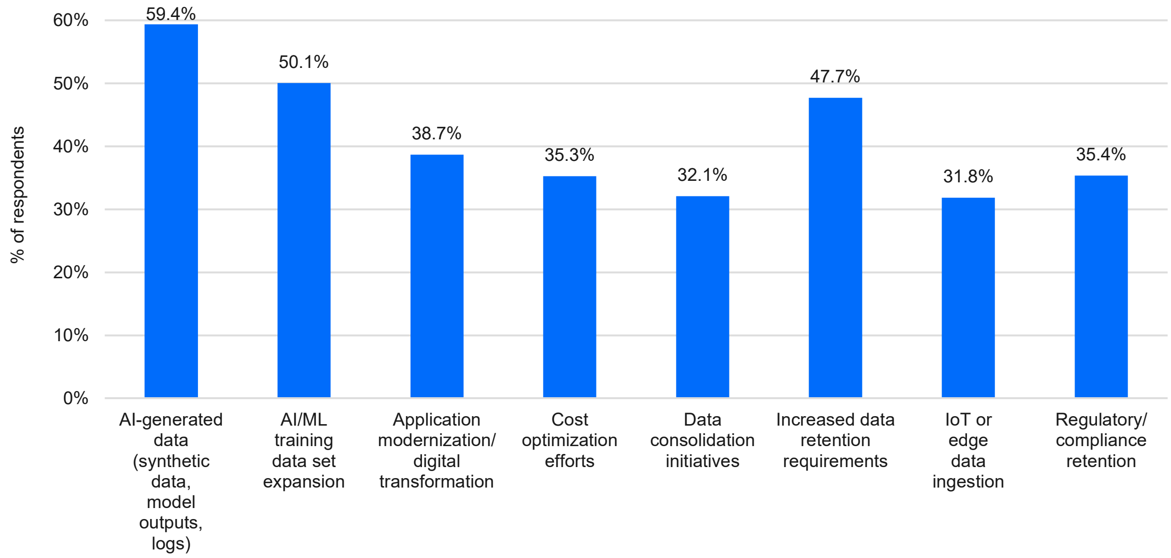
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

Organizations are not experiencing a temporary surge in data volumes but a sustained and compounding expansion driven by the nature of AI itself. For storage infrastructure, the implication is straightforward. Capacity requirements are not a problem to be solved once, but a challenge that will continue to grow in scale for the foreseeable future.

FIGURE 7

Drivers of data lake growth

Q. What are the primary drivers of change in your data lake volume?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

Shifts in data value and retention strategy

The value of enterprise data is increasing as a direct result of AI adoption, and this is changing how long organizations retain it. Study results are unambiguous on this point: nearly 95% of respondents say the value of their organization's data has increased due to AI and GenAI adoption, with a third describing that increase as significant (see Figure 8). This is not a perception shift without consequence. When data is more valuable, the calculus around retention changes fundamentally.

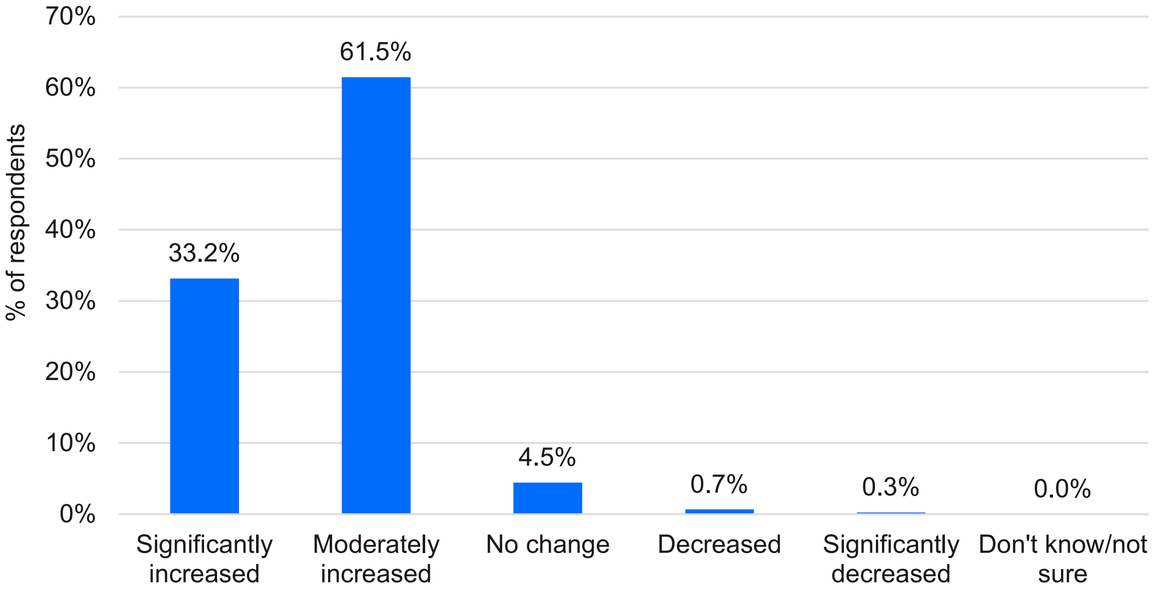
"Customers are now storing previously discarded data in anticipation of future AI value extraction — old X-rays, archived records, historical surveillance footage. No data should ever be thrown away."

— IT decision-maker, distribution and systems integration

Figure 8

AI/GenAI adoption increases data value

Q. How has the value of your organization's data changed as a result of AI/GenAI adoption?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

More than 74% of organizations say AI adoption has caused them to retain data for longer periods, and a meaningful share are now reserving substantial portions of their total data footprint specifically for AI and analytics purposes. The implications for storage infrastructure are direct: longer retention horizons and larger dedicated AI data reserves translate into sustained, cumulative growth in capacity requirements that compounds over time (see Figure 9).

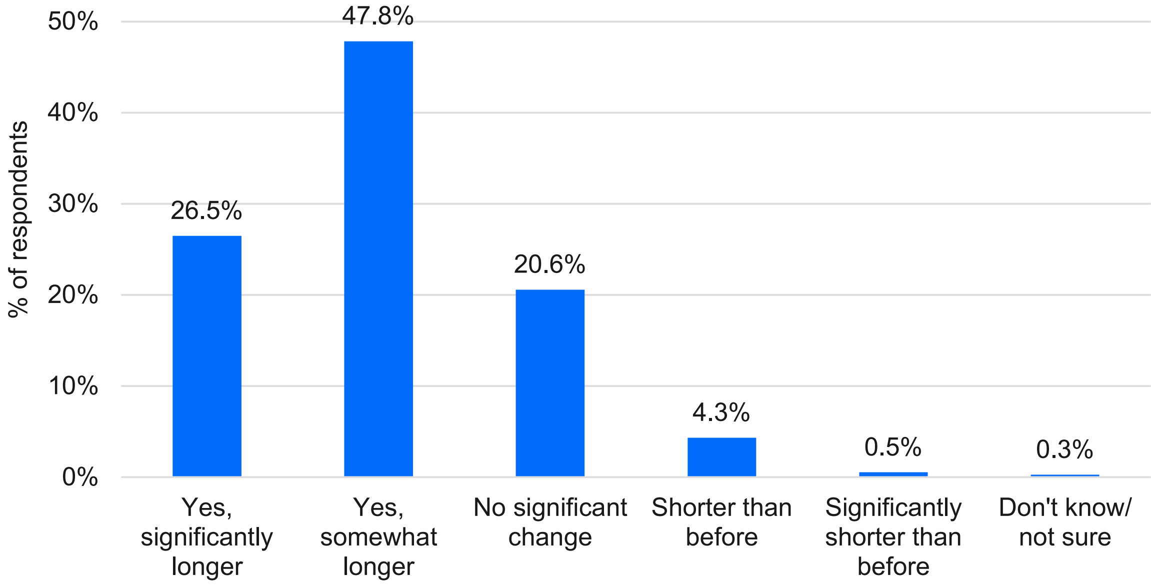
"Everything that every LLM ever generates is stored on HDD. Every prompt, every piece of media that's uploaded is being stored and retained."

— Senior technology executive, hyperscale cloud service provider

FIGURE 9

AI's impact on data retention

Q. Has AI/GenAI adoption changed how long your organization retains data?



n = 763 (unweighted base)

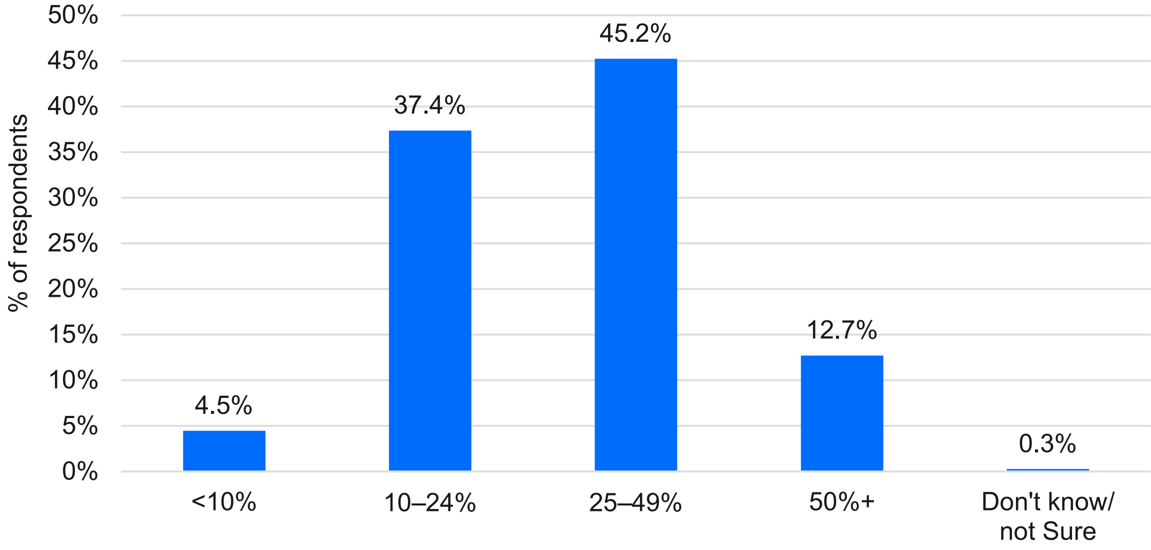
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

A critical emerging dimension of this shift is the growing strategic importance of what might be called the "cool" data tier, data accessed periodically rather than continuously, typically on a monthly or quarterly basis. As organizations retain data for periodic model retraining, compliance audit, and inference augmentation, this tier is expanding in both volume and relevance. It occupies a position that is neither hot enough to justify the cost of flash storage nor cold enough to be treated as pure archive. For storage architects, designing effectively for the cool tier is increasingly central to managing AI infrastructure economics, and it is precisely the tier where HDD economics are most compelling (see Figure 10).

