NetApp

U.S. Public Sector

NETAPP AI CAPABILITIES
FOR THE U.S. PUBLIC SECTOR

Overcome AI data challenges to improve public health and safety, enable smart policy making, gain decision advantage, and streamline bureaucratic processes.

The challenge

Whether hunting cyberthreats, operating military drones, tracking infectious diseases, or monitoring bacteria in water supplies, governments around the globe are fully embracing AI to do more with their mountains of data.

But many approaches to AI involve architectures that silo analytics, training, and inference workloads. Effective AI requires secure, seamless data movement throughout the pipeline, from ingest and data preparation all the way to analysis and tiering. Silos create unnecessary complexity, increase costs, and make it hard to scale. When a siloed infrastructure prevents data from flowing quickly and freely, AI and deep learning only scratch the surface.

Regardless of the task, government organizations need an IT infrastructure that provides high performance, easy scalability, and protection from ever-growing cyberthreats. The infrastructure must enable the organization to move faster, smarter, and more securely.

Get your data AI-ready with NetApp

For decades, NetApp has set the standard for hybrid cloud storage of unstructured data—the lifeblood of AI. Now, we're empowering government organizations to make any mission possible by using their data to unlock value from AI—enabling leaders to make informed decisions quickly and streamlining administrative processes.

NetApp® solutions get your data AI-ready with the most secure and high-performing storage on the planet. With NetApp, government agencies can get the right data and apps to the right place at the right time to meet their mission objectives. We lead the industry in high-performance, massively scalable, unified data storage that can power every segment of the hybrid AI data pipeline. Equally important, NetApp technology provides the simplicity, security, and governance required to mitigate risk so that you can confidently deploy safe, responsible AI.

Accelerate the entire AI pipeline

Eighty percent of an artificial intelligence/machine learning (AI/ML) project involves data management—from data collection and triage to data preparation and normalization and finally to algorithm training. A data management platform that aids data scientists throughout the machine learning operations (MLOps) workflow is a key component for a successful AI/ML solution.

NetApp high-performance solutions scale performance and capacity separately, so organizations get the speed and space they need to feed GPUs during model training and to optimize GPU utilization. Our solutions are built to handle massive data workloads without compromising on speed or latency.

It's also important to keep data moving and keep data scientists working instead of waiting for data to process. To keep your data flowing freely through the pipeline, NetApp solutions eliminate silos by consolidating AI workloads and existing enterprise apps. Government organizations can drive innovation and deliver faster results with the high-performance capabilities of our AI solutions.

To help you get started quickly, NetApp partners with NVIDIA to provide high-performance, specially engineered hardware that form the foundation of SuperPOD and BasePOD solutions. NetApp offers the only storage solutions that connect these pods to the rest of the enterprise and all major hyperscalers—simplifying the flow of data through the MLOps pipeline and to the GPUs.

With the NetApp DataOps Toolkit, data scientists can manage data directly within Jupyter Notebooks, an environment they're familiar with. They can create backups of entire datasets, restore previous data environments, and copy terabytes of data in seconds—so data flows more quickly, speeding everything from MLOps to warfighting. The ability to create thin clones of Jupyter Workspaces accelerates time to production and allows data scientists to innovate experimentally without space constraints.

Seamlessly mobilize all your data

A NetApp intelligent data infrastructure with AI-optimized data management enables you to instantly classify, access, move, and serve data natively across your hybrid multicloud environment. Data moves agnostically between edge, core data center, and hyperscaler cloud environments with ease—which is paramount to a successful AI/ML project.

During a project's lifecycle, data can be generated at the edge, and then moved to an on-premises data center for initial triage, to the cloud for algorithm development and testing, and back to an on-premises environment for production-level training. NetApp FlexCache® technology enables on-premises data to be cached in the cloud, where that data can be consumed by a cloud-based generative

KEY BENEFITS

Accelerate the entire AI pipeline

- Power all your AI use cases—and optimize GPU utilization—with high-performance, highly scalable NetApp solutions.

- Eliminate silos by consolidating AI workloads and existing enterprise apps—without trade-offs.

Seamlessly mobilize all your data

- Take advantage of intelligent data infrastructure with AI-optimized data management to instantly classify, access, move, and serve data natively across your hybrid multicloud environment.

- Streamline operations with AIOps-driven automation and unified, multicloud control delivered by BlueXP.

Get unmatched data security

- Deploy and protect AI workloads with the only hardened enterprise storage validated to store top-secret data.

- Keep AI models and data secure with real-time AI/ML-based ransomware protection and guaranteed data recovery.

- Enable safe, responsible AI and meet data privacy and governance requirements.

AI service. If internal data can't be exposed to the cloud because of regulations or governance concerns, data can be replicated or cached to an on-premises high-performance AI compute platform instead. Simplifying data movement and eliminating data silos helps keep data locality from becoming an MLOps bottleneck.

NetApp provides data management services as virtual instances for the edge, with engineered hardware solutions for on-premises requirements and first- or third-party solutions in every major hyperscaler cloud, including secure Department of Defense (DoD) and Intelligence Community regions such as Commercial Cloud Services (C2S), Commercial Cloud Enterprise (C2E), and Joint Warfighting Cloud Capability (JWCC).

