Meeting the Challenge of Volume, Velocity, Variety and Veracity

The classic "Four V's" of big data are profoundly demonstrated in AI/ML processing systems.

What are AI and ML?

Artificial intelligence (AI) is a branch of computer science concerned with creating self-learning systems that can perform tasks that normally require human intelligence. From self-driving cars to predicting the future, AI is revolutionizing the ways in which we can use data to shape our world.

Seth Kindley
Principal Data Architect for US Civilian Federal agencies, Public Sector, and the Western United States
Pure Storage

Nick Psaki
Principal Technologist in the Office of the CTO for North American Public Sector
Pure Storage

Machine learning (ML) is a subfield within AI that deals with computer algorithms that can improve themselves via training data without explicit programming. It's widely considered the most promising path for achieving true human-like artificial intelligence.

No matter the specificity of the workflow, Pure Storage® FlashBlade® can serve the needs of the most demanding pipelines and deliver insight at unrivaled speeds.

How does Pure Storage help?

AI and ML are transforming organizations everywhere. However, data doesn't analyze itself. Legacy systems often can't support the massive data pipelines required. Pure delivers:

- A massively parallel platform.
- Ultra-fast, all-flash performance for thousands of processes, to billions of objects and files.
- Faster insights and multiple 'runs' to be performed over a shorter duration.

Don't take our word for it, look at what Gartner has to say...

The Challenge of Volume, Velocity, Variety and Veracity

The classic "Four V's" of big data are profoundly demonstrated in AI/ML processing systems.

The unique architecture of FlashBlade delivers unprecedented scalability of I/O service to address all four dimensions of data access, resulting in faster training runs, faster time to deployment, and operations with ever larger data sets. Modern data pipelines leverage data services at a rate once the realm of high-performance computing.

The massive scalability of GPU compute allows thousands of processes to work on data sets simultaneously, creating tremendous stress on the data service infrastructure. Traditional storage devices are built on serial I/O subsystems. To truly leverage the potential of modern analytic pipelines and machine learning, the supporting data systems need to provide massive parallelism and concurrency. Pure Storage delivers this with the FlashBlade platform. ■

Learn more:
- [High-Performance AI for Federal Agencies](#)
- [Data-Driven Government](#)
- [Smarter Infrastructure for AI](#)

About the Authors

Seth Kindley

Mr. Kindley serves as Principal Data Architect for US Civilian Federal agencies, Public Sector, and the Western United States. He is a technical advisor to: engineering, product management, sales, marketing, and presales leadership in the field across the Continental United States.

Nick Psaki

Mr. Psaki is Pure Storage's Principal Technologist in the Office of the CTO for North American Public Sector, providing deep technical knowledge of data service architectures that enable business and technological transformation for government and large enterprises. Nick has over 25 years' experience in designing, developing, deploying, and operating information systems for data analysis, sensor integration, and large-scale server virtualization.

High-Performance AI/ML for Federal Agencies

BEST AI SOLUTION FOR BIG DATA

Accelerate mission success with a modern AI infrastructure

Artificial Intelligence (AI) is *transforming government* everywhere and we're already seeing its potential to improve national security, safety, health, and citizen services, and to drive better, more informed, data-driven decision-making. For data-rich agencies, shared data is infinitely more valuable than stored data. But data doesn't analyze itself and legacy systems often can't support the massive AI data pipelines required for modern analytics. AI requires completely new infrastructure, and the complexities of legacy solutions create roadblocks that prevent agencies from tapping into the full potential of the new era of intelligence.

Pure Storage® accelerates *modern analytics and AI* workloads, delivering a single, scale-out storage platform that drives faster time to insight for data-driven government.

Unlock Unstructured Data's Value with Machine Learning

AI workloads process massive amounts of data from structured and unstructured sources. Your unstructured data could be one of your greatest assets—but it will take machine learning (ML) to get to the real value hidden inside. To unlock this value through machine learning, you first need the right tools and that includes your storage platform. With Pure *FlashBlade®*, you get a unified fast file and object storage platform that serves as an all-in-one, high-performance, scale-out solution for unstructured data.

A Leader in Innovation

Pure is recognized as the *"Best AI Solution for Big Data"* by the AI Breakthrough Awards. Pure offers simple and flexible AI-first infrastructure with the ability to flexibly adapt to demands for AI and to launch agencies into a new era by enabling faster innovation and a competitive edge for AI and ML initiatives.

Find out more

PURESTORAGE®