FIGURE 10

Data retained for AI/ML purposes

Q. Approximately what percentage of your organization's total data is retained specifically for AI/ML or analytics purposes?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

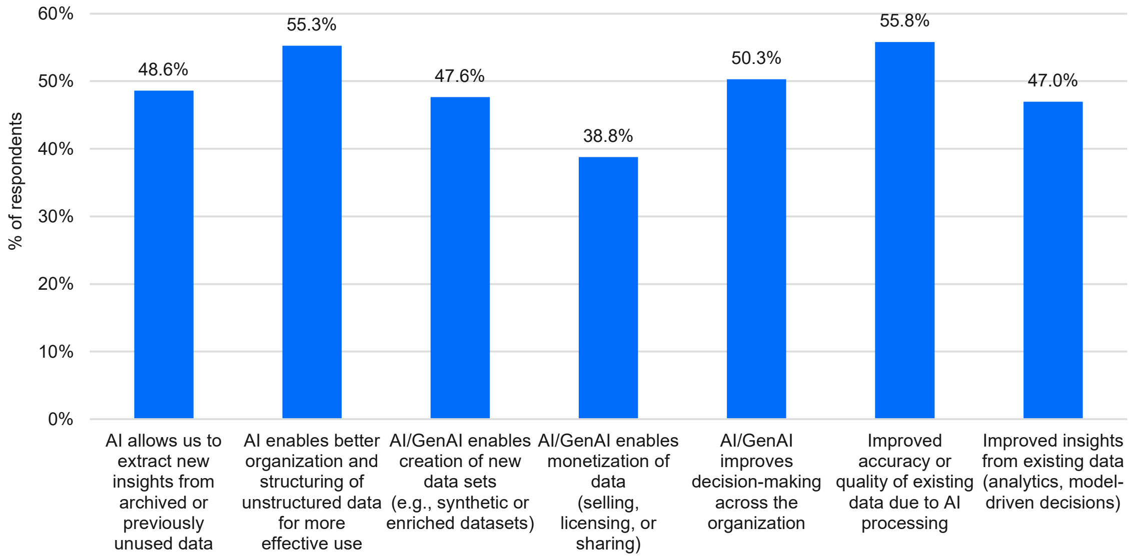
Why data value has increased, and what's driving retention

The reasons organizations cite for increased data value are broad and mutually reinforcing. At the top of the list are improvements in data quality and structure that AI processing itself enables, with more than half of respondents pointing to AI's ability to improve the accuracy of existing data and to organize and structure unstructured content in ways that were previously impractical. The creation of new synthetic and enriched data sets, improved decision-making, and the emergence of data monetization round out the picture, painting a consistent portrait of data whose value is being actively created and expanded by AI, not merely assumed (see Figure 11).

FIGURE 11

Why data value is increasing

Q. What are the main reasons you believe the value of your data has increased due to AI/GenAI?



n = 763, n = 722 (unweighted base)

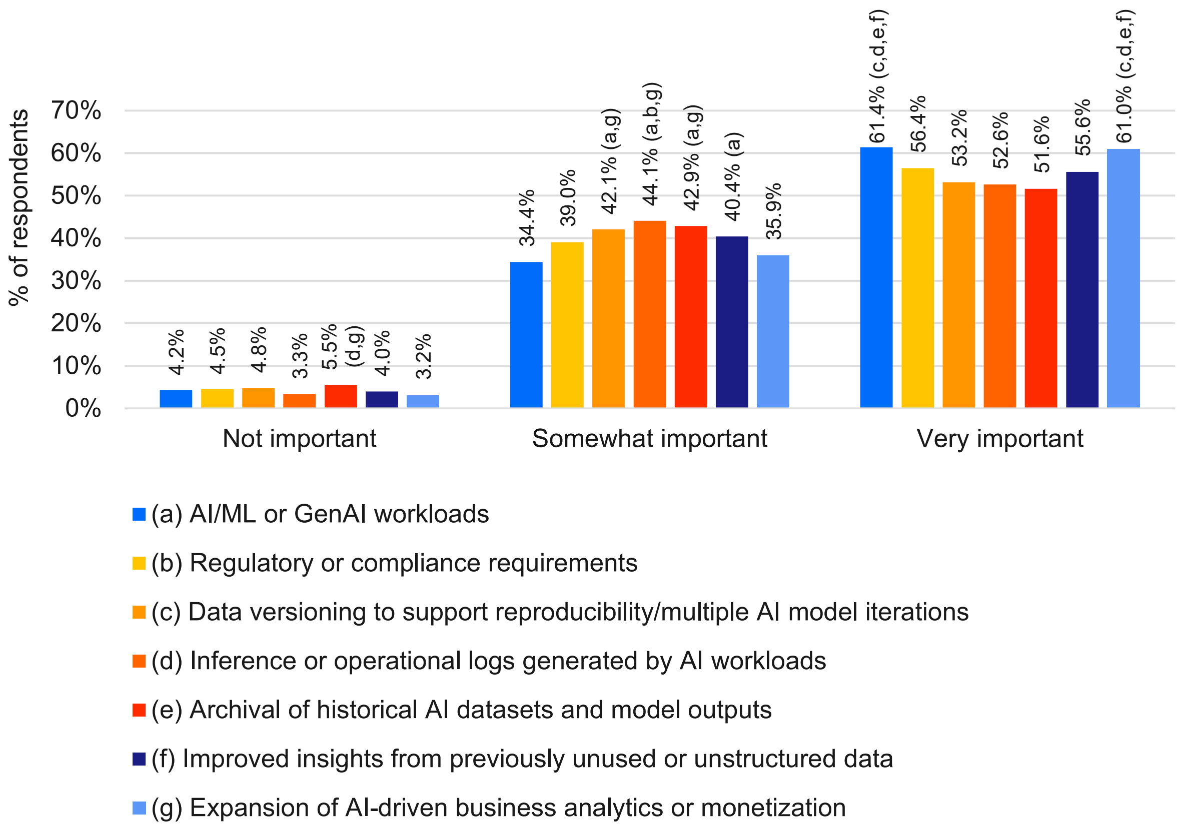
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

When asked to rate the importance of specific factors driving increased storage requirements, the results reinforce the same theme. AI and machine learning workloads and the expansion of AI-driven business analytics and monetization both rank at the top, with more than 60% of respondents rating each as very important. Regulatory compliance, improved insights from previously unused data, data versioning for reproducibility, inference logs, and archival of historical AI data sets and model outputs all follow closely, each rated very important by more than half of respondents. Taken together, these findings make a clear and important point: storage decisions are no longer primarily operational. They are increasingly strategic, reflecting the growing role of data as a long-term business asset and source of competitive advantage (see Figure 12).

FIGURE 12

Drivers of increased storage and retention

Q. Thinking about your organization's data storage, how important is each of the following factors in driving increased storage or longer retention?



n = 763, n =730 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

New data types and data set augmentation

The data types gaining the most strategic value in the AI era are predominantly high volume and unstructured. Structured business data, system and application logs, and user behavior data lead the list of types organizations expect to also appreciate in value, reflecting the breadth of inputs that modern AI systems draw on. Synthetic data generated by AI models, real-time IoT sensor data, images and scanned documents, and video all feature prominently as well, pointing to an expanding and increasingly diverse data footprint that places sustained pressure on storage capacity across multiple tiers (see Figure 13).

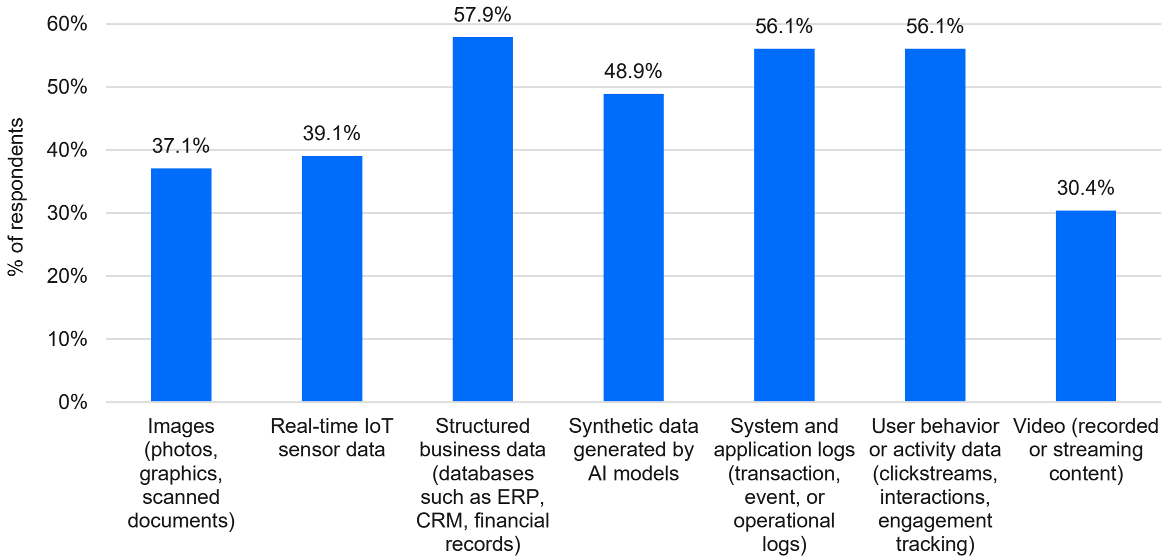
> "Universities are mining old data sets to improve AI model outputs. Governments are digitizing decades of physical records to feed training pipelines. Data that had no prior strategic value is now a training resource."
>
> — IT decision-maker, distribution and system integration

Synthetic data augmentation is a particularly significant driver of this expansion. Virtually all respondents confirm that augmenting data sets with synthetic data has either increased the value of existing data, led to the creation of entirely new data sets or both (see Figure 14). Meanwhile, AI augmented and enriched existing data sets, logs and telemetry, and user-generated content enhanced by AI are among the top data types driving actual storage growth today, alongside training data sets, inference outputs, and model checkpoints (see Figure 15).

FIGURE 13

Data types gaining the most value

Q. Which types of data do you expect will experience the greatest increase in value due to AI/GenAI?



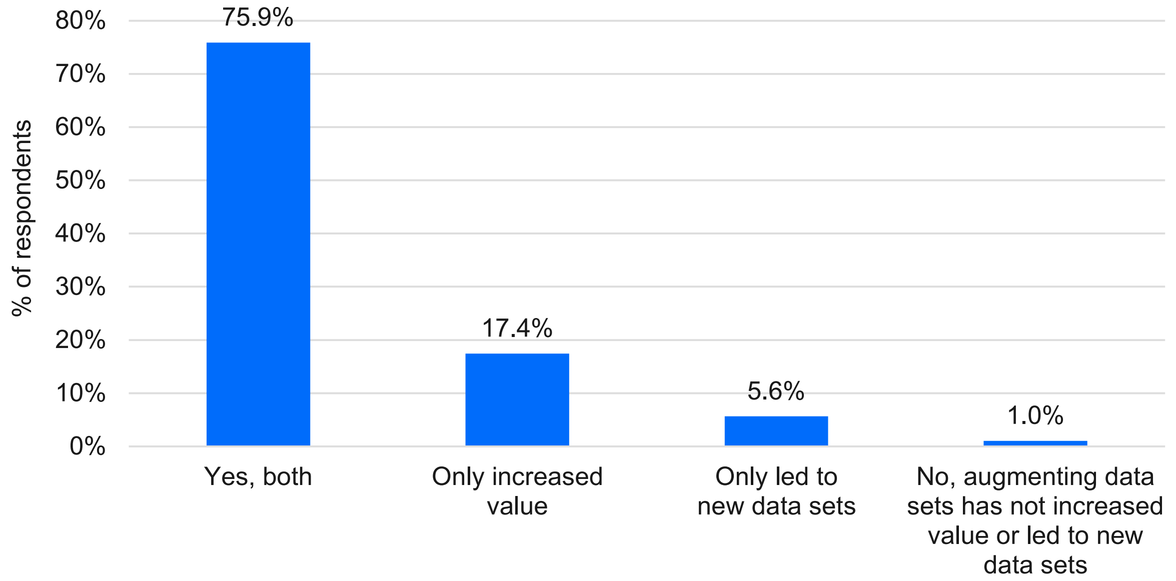
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