Our intelligent data management solutions optimize storage, simplify data workloads, and enhance collaboration. AIOps-driven automation and unified multicloud control delivered by NetApp BlueXP™ services streamline operations. And with intelligent data tiering, you always have the right workload in the right environment at the right time.

Get unmatched data security

With any data, protection and security are top of mind, but security in an MLOps pipeline is heightened, and working with highly classified government data makes security even more critical. Any manipulation of the training set can cause unpredictable AI/ML outcomes.

NetApp provides certified security solutions in every deployment environment needed for a successful MLOps program—NIST FIPS 140, DoDIN APL, and CSfC. We offer the only hardened enterprise storage validated to store top-secret data. Besides being CSfC certified, our solutions feature quantum-resistant encryption.

Not only is the hardware and NetApp ONTAP® data-management software certified, but NetApp also provides extended features within the software to help secure your environment, including:

- Multi-admin verification
- Creation of write once, read many (WORM) volumes
- Data classification
- Immutable backups
- Built-in anti-ransomware protection

The latest NFS version, NFS 4.2, supports a new feature called extended attributes (xattrs). This feature allows extensible data tagging with NFS, like the tagging done with objects. With user-defined tags, you can label data to automate the management and movement of data in an AI workflow. You can also segment data based on user-defined authorization tags when you use public-key infrastructure (PKI) services. These tags can also be used to provide dataset-to-model traceability. When a model starts to underperform, display bias, or drift, you can use xattr tags to investigate the dataset used to train the model. NetApp is the only data management solution that currently supports NFS extended attributes.

NetApp offers the most secure storage on the planet—no other storage solution provides equivalent security protection. We deliver unmatched data security with built-in versioning and traceability, built-in cybersecurity, and real-time, AI-powered ransomware protection with 99% detection accuracy and guaranteed recovery.

NETAPP AI SOLUTIONS IN ACTION

CURING CANCER

Yale New Haven Health built an AI environment on NetApp technology that enables them to securely access cutting-edge large language models (LLMs) to get data out of clinical notes and accelerate research.

By integrating our DataOps Toolkit with NVIDIA's NeMo Retriever, Yale researchers are enhancing clinical and research data management for LLM training and deployment. This advancement leads to automated data extraction that's more accurate and efficient, which is crucial in the quest to cure cancer.

ENHANCING EDUCATION AND RESEARCH

Supported by a $75 million investment from New York State, University at Albany has launched a supercomputing initiative called AI Plus. This initiative will significantly expand the AI supercomputing resources available in New York for teaching and research well beyond traditional STEM applications, including the arts and humanities.

Though it's early in their journey, they chose NetApp all-flash storage systems to deliver ultra-high performance to their NVIDIA DGX clusters. They selected NetApp for our ability to provide an enterprise-ready, multitenant environment on premises with the potential of adding cloud tools in the future. Having private or "near-cloud" solutions will play a larger role as Albany hires new faculty and expands its DoD- and federal civilian-funded customer base.

PREDICTING AND PREVENTING WILDFIRES

Lockheed Martin has built an AI Integrations Lab: an internal ecosystem for developing AI solutions at scale for civilian and military applications. The entire lab runs largely on NetApp all-flash storage. They use NetApp Keystone® services for the flexibility to grow and innovate quickly without large capex. With NVIDIA Omniverse Nucleus, the lab can use AI and ML to ingest and make sense of all this ongoing data collection in real time, so researchers can collaborate and data can be shared across tools.

One example is wildfire prevention and prediction. Lockheed Martin's AI systems can use information about the current state of a fire and the local environment to predict its future behavior in minutes. This information can help deliver critical intelligence to assist firefighters in making faster, more accurate decisions.

Achieve AI mission success with a partner you can trust

NetApp has a proven track record of serving the U.S. public sector with reliable and secure AI solutions that remove bottlenecks at the edge, at the core, and in the cloud to enable more efficient data collection, faster AI workloads, and smoother cloud integration. Our unified data management solutions support seamless, cost-effective data movement across hybrid multicloud environments.

NetApp core AI capabilities include:

- **Storage-efficient copies.** Create complete, writable copies of datasets in seconds, regardless of size. Administrators can create unique repositories for teams or individual users without requiring 5 or 10 times the storage capacity.

- **Complete data protection.** Make hundreds of point-in-time backups of data volumes with NetApp Snapshot™ technology. Each Snapshot copy is an immutable set of references that can be mounted for read-only use or replicated to other systems or locations for long-term archiving or use by remote teams.

- **Simplified data management.** Use the DataOps Toolkit to access a set of open-source Python libraries that data scientists can load into their Jupyter Notebooks, allowing them to manage data, make backups, and manage versioning—all within their own Jupyter environment.

- **Integrated data and model management.** Go beyond open-source tools. NetApp ONTAP AI integrates with third-party toolsets and tool management solutions, so data scientists can use the tools that make sense for each workload. This integration also makes it easier to track model drift/traceability.

A unified data storage environment allows us to focus on the development of our AI-powered wildfire prediction and suppression system with more efficient and tactical decision-making.

Will Karavites, Chief AI/ML & Cloud Architect, Lockheed Martin

Get started today

Only NetApp can help you build an intelligent data infrastructure that satisfies the needs of complex, high-performance AI data pipelines across your entire hybrid multicloud environment—and enables safe, responsible AI that meets data privacy and governance requirements.

Learn more about NetApp AI solutions for government and public sector.

Contact Us

NetApp