VAST DATA FEDERAL

Why federal AI requires a new data foundation

The promise of artificial intelligence in the federal government is being stifled by a relic of the past: legacy data architecture

Randy Hayes | VAST Data Federal

For decades, agencies have been forced to manage data through a complex hierarchy of storage tiers, juggling multiple vendors for flash, hard drives, and tape to balance cost and performance. However, in the era of agentic AI, this fragmented approach is no longer just inefficient, it is a critical failure point.

Data is the lifeblood of decision-making. Yet for most agencies, that data is scattered across a chaotic landscape of legacy silos split between on-premises data centers, various cloud providers and disconnected edge environments. This fragmentation was born from decades of tactical purchasing, but in the age of AI, it has become a strategic liability.

The mission impact of data fragmentation

Although the industry once focused on "tiering" to manage costs, the real cost of this fragmented architecture is now measured in lost insights. AI is a reflection of the data it consumes. When information is trapped in silos, agencies are often forced to train models on only a slice of their total data corpus. This selective ingestion creates a "sampling bias" where AI models make decisions based on incomplete context, a dangerous proposition for applications ranging from predictive maintenance to national security.

To bridge these fragments, federal IT teams are forced to stitch together independent systems for storage, databases and compute. This creates fragile, expensive ETL pipelines that introduce significant latency. By the time the data is ready, the window for real-time action has often closed.

VAST Data Federal is replacing the fragmented mess with a single, high-performance data fabric that eliminates the trade-offs between accessibility, scale and mission speed. We have spent the last several years evolving that foundation into a full-stack AI operating system.

Because machines do not perceive the world through files and folders, we integrated an exascale vector database that converts complex information into vectors that AI can read instantly. We expanded this ecosystem with the DataSpace, creating a single namespace that spans the edge, the core data center and the cloud. To automate mission-critical workflows, our DataEngine allows agencies to run "triggers" to create instant actions and workflows. Furthermore, as federal employees begin to work alongside AI agents, our AgentEngine provides the framework to spawn and manage these autonomous entities effectively.

Solving the inference crisis: The KV cache challenge

As we work with agencies to build "AI factories," we are tackling a new, more granular bottleneck: high-performance inference. The limiting factor in AI performance is no longer just GPU compute; it is the system's ability to move and manage key-value (KV) cache efficiently. As AI models generate output, they often repeat redundant calculations, and although KV caching speeds this up by storing previous steps, it typically leads to redundant data piling up in expensive, isolated GPU memory.

VAST eliminates these redundant prefill computations. By allowing the KV cache to be persistent, shareable and scalable across GPUs and nodes—rather than trapped in local memory—we significantly reduce operational costs and increase throughput. Agencies no longer need to reprocess the same input tokens in retrieval-augmented

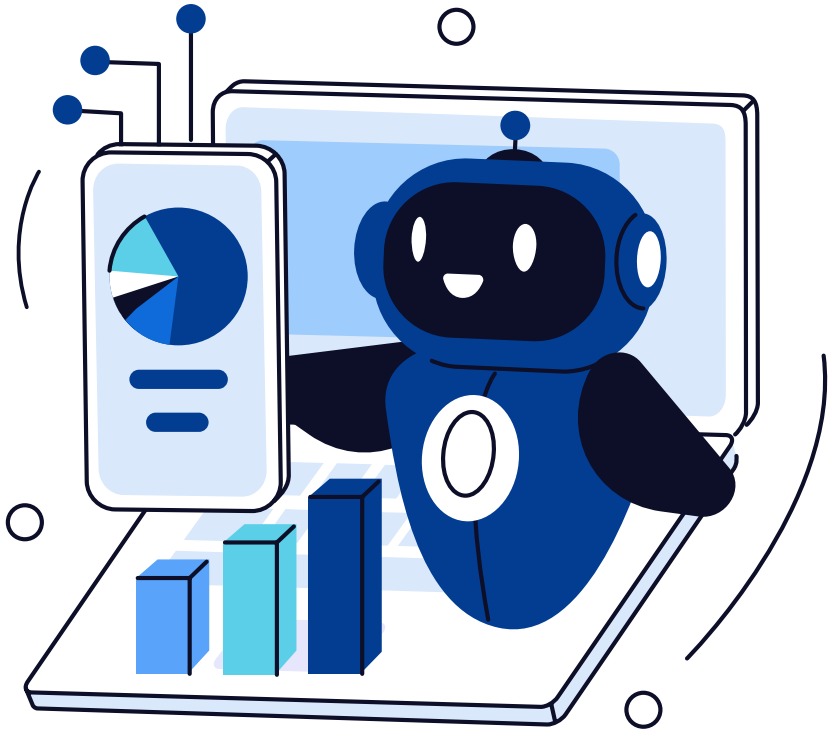
> "When information is trapped in silos, agencies are often forced to train AI models on only a slice of their total data corpus."

generation applications. They can compute it once, store it in the KV cache and retrieve it instantly.

Because VAST Federal is a software-first company, our platform is designed to be infrastructure agnostic. It runs on an agency's existing hardware or in the cloud, supporting legacy storage systems, servers and databases without requiring a "rip and replace" strategy.

The goal is simple but profound: to help agencies move past the era of managing tiers and into the era of managing intelligence, ensuring they get their AI infrastructure right the first time. ■

***Randy Hayes** is vice president of public sector at VAST Data Federal.*



ADOBE STOCK//IRINA STRELNIKOVA