FIGURE 14

Impact of synthetic data augmentation

Q. Has augmenting datasets (e.g., synthetic data) increased value of existing data or led to new data sets?



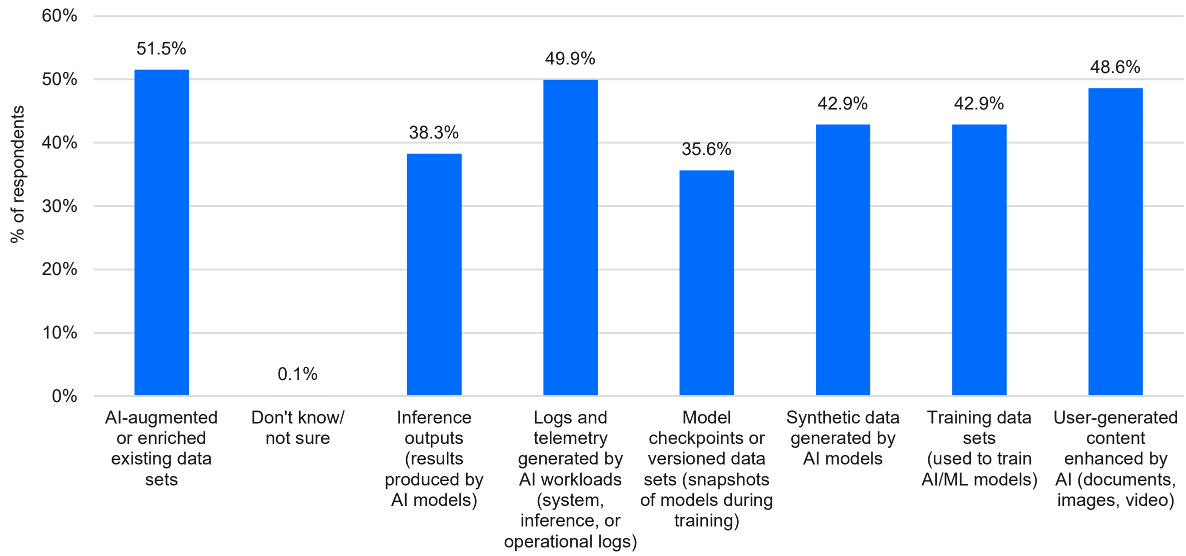
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decisionmaker Survey,* 2026

FIGURE 15

Top contributors to AI-driven data growth

Q. Which types of AI-generated data are contributing most to your organization's data growth?



n = 763 (unweighted base)

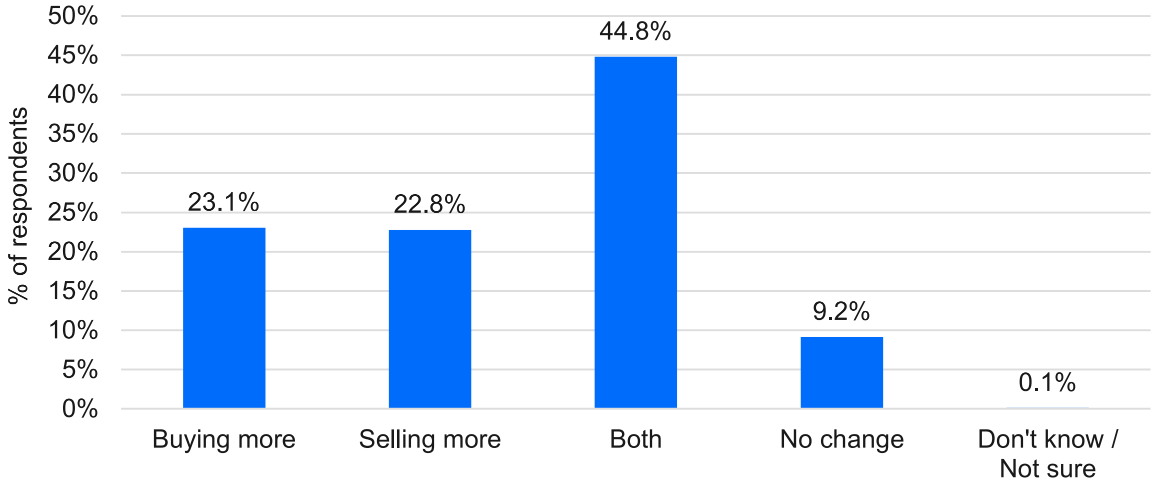
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

Organizations are also increasingly treating data as a tradeable asset rather than a purely internal resource. More than 90% report buying or selling more data sets than in previous years, with nearly half doing both, signaling the emergence of a data economy in which the quality and scale of an organization's data holdings carry direct commercial value (see Figures 16 and 17).

FIGURE 16

Buying and selling data sets

Q. Are you buying or selling more datasets than in previous years?



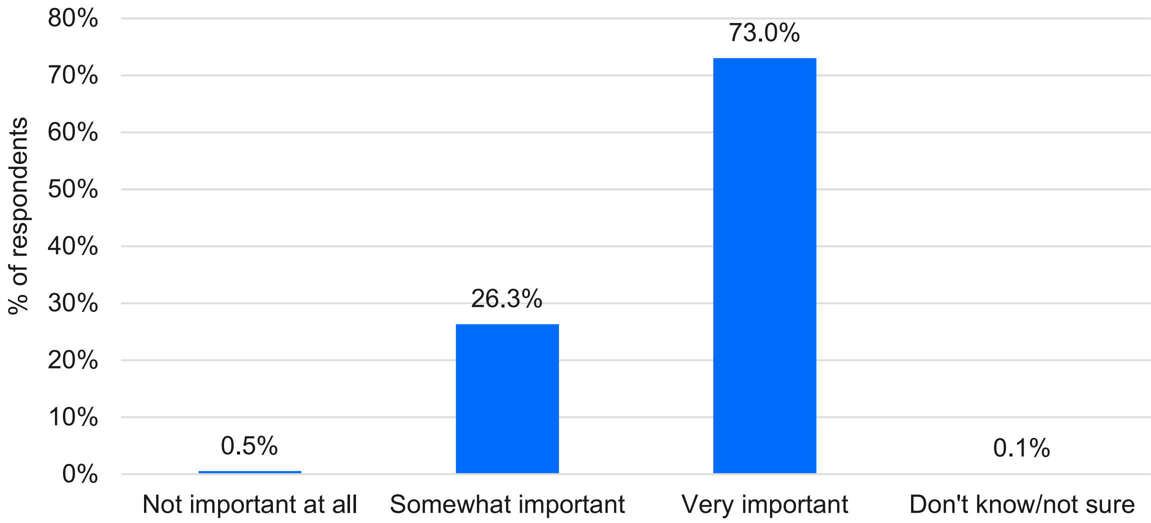
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

FIGURE 17

Data quality's role in AI accuracy

Q. How important are bigger, higher-quality data sets for improving AI model accuracy?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

Changing perceptions and storage requirements

AI is not only changing how much storage organizations need but changing how they think about storage. Storage is no longer passive infrastructure, rather it is becoming an active component of AI architecture. Where storage decisions were once driven primarily by capacity and cost, the AI era is introducing a more complex set of requirements: reliability at scale, operational simplicity across tiered environments, energy efficiency, and the ability to retrieve data rapidly from what was previously considered passive archive. Workload type and AI life-cycle phase are emerging as the primary variables shaping storage decision-making, and the data shows clearly that different phases of the AI life-cycle place materially different demands on storage infrastructure. As a result, organizations that resist one-size-fits-all approaches and instead architect storage around workload-specific requirements will build infrastructure that is both more cost effective and more resilient.

"You can think of having more storage as an optimization for keeping the GPUs busy. The biggest challenge is to ensure that the GPUs are not idle waiting for data to show up — hours of GPU idle time becomes very expensive, very fast."

— Senior product manager, hyperscale cloud service provider

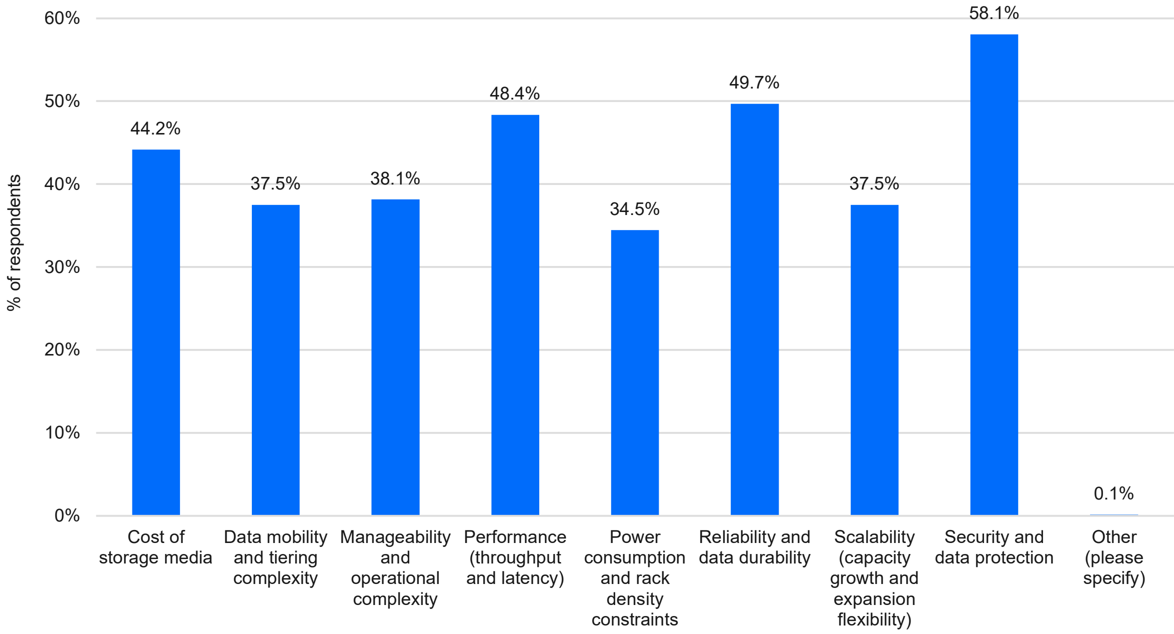
Customer perceptions of storage in the AI era

Security and data protection, reliability, and performance top the list of storage concerns among AI decision-makers, but what is striking about the full picture is how evenly distributed concern is across a broad range of dimensions. Cost of storage media, manageability and operational complexity, data mobility and tiering complexity, scalability, and power consumption all register as significant concerns for between a third and a half of respondents. No single issue dominates, and no single technology addresses all of them simultaneously. This is the multidimensional reality that storage architects in the AI era are navigating (see Figure 18).

