

Artificial Intelligence

AI is changing infrastructure planning — here's why

As AI workloads diversify, data requirements are becoming the primary driver of infrastructure decisions

Data is becoming the main driver of infrastructure decisions as AI workloads diversify

Infrastructure planning has long been built on stability as organizations selected platforms, qualified technologies, standardized architectures, and then scaled them consistently and efficiently.

That's changing as more AI initiatives move beyond pilots and into production.

Teams now support a growing variety of workloads, deployment environments, and business requirements while also meeting governance, compliance, and sovereignty requirements. Some AI applications run in public clouds, while others remain on-premises or at the edge.

Meanwhile, the accelerating pace of technology is creating more model choices and deployment options than ever. Infrastructure decisions are becoming less about identifying a single standard architecture and more about maintaining the flexibility to adapt.

Models are becoming more interchangeable within AI workflows. Organizations can route different tasks to different models based on cost, performance, or business requirements. Capable AI models are appearing in a broader range of environments, from hyperscale infrastructure to local devices. Infrastructure providers are also planning for more modular and replaceable compute architectures.

This all points to a broader shift: AI infrastructure is becoming more dynamic, and infrastructure decisions are becoming increasingly workload-specific. In many cases, data is the reason why.

The changing value of data

AI is creating new opportunities to reuse data across systems, workflows, and applications. Those applications have different requirements for where data resides, how it is governed, who can access it, how often it moves, and how long it needs to be retained. Privacy, sovereignty, latency, economics, and operational considerations also influence these architecture decisions.

This means infrastructure is increasingly shaped by the requirements of the data rather than a one-size-fits-all approach to deployment. Historically, organizations collected data to answer known business questions. Analytics helped forecast future outcomes, but the questions themselves were usually understood in advance.

With AI, customer interactions, operational history, governance records, organizational knowledge, and business context can be reused across multiple systems, workflows, and applications.

Because new deployments generate new data that can inform future models, workflows, and decisions, retained data becomes an input into future value creation.

When data outlasts infrastructure

In my industry conversations, it has become increasingly clear that data persists long after individual compute cycles end. Models, infrastructure, and applications may change, but the data remains.

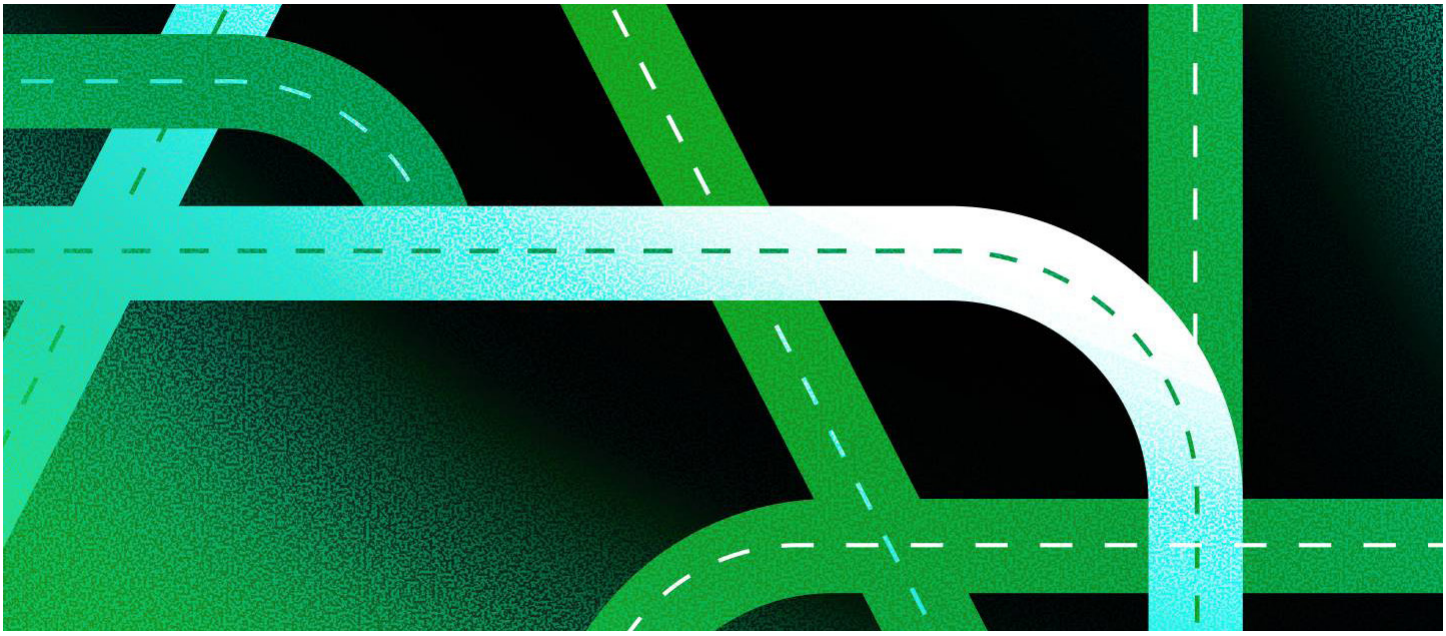
As a result, many organizations are finding value in combining commercially available models with proprietary data. When that happens, the ability to retain, govern, access, and reuse data grows in importance. Because data often outlasts the infrastructure used to process it, organizations are placing greater emphasis on architectures that can adapt as technologies evolve.

They want the ability to adopt new technologies and qualify infrastructure faster. They want architectures that allow them to respond to changing requirements without rebuilding everything from the ground up.

I'm also seeing more interest in adaptable infrastructure models that reduce rigid dependencies between technology layers. This flexibility is meant to create enough optionality to respond to changing workloads, technologies, and business priorities.

As AI becomes more embedded in business operations, infrastructure decisions will increasingly be shaped by data: where it resides, how it is governed, how long it remains valuable, and how effectively it can be reused across a growing range of applications and workloads.

Building flexible infrastructure around the long-term value of data will be a defining advantage in the future of AI.



© 2026 Seagate Technology LLC. All rights reserved. Seagate, Seagate Technology, Exos, and the Spiral logo are registered trademarks of Seagate Technology LLC in the United States and/or other countries. All other trademarks or registered trademarks are the property of their respective owners. Seagate reserves the right to change, without notice, product offerings or specifications.