Get productive with AI

NVIDIA AI Computing by HPE

AI is everywhere, disrupting every industry and creating limitless opportunities. HPE and NVIDIA® are collaborating to deliver co-developed solutions to help you accelerate the adoption of generative AI.

Hewlett Packard Enterprise | **NVIDIA**

Overcoming the challenges with AI adoption

Complex, fragmented toolsets inhibit AI and IT teams from innovating quickly

Enterprise architecture and workflow are complex

Data safeguards

Issues controlling security, governance, and compliance across hybrid estates

How to start small, grow, and adapt as business needs evolve

Scaling and financial commitment

Management complexity

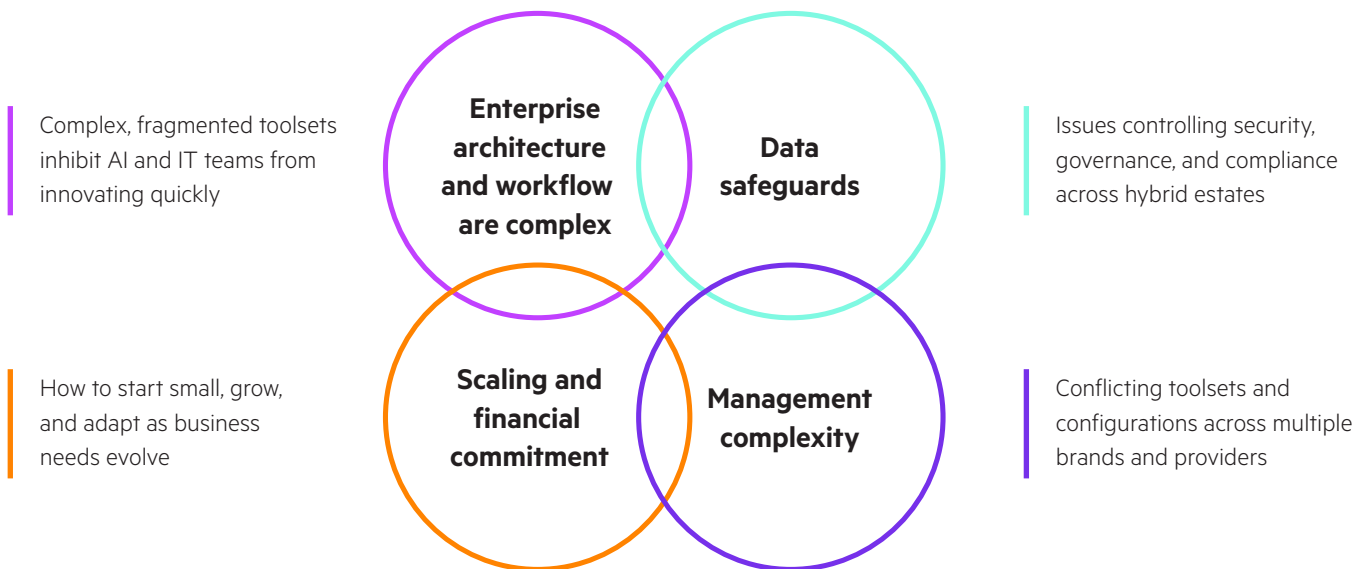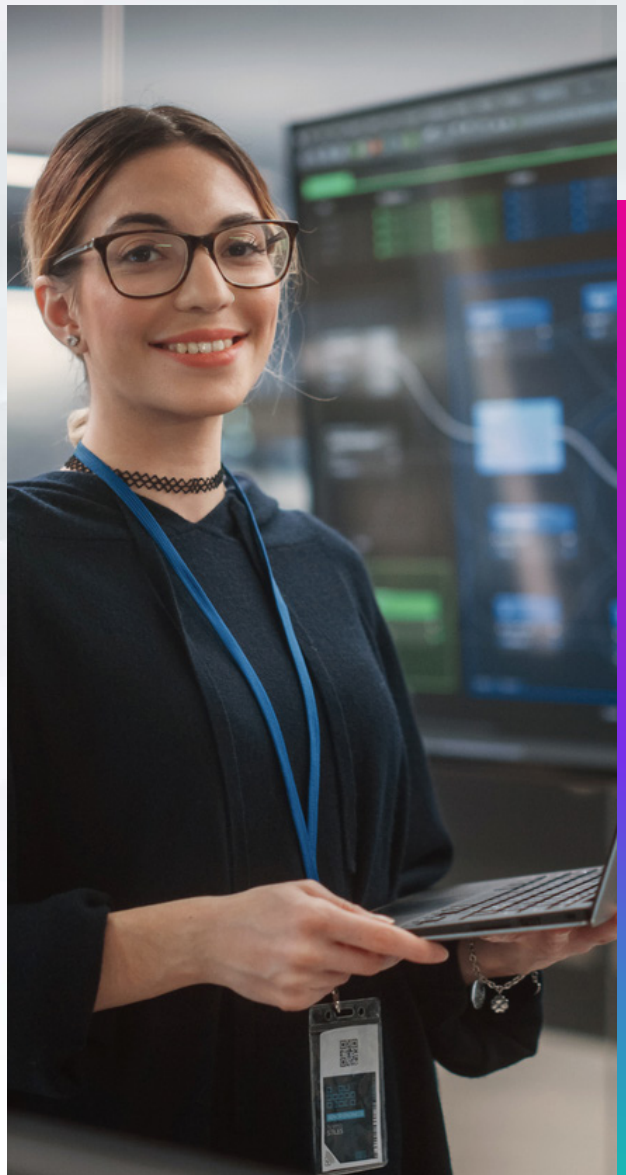Conflicting toolsets and configurations across multiple brands and providers