FIGURE 18

Top storage concerns for AI workloads

Q. What are your organization's biggest concerns regarding storage for AI workloads?



n = 763 (unweighted base)

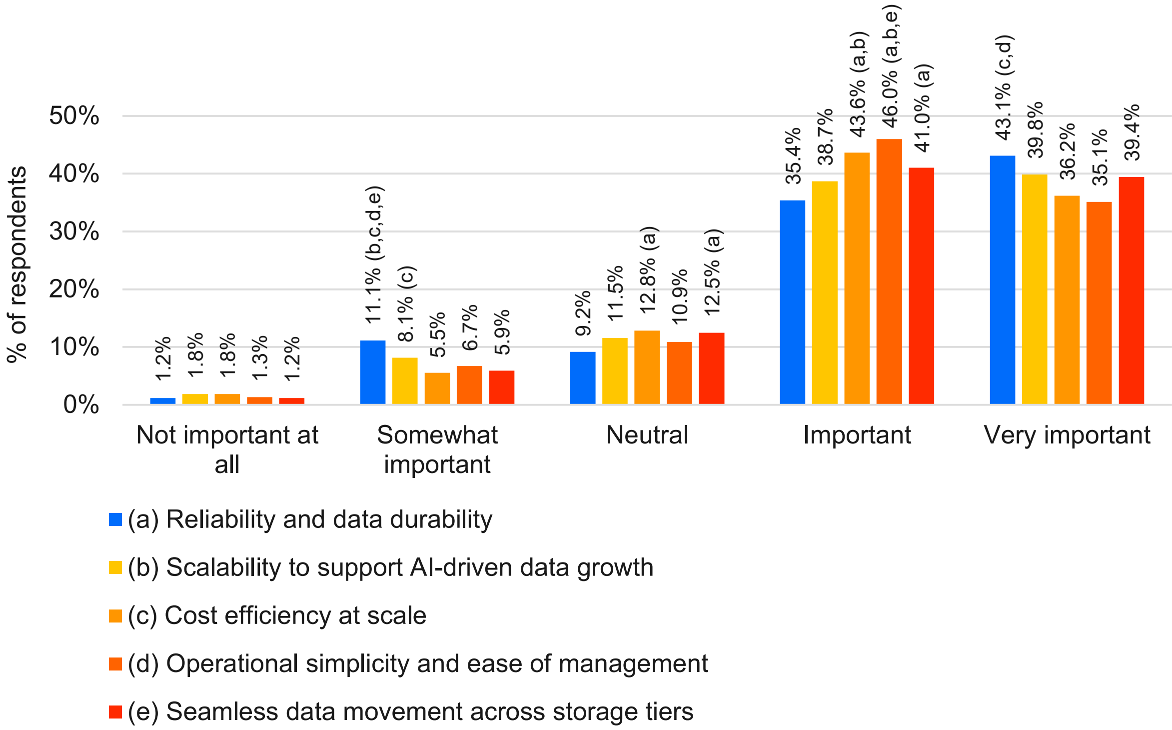
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

When asked to rate the importance of specific storage platform attributes, respondents rate operational simplicity and ease of management highest, with more than 81% rating it important or very important, marginally ahead of seamless data movement across storage tiers, cost efficiency at scale, reliability, and scalability. The consistency of these ratings across attributes is itself informative: organizations are not willing to trade one dimension for another. They want storage platforms that deliver on simplicity, cost, reliability, and scalability together (see Figure 19). That is a demanding combination, and it goes a long way toward explaining why more than half of respondents describe managing AI storage tiers as complex, with 1 in 10 describing it as very complex. The operational challenge of multitier AI storage is real, widely felt, and not yet resolved for most organizations (see Figure 20).

FIGURE 19

Key attributes for AI storage platforms

Q. Which attributes are most important when selecting a storage platform for AI workloads?



n = 763 (unweighted base)

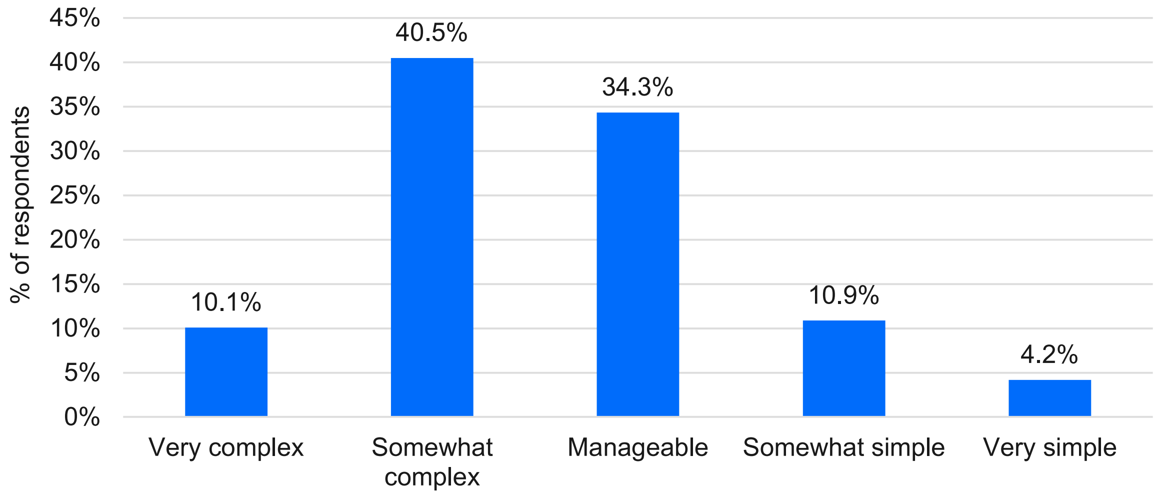
Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

What this data also reveals is a macrotrend that deserves explicit recognition: AI investments are increasingly constrained not only by compute but by storage architecture. When GPUs sit idle waiting for data, the cost of that idle time compounds rapidly, and the bottleneck is not processing power, but the storage infrastructure feeding it. For enterprise leaders, this reframes storage from a supporting cost into a primary determinant of AI ROI.

FIGURE 20

Complexity of managing AI storage tiers

Q. How would you rate the complexity of managing multiple storage tiers for AI workloads?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

How organizations balance simplicity and cost

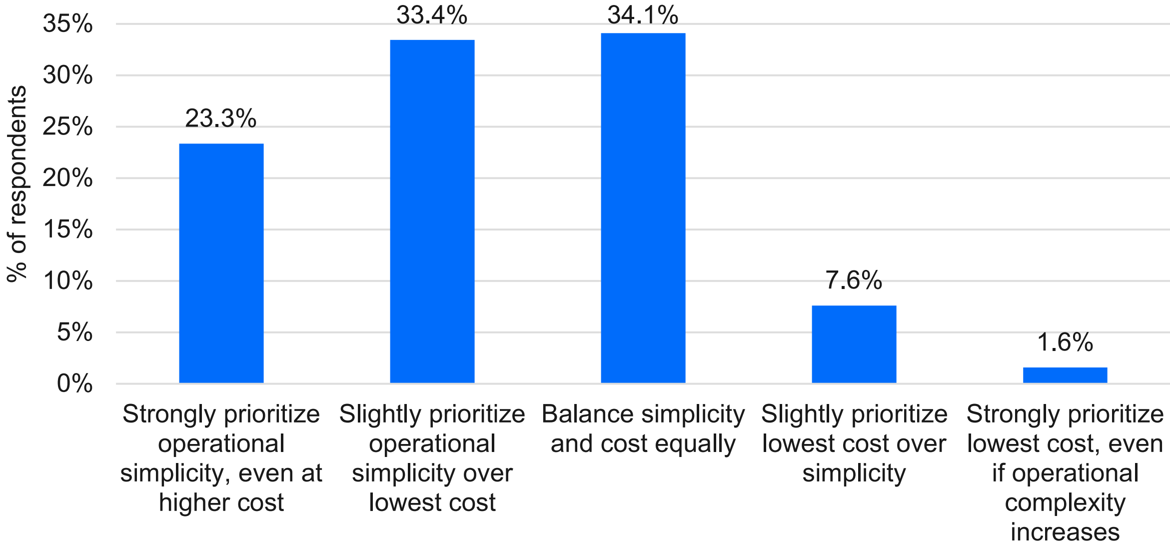
One of the more revealing findings in this study concerns how organizations trade off operational simplicity against cost minimization when evaluating AI storage solutions. The conventional assumption might be that cost dominates enterprise storage decisions, particularly at scale. The data tells a more nuanced story. More than 90% of respondents prioritize operational simplicity at least equally alongside cost, with the largest single group, at just over a third of respondents, saying they balance both equally. A further 23% strongly prioritize simplicity even at higher cost, and another third slightly favor simplicity over the lowest price. Fewer than 1 in 10 are predominantly cost-first buyers willing to accept greater complexity in exchange for lower spend (see Figure 21).

For HDD platform vendors, this is a significant finding. It suggests that the market for integrated, easy-to-manage storage solutions is considerably larger than a pure cost-per-terabyte argument would imply. Organizations indicate in the survey results that they are not simply looking for the cheapest storage. They are looking for storage that is easy to operate, easy to scale, and easy to integrate into increasingly complex AI environments. Vendors that can deliver on both dimensions simultaneously, cost efficiency, and operational simplicity are addressing the priorities of the overwhelming majority of the market.

FIGURE 21

Simplicity versus cost in storage decisions

Q. When evaluating AI storage solutions, how does your organization balance operational simplicity versus minimizing cost at scale?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

Storage platform usage for AI workloads

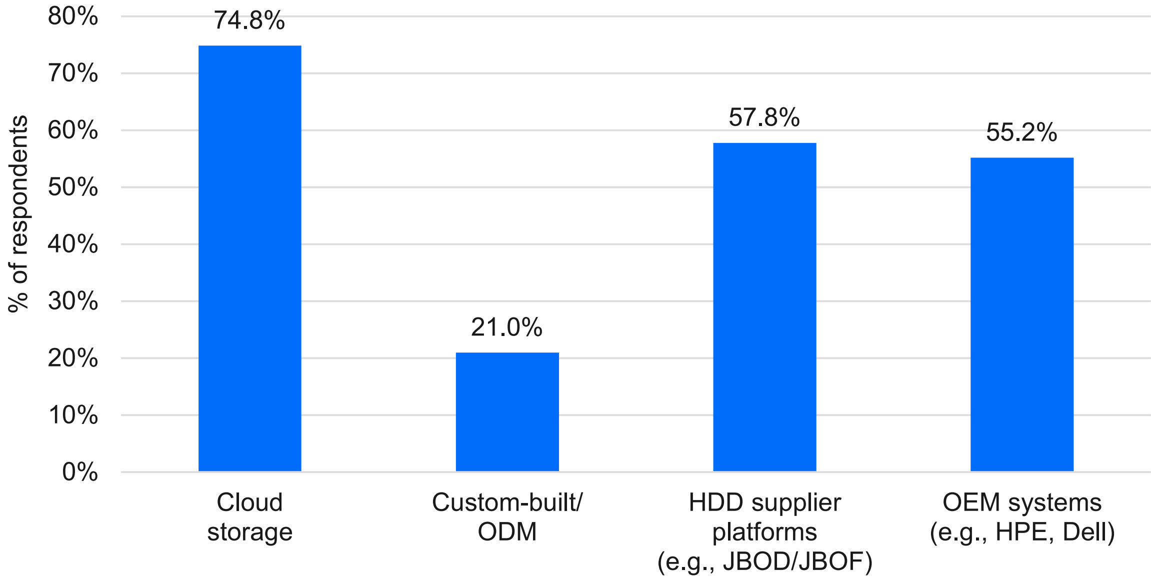
The storage platform landscape for AI workloads is heterogeneous, reflecting the diversity of deployment environments and workload types that characterize enterprise AI today. Cloud storage leads, deployed by nearly 75% of respondents, while on-premises HDD supplier platforms including JBODs and JBOFs are deployed by more than half, placing them second in the rankings and ahead of traditional OEM systems from vendors such as HPE and Dell. For organizations evaluating their AI storage architecture, this breadth of HDD deployment is an encouraging indicator of real-world adoption across a range of enterprise environments (see Figure 22).

That said, a balanced reading of this data is warranted. The platforms and systems business for HDD suppliers remains relatively small, highly competitive, and relatively nascent compared with the entrenched positions of cloud hyperscalers and established OEM vendors. Survey results may reflect the perception that an HDD platform consists of the drives themselves rather than an actual system even though JBOD and JBOF were specified. In any case, the results reflect market deployment, not market share, and cloud and OEM players bring considerable scale, ecosystem depth, and enterprise relationships to this space. The signal is promising, but the competitive landscape for HDD-based platforms remains demanding.

