



New Research Reveals the Reshaping of AI Infrastructure Economics as Data Persists and Compounds

September 9, 2026

WD-sponsored global study finds AI is driving rapid data growth, longer retention, and renewed demand for historical data, reshaping infrastructure economics at scale

SAN JOSE, Calif.--(BUSINESS WIRE)--Sep. 9, 2026-- The AI infrastructure conversation has largely been defined by compute. But new global research from IDC reveals another critical factor organizations must address as they scale AI: data that persists and compounds long after individual compute cycles are complete.

The IDC White Paper, [*Built for Scale: The Enduring Role of HDDs in the AI Era*](#), sponsored by WD, finds that among surveyed organizations, AI is creating a structural expansion in data storage requirements. Organizations are generating more data, assigning greater value to it, retaining it longer, and increasingly bringing historical data back online for new AI workloads.

The result is a compounding data cycle: AI requires data and increases the potential value of data already stored. While compute requirements vary by workload and investment cycle, the data created by AI persists and accumulates, making storage capacity, accessibility and economics increasingly important considerations in AI infrastructure design.

Among the Key Findings from the White Paper:

AI is Creating a Compounding Data Cycle

- 94.7% of surveyed organizations are storing more data because of AI and generative AI adoption over the past 12 months.
- 61% experienced data growth of 25% or more over the past year due to AI, while 74% expect data volumes to grow 25% or more over the next three years.
- 85.4% reported growth in data lake volumes over the past 12 months.
- 59.4% identified AI-generated data, including synthetic data, inference outputs and model logs, as the leading driver of data lake growth.

Data is Persisting Longer, Becoming More Active and Valuable

- Nearly 95% say the value of their organization's data has increased as a result of AI and GenAI adoption.
- 74.3% say AI and GenAI have caused them to retain data longer.
- 75.9% report bringing increasing volumes of archived cold-tier data back online to support AI workloads.
- 96% anticipate needing faster archive retrieval to support AI inference and retrieval-augmented generation (RAG) applications.

AI Infrastructure Must Be Designed for the Full Data Lifecycle

- For the surveyed organizations, 74.6% of enterprise data resides in warm, cool and cold storage tiers.
- More than 60% of data lake volume consists of cold or infrequently accessed data.
- 98.2% consider total cost of ownership per terabyte important or very important when making storage decisions.

"For the last few years, the AI infrastructure conversation has centered on compute. But AI runs on data," said Irving Tan, CEO, WD. "Organizations are generating more data, keeping it longer, and finding new ways to create value from the information they already have. Compute requirements will evolve over time, but the need to store, manage and access data at scale is only growing. That foundation will play a critical role in determining how far AI can go."

From Compute Cycles to a Persistent Data Lifecycle

AI workloads create data throughout their lifecycle, from training datasets and model checkpoints to inference outputs, logs and synthetic data. Unlike the compute cycles that process it, much of that data remains after the workload is complete and can become input to future AI applications. The findings point to a structural expansion of storage requirements as organizations retain more data for longer periods and find new uses for information they already hold.

AI Is Bringing Historical Data Back Into Play

The research also shows that the traditional boundaries between active and archived data are changing. As historical data becomes an increasingly valuable input for AI workloads, information that once sat dormant is reentering the active data lifecycle, changing expectations for how data is retained, managed and made available when needed.

AI Infrastructure Is a System, Not a Single Storage Tier

These shifts reinforce the need to architect storage across the full AI data lifecycle. As data volumes grow, retention periods lengthen and historical data becomes active again, organizations increasingly need to balance performance, capacity, accessibility and economics according to workload requirements. At AI scale, where data lives — and what it costs to store and access it — becomes an increasingly important architectural consideration.

The ability to manage data economically at scale is no longer a secondary infrastructure consideration. It is becoming a critical factor in AI success.

The full IDC White Paper, *Built for Scale: The Enduring Role of HDDs in the AI Era*, sponsored by WD, is available at <https://www.westerndigital.com/resources/white-paper/idc-hdds-ai-era>.

Research Methodology

The IDC White Paper is based on a quantitative survey of 763 IT and business decision-makers at the 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 three senior storage industry leaders to provide additional perspective on deployment realities and evolving customer priorities.

Source: *IDC White Paper, Sponsored by Western Digital Corporation, Built for Scale: The Enduring Role of HDDs in the AI Era, Doc. #US54789026, September 2026.*

About WD

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.

©2026 Western Digital Corporation or its affiliates. All rights reserved.

WD, the WD design, and Western Digital are registered trademarks or trademarks of Western Digital Corporation or its affiliates in the US and/or other countries. All other marks are the property of their respective owners.

View source version on businesswire.com: <https://www.businesswire.com/news/home/20260909414254/en/>

Media Contact

WD.Mediainquiries@wdc.com

Source: Western Digital Corporation