Figure 1. Overcoming the challenges with AI adoption

NVIDIA AI Computing by HPE can help

NVIDIA AI Computing by HPE is a portfolio of co-developed AI solutions and joint go-to-market integrations that enable you to accelerate the adoption of generative AI. The portfolio delivers:

☑ **People:** Work with experts and develop new skill sets

- HPE Services provides global AI expertise and capabilities throughout the entire AI lifecycle

- Partnerships with Global System Integrators to meet our joint customers' AI needs

☑ **Technology:** Co-developed, turnkey solutions

- HPE Private Cloud AI, a full-stack, turnkey private cloud for production generative AI built on NVIDIA AI computing, networking and software with HPE's AI storage, compute and the HPE GreenLake cloud, to accelerate time to value

☑ **Economics:** Start small and grow as your needs evolve. Flexible consumption as CapEx or OpEx and optimize AI costs long term

- Bring enterprise-grade control and full visibility into utilization and cost with HPE GreenLake Private Cloud

NVIDIA AI Computing by HPE

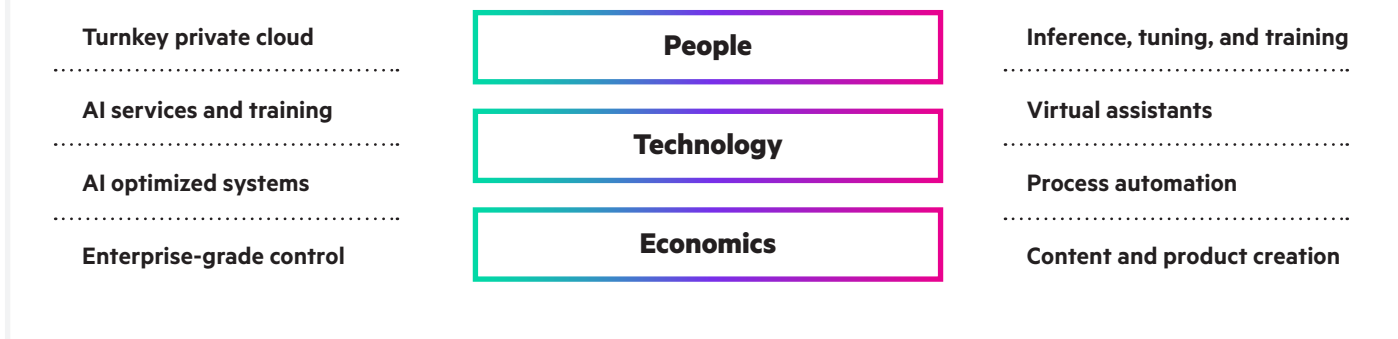Co-developed solutions to simplify enterprise AI

Turnkey private cloud		People		Inference, tuning, and training
AI services and training		Technology		Virtual assistants
AI optimized systems		Economics		Process automation
Enterprise-grade control				Content and product creation

Figure 2. NVIDIA AI Computing by HPE

Advance your organization with HPE Private Cloud AI

Part of the NVIDIA AI Computing by HPE portfolio, this first-of-its-kind solution enables enterprises of every size to gain an energy-efficient, fast and flexible path for sustainably developing and deploying generative AI applications.

HPE Private Cloud AI includes a self-service cloud experience with full lifecycle management and is available in four right-sized configurations to support a broad range of AI workloads and use cases.

Designed to support inference, RAG, fine-tuning and small model training use cases leveraging data across hybrid cloud environments, it can be deployed on-premises, colos, or edges, ensuring data is secure and under complete control of enterprise IT. HPE Private Cloud AI supports a wide array of use cases across multiple sectors, including virtual assistants for increased productivity, automation to streamline operations, and accelerate revenue growth through AI-driven product development and content creation.