FIGURE 22

Storage platforms used for AI workloads

Q. Which storage platform types do you use for AI workloads?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

Reliability, performance, and trade-offs at scale

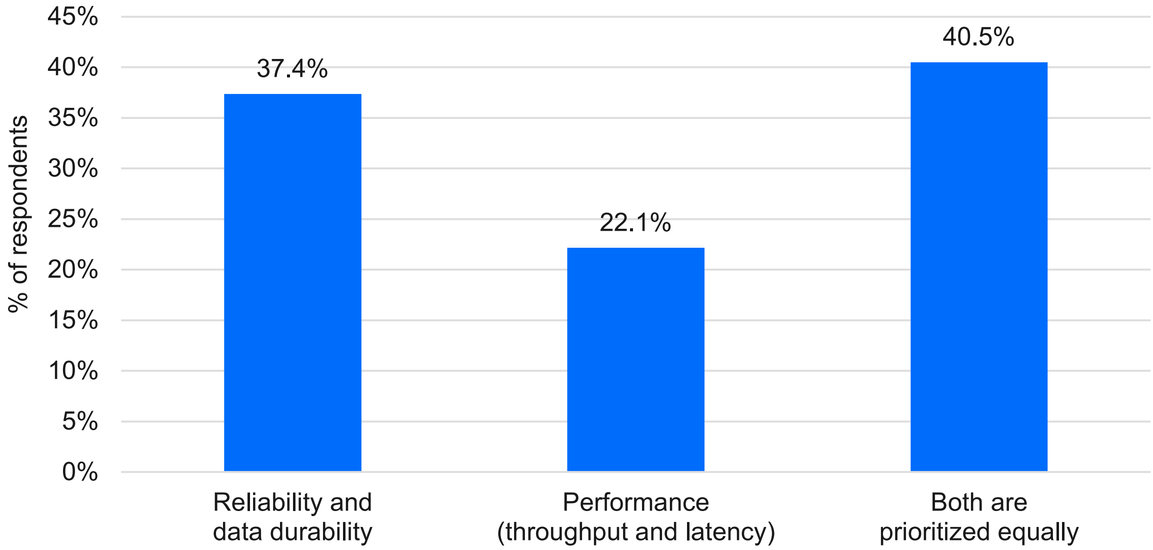
When storage costs must be controlled at scale, organizations make clear and consistent trade-offs. Simplicity of management is the attribute most commonly sacrificed, cited by 27% of respondents, followed closely by energy efficiency and peak performance. The willingness to compromise on peak performance is worth noting, as it reflects a pragmatic and workload-appropriate approach to storage decision-making in which raw speed is not always the overriding concern at the capacity tier.

When asked to choose a single overriding priority for AI storage design overall, reliability and data durability lead performance by a wide margin among those who select one over the other, with 37% prioritizing reliability compared with 22% prioritizing performance. The largest single group, just over 40%, rates both as equally important. These findings suggest that organizations are making deliberate trade-offs when designing storage infrastructure for AI, balancing performance requirements against the realities of scale, reliability, and cost (see Figure 23).

FIGURE 23

Reliability versus performance priorities

Q. When designing storage infrastructure for AI workloads overall, which attribute does your organization prioritize more?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

Storage priorities across the AI life cycle

Storage attribute priorities shift meaningfully depending on where an organization is in the AI life cycle. Model training is the one phase where performance clearly dominates, with 45% of respondents prioritizing throughput and latency, reflecting the sequential read demands of GPU-driven training workloads. Everywhere else, the picture looks different. At data ingestion, reliability leads. At inference and production deployment, the largest group rates both equally. For AI-generated data storage including logs, outputs, and synthetic data, the split is near perfect. And for long-term retention and archive, reliability leads performance comfortably.

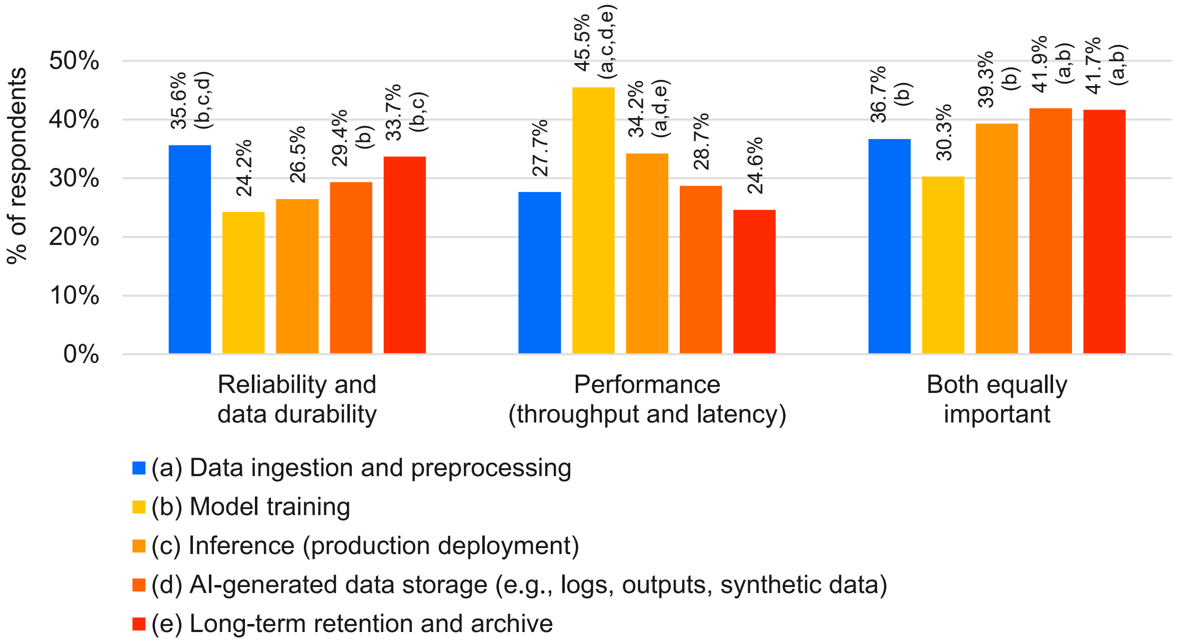
> "Checkpointing is mostly an HDD workload. We run hundreds of terabytes of checkpoint data when those training jobs are running, and we're doing it to HDD."
>
> — Senior engineer, hyperscale cloud service provider

The practical implication is that performance-first storage is the right answer for one phase of the AI life cycle and a secondary consideration for the remaining four. For the vast majority of AI data by volume, which accumulates across ingestion, inference, generated outputs, and long-term retention, reliability, capacity, and cost are the attributes that matter most (see Figure 24).

FIGURE 24

Storage priorities across the AI life cycle

Q. For each phase of the AI life cycle, which attribute is more critical from a storage perspective?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

The archive retrieval imperative: RAG and AI inference

One of the more consequential shifts surfaced by this study is the growing expectation that archived data must be retrievable faster than traditional cold storage architectures were designed to support. Retrieval-augmented generation pipelines and AI inference workloads are increasingly drawing on historical and archived data in near real time, effectively blurring the boundary between cold and warm storage tiers. Nearly all respondents, 96%, anticipate needing faster archive retrieval to support these workloads, with 45% saying the need is significant.

"Customers deploy large SSD arrays for model training — then post-training they dismantle the SSDs and install HDDs at the back as backbone storage for inference."

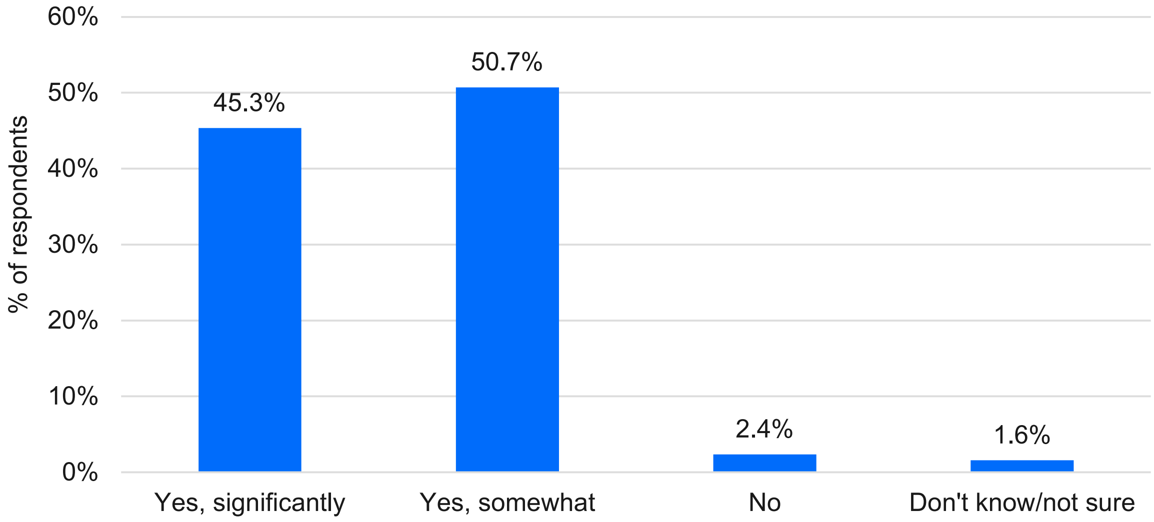
— IT decision-maker, distribution and systems integration

This has direct implications for how organizations design and manage high-capacity storage archives going forward. Storage that was previously treated as passive and access-infrequent is being asked to participate more actively in live AI workflows. This is a reflection of the broader shift from storage as passive repository to storage as active infrastructure. The response is not necessarily to replace capacity-tier archive infrastructure with flash, but to architect large-scale storage deployments with retrieval performance and data accessibility in mind, ensuring that the economics of high-capacity storage do not come at the expense of the responsiveness that modern AI pipelines increasingly demand (see Figure 25).

FIGURE 25

Demand for faster archive retrieval

Q. Do you anticipate needing faster archive retrieval to support AI inference or retrieval-augmented generation (RAG) pipelines?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

The value proposition of HDDs in the AI era

HDDs offer a compelling and durable value proposition in the AI era, not despite the rise of AI, but because of it. The explosion in data volumes, the extension of retention horizons, the reactivation of previously dormant archives, and the emergence of new data types across the AI life cycle all point toward a sustained and growing need for high-capacity, cost-effective, and reliable storage at scale. These are the conditions under which HDDs have always performed best, and the AI era is amplifying that and reshaping storage economics in their favor. The case rests on five dimensions: total cost of ownership, reliability and scalability, performance alignment with the majority of AI workloads, energy efficiency, and the economic comparison with flash storage for capacity-tier use cases. What the survey data and practitioner insights in this study consistently show is that across each of these dimensions, the requirements of the AI era align more closely with HDD strengths than is commonly assumed. Because AI is shifting storage from a capacity management problem to an infrastructure economics problem, that makes HDDs not merely relevant but indispensable.

TCO and acquisition cost advantages

TCO per terabyte is a near-universal decision criterion in enterprise storage, rated important or very important by more than 98% of respondents, a finding that itself reflects how AI is changing the economics of storage decisions. When asked which medium offers lower TCO per terabyte across different workload tiers, respondents give SSD a modest edge in hot and warm tiers, reflecting the performance premium those workloads command. In the cool tier, the two are statistically tied, and in the cold tier, HDD edges ahead. For capacity-tier workloads where the majority of AI data resides, the TCO case for HDD is at minimum competitive and in many cases favorable.

"SSDs at 10 times the price of HDD cannot address exabyte-scale AI storage needs economically. The AI inflection point re-diverged SSD/HDD price-per-terabyte trajectories rather than closing them."

— IT decision-maker, distribution and systems integration

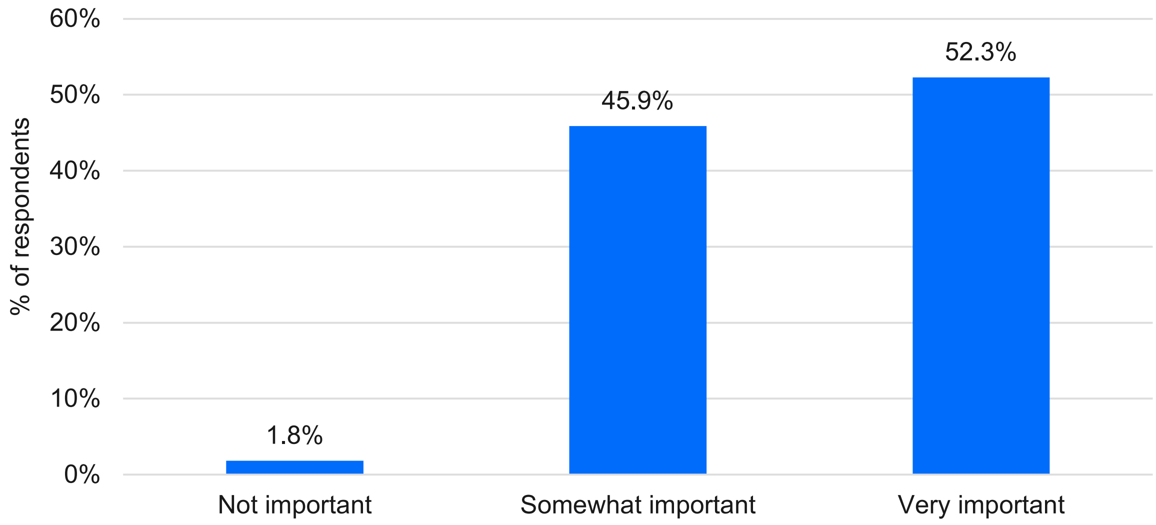
"HDD is 4 to 15 times cheaper (than flash), no matter what you do. We know the math, we have the financial models, and we can just plug it in, and if you don't make ridiculous assumptions, HDD is just far and away better."

— Technical fellow, hyperscale cloud service provider

FIGURE 26

TCO per terabyte as a decision factor

Q. How important is TCO per terabyte in your storage decisions?



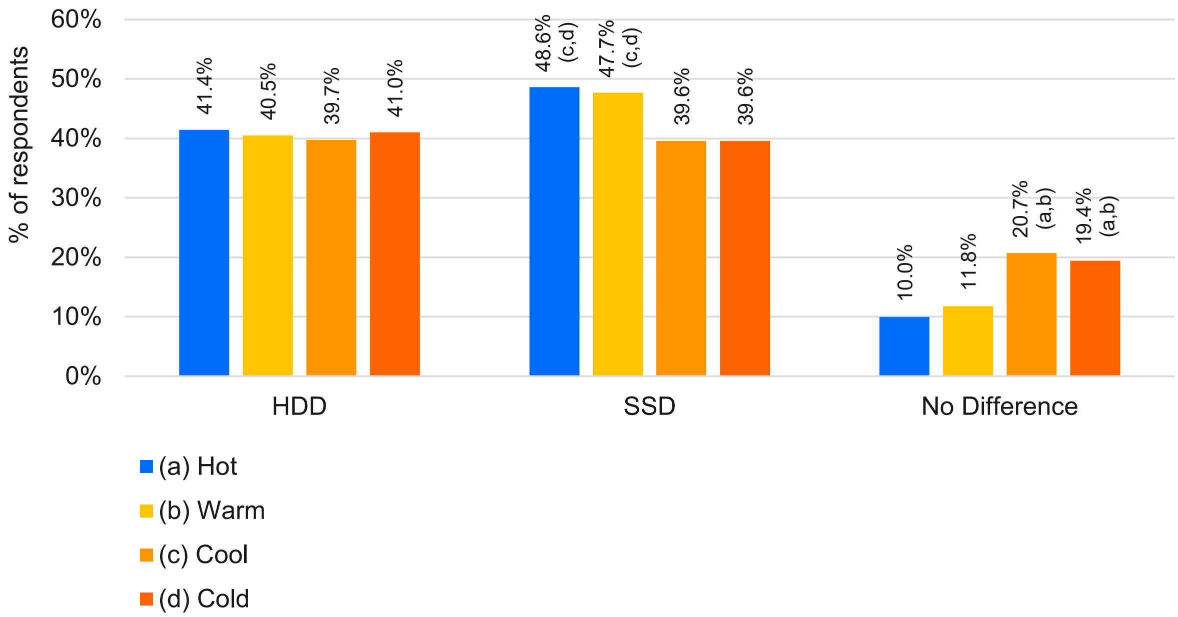
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

FIGURE 27

HDD versus SSD: TCO by access tier

Q. Which has lower TCO per terabyte (TB) for a given workload?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

Reliability, performance, and energy efficiency at scale

AI workloads require long-term, persistent data retention, and reliability is the attribute organizations most consistently prioritize when designing storage infrastructure for AI. As established earlier in this study, reliability leads performance as the top single storage design priority, and HDD's reliability profile at scale, combined with a technology road map that continues to advance capacity without sacrificing durability, makes it well suited to the sustained retention requirements of the AI era.

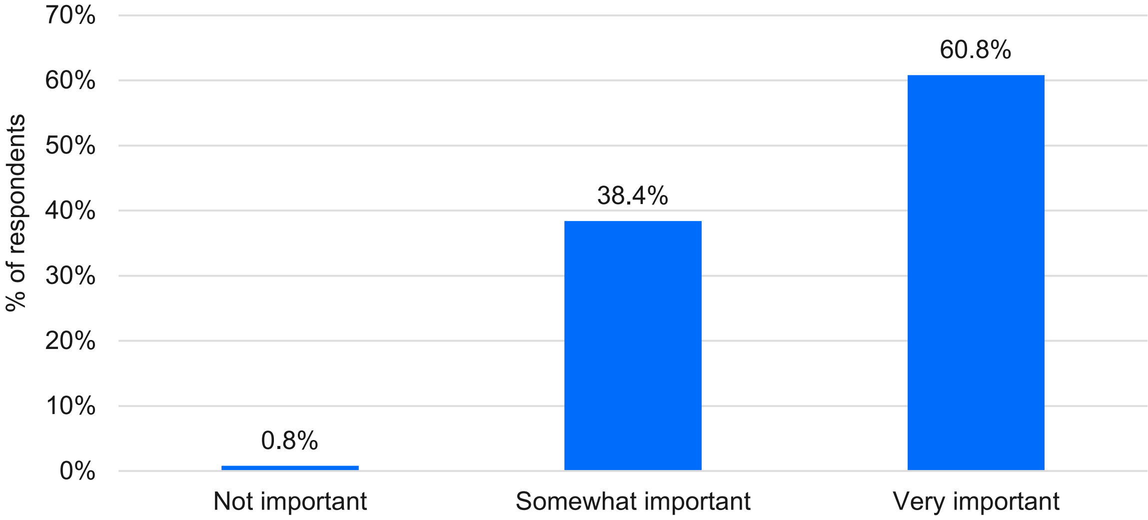
On performance, the relevant question is not whether HDD can match SSD across every workload, it cannot and is not expected to. The relevant question is whether HDD throughput is sufficient for the workloads where it will actually be deployed. The life-cycle phase data in this study shows that performance clearly dominates only in model training. Across every other phase, including data ingestion, inference, AI-generated data storage, and long-term retention, reliability leads or ties, and the economic case for HDD is strong.

Energy efficiency is also emerging as a meaningful purchasing criterion, rated important or very important by more than 99% of respondents, with 61% rating it very important. SSD holds a modest perception advantage over HDD on this dimension, but the gap is narrow at under 9 percentage points, and it narrows further when evaluated on an energy-per-terabyte basis at capacity-tier scale. HDD architectures, when measured against the absolute power draw required to store exabytes of data, offer a compelling efficiency case that the market has not yet fully recognized (see Figures 28 and 29).

FIGURE 28

Energy efficiency in storage selection

Q. How important is energy efficiency in storage selection for AI workloads?



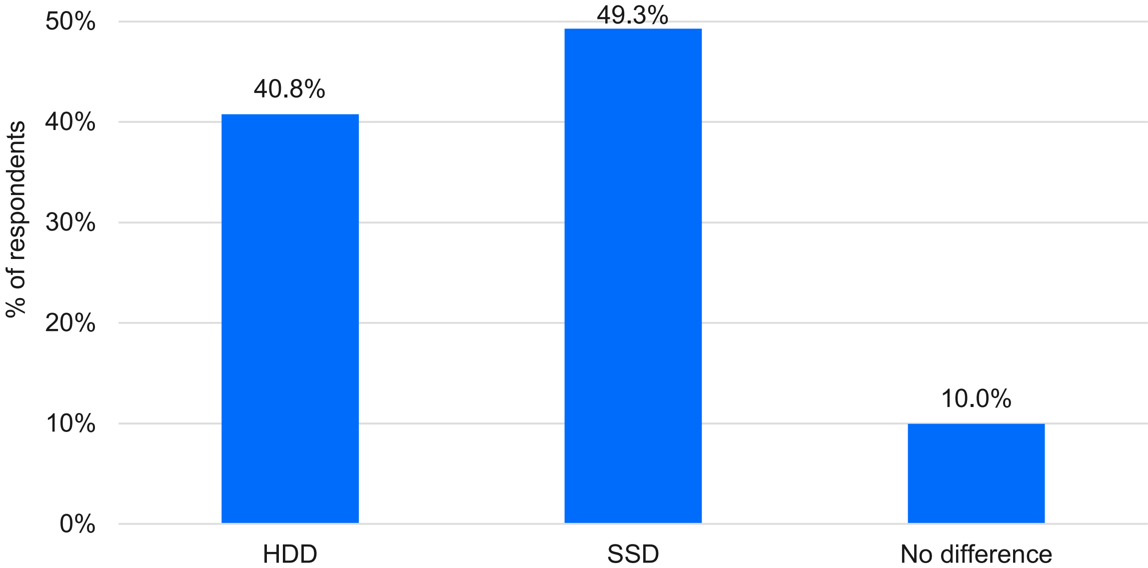
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

FIGURE 29

HDD versus SSD: Energy efficiency perceptions

Q. Which do you perceive as more energy efficient for your workloads?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

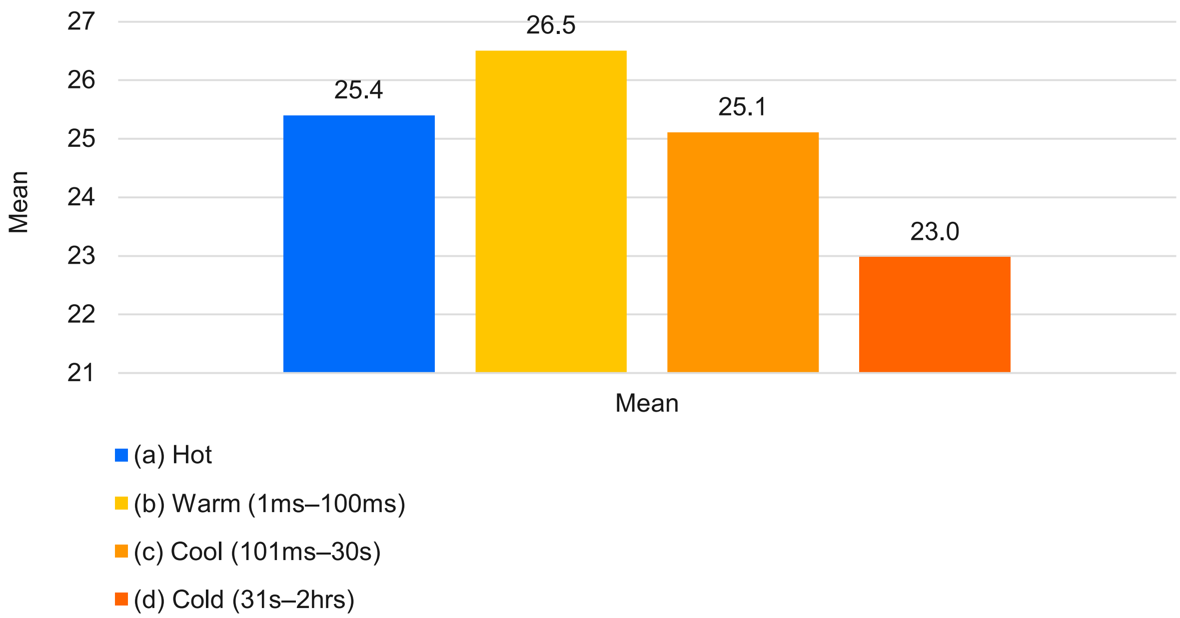
HDDs and SSDs are complementary

Most enterprises already operate a naturally tiered storage architecture in which HDD and SSD play distinct and complementary roles. Survey data confirms that HDD has meaningful presence across all four access tiers, not just cold storage, with nearly half of organizations deploying HDD in warm-tier environments. Warm, cool, and cold tiers together account for nearly 75% of total stored data on average, and more than 60% of data lake volume is currently cold or infrequently accessed. All four tiers are expected to grow over the next five years, with the warm tier representing the single largest incremental opportunity for cost-efficient, high-capacity storage (see Figures 30 and 31).