Managed through HPE GreenLake cloud, it includes a powerful data science platform with the NVIDIA AI Enterprise software and NVIDIA NIM inference microservices to experiment and operationalize AI while keeping data and models secure. As enterprise AI needs expand and end users needs for new AI capabilities and applications increases, HPE Private Cloud AI can grow to meet the demand.

HPE Private Cloud AI provides:

- Instant AI productivity: Unique, private cloud experience with NVIDIA AI Computing accelerates time to value of enterprise-grade generative AI solutions.

- Unify access to all your data: Built-in HPE Ezmeral Data Fabric simplifies data management by seamlessly integrating, organizing, and governing enterprise data with one global namespace for seamless access to data, anywhere.

- Enterprise-grade confidence and control: Comprehensive, multi-layered controls protect sensitive data and models and maintain high performance, reliability, and utilization of AI infrastructure.

- Cloud experience that keeps data private: Managed experience with HPE GreenLake cloud, deployed on-premises and designed for hybrid, HPE Private Cloud AI is flexible and modular to expand and grow with AI demand.

HPE Private Cloud AI components
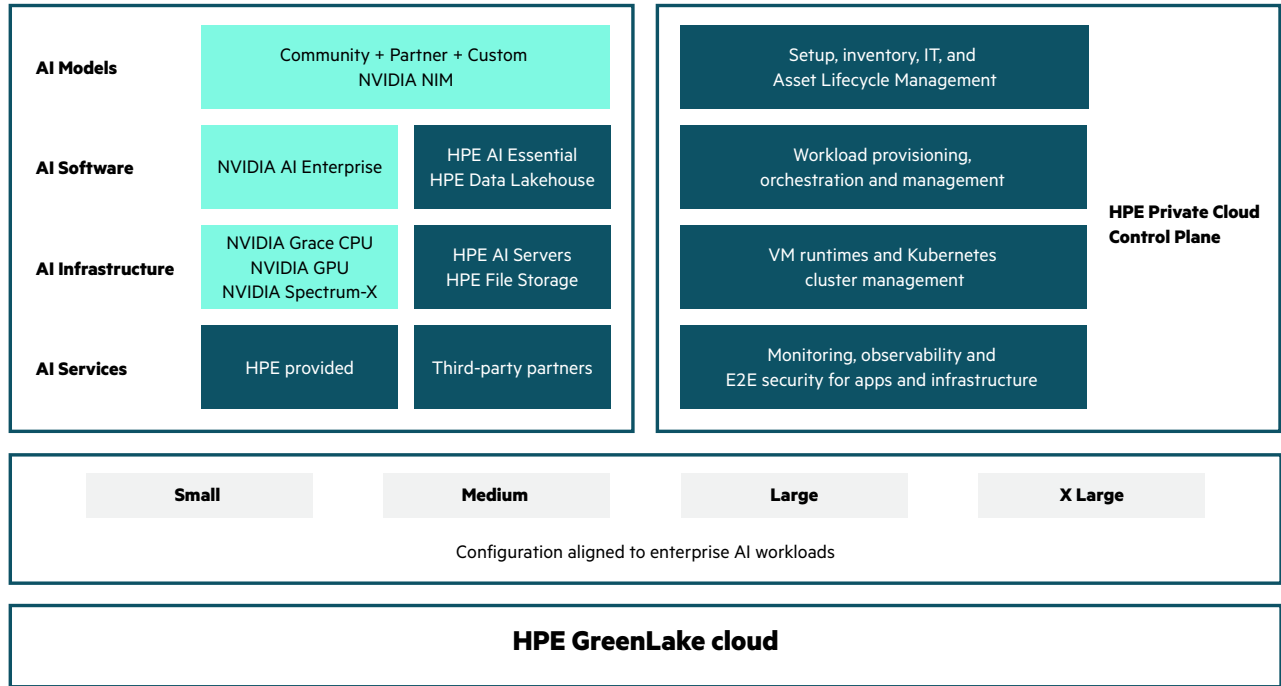
NVIDIA AI computing by HPE
HPE Private Cloud AI

			HPE Private Cloud Control Plane
AI Models	Community + Partner + Custom NVIDIA NIM		Setup, inventory, IT, and Asset Lifecycle Management
AI Software	NVIDIA AI Enterprise	HPE AI Essential HPE Data Lakehouse	Workload provisioning, orchestration and management
AI Infrastructure	NVIDIA Grace CPU NVIDIA GPU NVIDIA Spectrum-X	HPE AI Servers HPE File Storage	VM runtimes and Kubernetes cluster management
AI Services	HPE provided	Third-party partners	Monitoring, observability and E2E security for apps and infrastructure

Small	Medium	Large	X Large

Configuration aligned to enterprise AI workloads

HPE GreenLake cloud

Figure 3. HPE Private Cloud AI

Start unlocking AI value now

a50010571ENW, Rev. 1

2892571-1224

Hewlett Packard Enterprise