FIGURE 30

Data distribution across access tiers

Q. Approximately what percentage of your organization's total stored data falls into each access tier today?



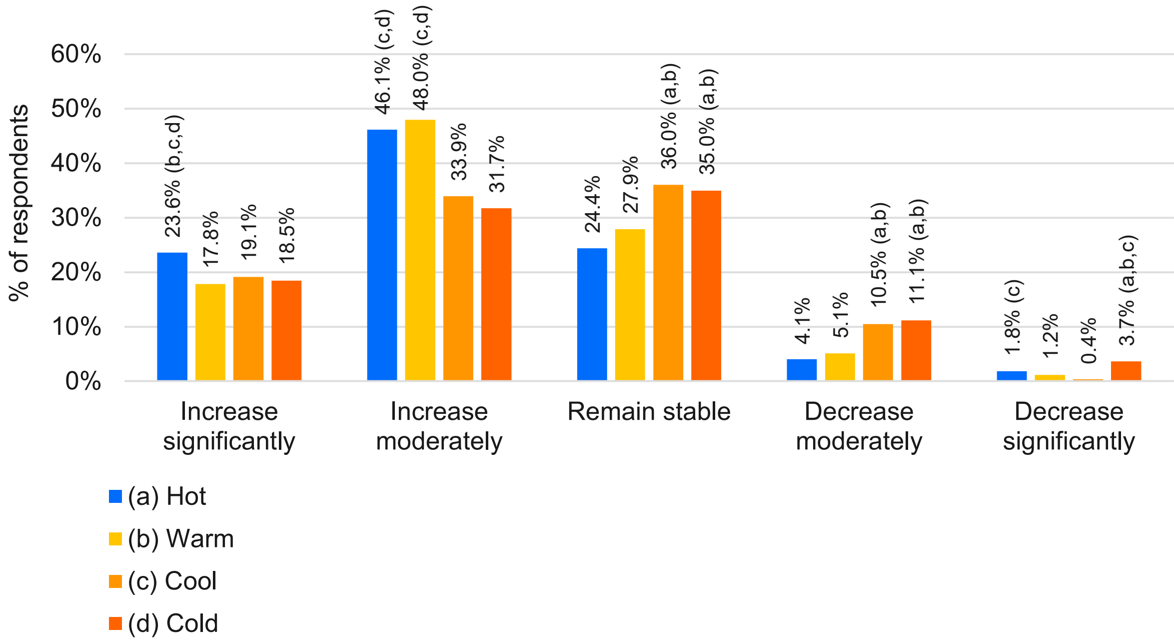
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

FIGURE 31

Access tier growth: Next five years

Q. How do you expect the amount of data in each tier to change over the next five years?



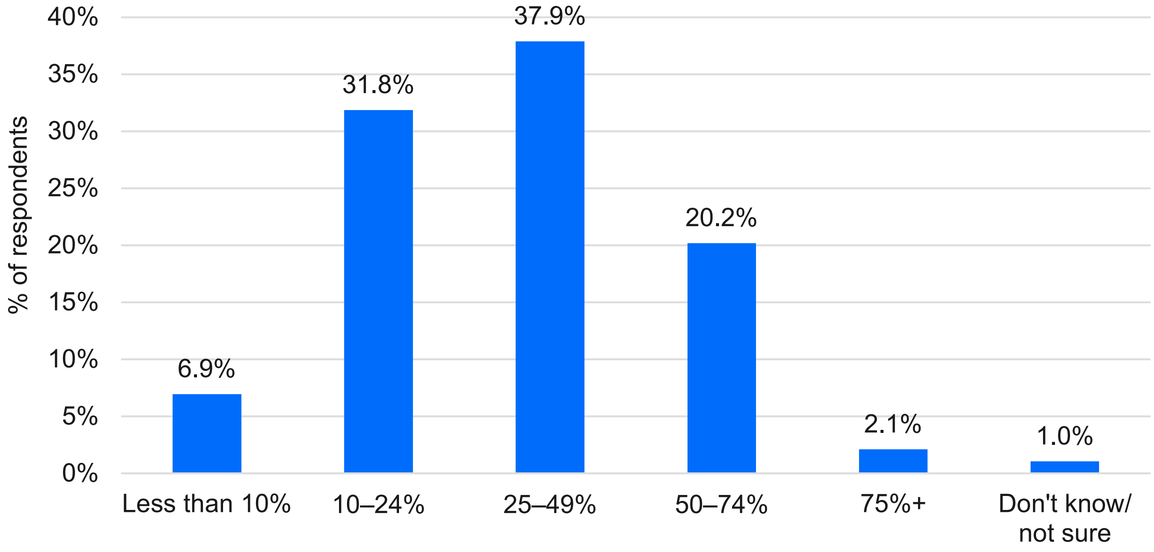
n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

The position of this paper is not that HDD should replace SSD. Flash storage will continue to play an essential role where latency and throughput are genuinely critical. The argument is more straightforward. A well-architected AI storage stack uses each medium where it performs best, and for the vast majority of AI data by volume, HDD delivers the capacity, economics, and reliability that the workload requires (see Figure 32).

FIGURE 32
Cold data share of the data lake

Q. Approximately what percentage of your organization's data lake is currently considered cold or infrequently accessed?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

HDDs in the AI life cycle

The narrative that HDDs are an archival technology waiting at the end of the data pipeline understates their role considerably. In practice, HDDs are present and active across multiple phases of the AI data life cycle, from the initial ingestion of raw and historical data through training data set staging, model checkpointing, inference backbone storage, data creation, and long-term retention. The economics and reliability characteristics that make HDDs compelling for archive apply equally across these phases, and the volume of data involved at each stage means that the cost implications of media choice compound quickly at scale. The in-depth interviews conducted for this study surface operational examples that bring this to life across healthcare, government, and enterprise technology environments, illustrating how organizations are deploying HDD-based infrastructure not as a passive repository but as an active and load-bearing component of their AI stack. Storage, in other words, is no longer passive infrastructure. It is an active component of AI architecture.

Role of HDDs across AI life-cycle phases

At the data ingestion stage, HDD serves as the natural landing zone for the raw, unstructured, and historically archived data that organizations are increasingly reactivating for AI training purposes. During data preparation, the economics of staging large training data sets at scale consistently favor HDD over flash. In model training, HDD provides the high-capacity backdrop against which active computation runs on SSD and GPU memory, storing the training data that feeds the pipeline without requiring the latency profile of flash storage. Checkpointing, the frequent preservation of model state during training, is a high-volume sequential write workload that utilizes both SSD and HDDs depending on the use case.

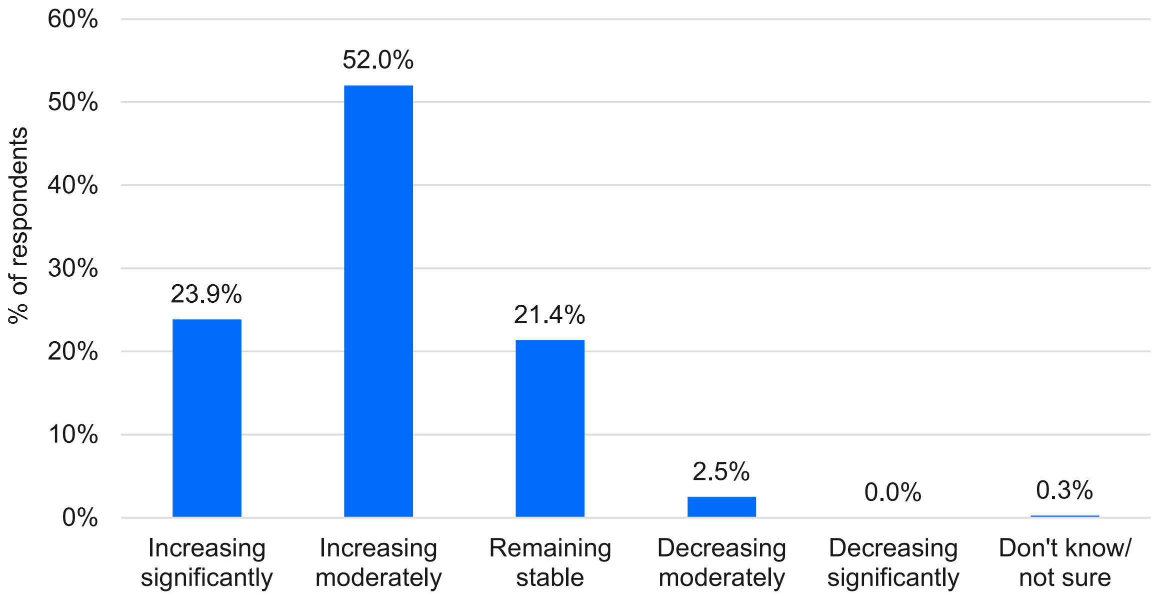
Beyond training, HDD's role extends into inference and ongoing operations. AI-generated content, including synthetic outputs, logs, inference results, and versioned data sets, accumulates continuously in the capacity tier, creating persistent and growing storage demand that favors HDD economics. The growing requirement for faster archive retrieval to support RAG pipelines means that HDD-based archives are being asked to play a more active role in live AI workflows than their traditional classification as cold storage would suggest. Survey results reinforce this shift: more than 75% of organizations report increasing the volume of archived cold-tier data being brought back online to support AI workloads, with nearly a quarter reporting that increase as significant (see Figure 33).

> "Archive data is being reactivated. Hospitals are loading HDD archives into JBOD racks, and cold data is being warmed up to serve AI training and inference workloads."
>
> — IT decision-maker, distribution and systems integration

FIGURE 33

Cold data reactivation for AI workloads

Q. How is your organization changing the volume of archived (cold-tier) data being brought back online or accessed more frequently to support AI workloads?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey,* 2026

Workload-specific fit: When and why HDDs are preferred

HDD is the preferred medium wherever capacity, cost, and sustained sequential throughput matter more than sub-millisecond latency, and that description encompasses the majority of AI data by volume across the full life cycle. The tier distribution data in this study shows that warm, cool, and cold storage together account for nearly 75% of total stored data, and more than 60% of data lake volume is cold or infrequently accessed. These are not edge cases or niche workloads. They represent the bulk of what organizations actually store, and they are the tiers where the economics of high-capacity storage are most compelling.

The life-cycle phase data reinforces this picture from a different angle. Of the five AI life-cycle phases examined in this study, performance is the overriding concern in only one: model training. Across data ingestion, inference, AI-generated data storage, and long-term retention, reliability leads or ties with performance as the primary storage attribute. This means that for four out of five life-cycle phases, the attributes that drive storage selection are precisely those where HDD excels. The performance conversation that dominates much of the HDD versus SSD debate is, for the majority of AI workloads by volume, largely beside the point. What matters is capacity, reliability, cost per terabyte, and the ability to sustain throughput at scale over long retention horizons, and on all of those dimensions, HDD continues to deliver.

HDDs will continue to meet the majority of enterprise storage exabyte needs over the long term

The long-term outlook for HDD is not managed decline but sustained and expanding relevance. The forces driving storage demand in the AI era, compounding data growth, longer retention horizons, archive reactivation, and the proliferation of new AI-generated data types, are structural rather than cyclical, and they favor high-capacity, cost-efficient storage in ways that are unlikely to diminish over the long-term forecast period. At the same time, HDD technology is not standing still. Continued advances in areal density through innovations such as heat-assisted magnetic recording (HAMR) are extending capacity leadership while maintaining the cost-per-terabyte advantages that make HDD the default choice for the capacity tier. And WD's recent HDD innovations focusing on increasing HDD throughput, a key requirement of AI workloads, as well as bandwidth and power efficiency will help to further extend HDD competitiveness into the future and across most of the phases of the AI life cycle. Meanwhile, the economics of flash storage at exabyte scale remain prohibitive for the foreseeable future, and the price trajectory required for QLC SSD to close the gap with HDD for capacity-tier workloads is neither imminent nor assured. Taken together, these factors point to a storage landscape in which HDDs not only retain

> "There's still a stigma to SSDs being a far better-performing warm tier storage than HDDs ... there's no way we can swap that perception. But that's not where HDD needs to win."
>
> — IT decision-maker, distribution and systems integration

their relevance but see it reinforced by the very trends that some observers have suggested might erode it.

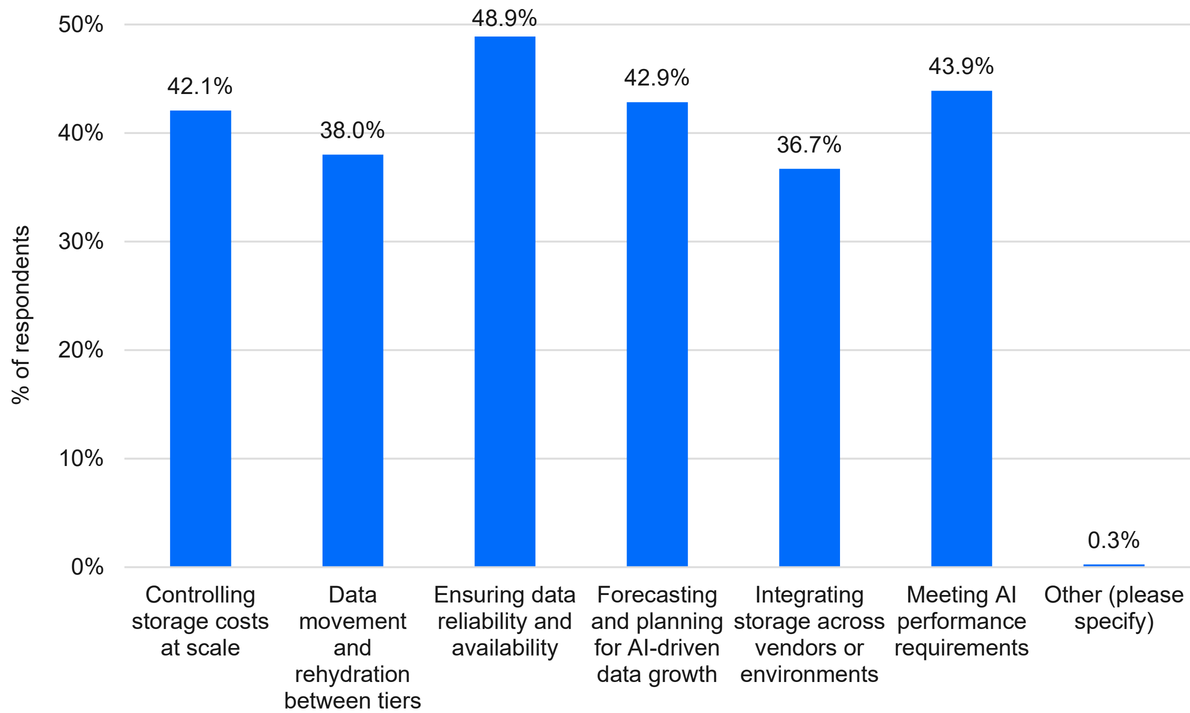
Strategies for reinforcing HDD value in hybrid storage architectures and tiering

Managing storage tiers for AI workloads is operationally demanding. Ensuring data reliability and availability tops the list of operational challenges, cited by nearly half of respondents, followed closely by meeting AI performance requirements, forecasting for AI-driven data growth, controlling costs at scale, and managing data movement between tiers. This list spans technology, economics, and operations simultaneously, reflecting the genuine complexity of multitier AI storage environments (see Figure 34).

FIGURE 34

Biggest operational challenges in AI storage

Q. What are the biggest operational challenges your organization faces in managing storage tiers for AI workloads?



n = 763 (unweighted base)

Source: IDC's *AI and Data Infrastructure Decision-Maker Survey*, 2026

Integrated HDD platform solutions are well positioned to address several of these challenges in combination, reducing the vendor complexity and integration overhead that organizations consistently cite as friction points. For the large segment of organizations that prioritize simplicity and cost equally, a platform that delivers on both dimensions addresses the core of what they are looking for, and that is where HDD vendors have a meaningful opportunity to differentiate.

Recommendations for enterprise customers

Enterprise customers can capitalize on the increasing value of data in the AI era by building storage strategies that are workload-aware, life-cycle-informed, and economically grounded. HDDs are a core enabler of that strategy, not as a legacy fallback but as the deliberate choice for the capacity tier at scale. IDC recommends that organizational leaders and IT decision-makers consider the following recommendations, detailed in the sections that follow.

Architect storage for the full AI data life cycle, not just the hot tier

Storage architecture decisions are too often made with hot-tier performance requirements in mind, while the majority of AI data by volume accumulates across warm, cool, and cold tiers. Design infrastructure that reflects the full life-cycle map, with tiered storage that matches media to workload across data ingestion, training, inference, and long-term retention. A hot-tier-centric architecture will systematically overinvest in the wrong medium for the majority of AI data.

Rethink cold data because its strategic value is increasing

Cold storage is no longer truly passive. AI adoption is extending retention horizons, archive reactivation is accelerating, and the growing use of RAG pipelines means that historically archived data is increasingly being called upon to support live AI workflows. Build retention policies that account for the future AI value of archived data, and ensure HDD-based archives are designed for retrieval performance and accessibility, not just low-cost storage.

Treat data retention as a strategic asset management decision, not an IT housekeeping task

The organizations best positioned to extract long-term value from AI are those that approach data retention with the same deliberateness they apply to other strategic assets. Establish clear governance frameworks that define what data is worth keeping, for how long, and at what storage tier. The cost of retaining data on appropriately tiered HDD infrastructure is low relative to the potential value of having that data available for future model training, analytics, or compliance purposes.

"Customers are beginning to treat data as core IP. Data that had no value two years ago is now a training asset — and nobody is throwing anything away."

— Senior storage channel leader, global technology distribution

Future outlook: The enduring role of HDDs in the AI era

The AI era is a data era, and the data era is a storage era. AI is fundamentally changing the economics of storage, and that shift is structural, not cyclical. As AI adoption deepens and matures, the structural conditions that favor HDD as the dominant capacity-tier storage medium are strengthening, not weakening.

Several forces will amplify this trajectory. AI adoption is expanding rapidly beyond the technology sector into healthcare, government, financial services, manufacturing, and beyond, each bringing with it large volumes of domain-specific data that must be ingested, retained, and made accessible for AI workloads. New verticals are arriving ahead of earlier adoption forecasts, and with them new categories of data from medical imaging repositories and digitized public records to industrial sensor archives and financial transaction histories that are well suited to high-capacity, cost-efficient HDD storage. At the same time, the growing use of RAG pipelines and AI inference is transforming previously passive archives into active infrastructure assets, extending the relevance of HDD-based storage deeper into the AI workflow than the conventional archive narrative would suggest.

The technology road map further supports this outlook. Continued advances in recording technologies, including HAMR and related innovations, will further extend the role of high-capacity storage in AI infrastructure, while maintaining the cost-per-terabyte advantages that make it the default choice for the capacity tier. The price trajectory required for flash storage to close the gap with HDD for capacity-tier workloads remains a long-horizon proposition, and the absolute economics of storing exabytes of AI data on flash infrastructure remain prohibitive for the foreseeable future.

The survey data that underpins this paper tells a consistent and convergent story. Organizations are storing more data, retaining it longer, assigning it higher strategic value, and prioritizing reliability alongside cost and operational simplicity across the majority of their AI workloads. HDD platforms are already the second most commonly deployed storage infrastructure for AI workloads among respondents, behind only cloud storage. And as organizations accumulate more direct experience managing AI storage at scale, the perception of HDDs as a legacy technology is likely to give way to a more accurate understanding of their role. The data in this study suggests that understanding is already forming. HDDs are not navigating disruption. They are being validated by it, and the conditions driving that validation are structural, compounding, and still in their early stages.

> "Compute is passé. Storage is the big problem we need to resolve in order to get AI moving efficiently."
>
> — Senior storage channel leader, global technology distribution

Appendix

Methodology

The findings presented in this paper are based on primary research conducted by IDC on behalf of WD. The research program comprised two components: a quantitative survey and a series of qualitative in-depth interviews.

The quantitative survey was conducted online using a computer-assisted web interviewing methodology in May 2026. A total of 763 qualified respondents completed the survey across 7 countries: the United States, Kingdom of Saudi Arabia, Germany, India, China, Japan, and South Korea. All respondents held director-level or above positions within their organizations and had direct involvement in AI infrastructure or data storage decisions. Participating organizations employed 500+ people. The survey was conducted on a double-blind basis, meaning WD's role as sponsor was not disclosed to respondents, and individual responses are not presented or attributed on an individual basis. Stats testing was conducted at the 95% confidence interval.

Respondents were drawn from a range of industries including technology, financial services, healthcare, manufacturing, retail, government, and telecommunications and from organizations spanning midmarket to large enterprise in terms of employee count. AI deployment maturity varied across the sample, from organizations in early stages of AI exploration to those with mature, production-scale AI workloads in operation.

The qualitative component consisted of in-depth interviews with two senior storage channel leaders with extensive experience in enterprise storage distribution and systems integration across multiple geographies and sectors, as well with leaders at a large cloud hyperscaler. All interviews were conducted on a named but not publicly attributed basis, and insights and perspectives from these conversations are referenced throughout this paper in aggregate or with role identifiers only, in accordance with IDC privacy and compliance policy.

IDC's Global DataSphere and Global StorageSphere forecasts are referenced in the Introduction section of this paper to provide broader market context for the primary research findings. These are independent IDC syndicated research programs and are not specific to this study or its sponsor.

About the IDC analysts

Ed Burns

Research Director, Semiconductors and Enabling Technologies

Edward Burns is research director within IDC's enterprise infrastructure global research domain and part of the storage and data management subdomain. He covers storage and flash technologies, with a focus on insights and market data for the worldwide hard drives and solid-state (flash) storage market. His core research coverage includes all market applications for HDDs and SSDs, as well as assessing demand for key components. In addition, Edward's research includes analyzing HDD/SSD technology trends, trends in competing storage technologies, and the market opportunity for online, nearline, and offline storage technologies across all markets worldwide.

Adam Wright

Senior Research Manager, Global DataSphere and Global StorageSphere

Adam Wright is a senior research manager within IDC's enterprise infrastructure global research domain and part of the storage and data management subdomain. He leads IDC's Global DataSphere and Global StorageSphere research. The Global DataSphere measures the total volume of data created worldwide each year. The Global StorageSphere measures the installed base of storage capacity worldwide and the amount of data stored each year.

Message from the sponsor

WD, also known as Western Digital, builds the storage infrastructure that powers certainty in the AI-driven data economy. At the forefront of innovation, WD partners with the world's leading hyperscalers, cloud service providers, and enterprises to enable reliable storage solutions that are proven and trusted at scale. Driven by a culture of innovation and execution, WD helps customers store, protect, and use the world's data with confidence. Follow WD on LinkedIn and learn more at www.wd.com.

IDC Custom Solutions

IDC Research, Inc.
One Beacon Street, Suite 33100, Boston, MA 02108, USA
T +1 508 872 8200

blogs.idc.com | LinkedIn @IDC | www.idc.com