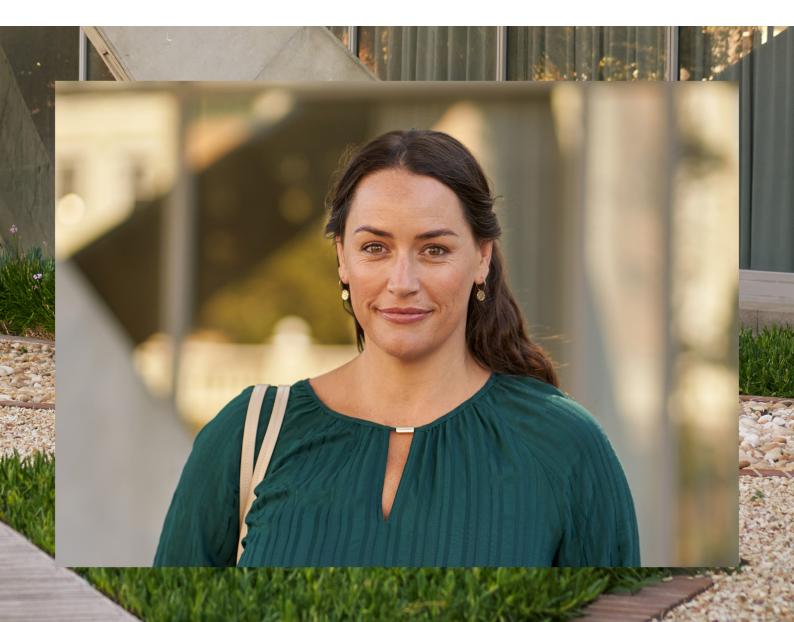# Generative AI solution for enterprise

Unlock AI capabilities that are shaping the future of business

How can GenAI transform your business?

Artificial intelligence (AI) is profoundly changing how business is done, and while enterprises understand the transformative potential of AI, implementing the rapidly evolving technology is a major challenge. Generative AI (GenAI) and the adoption of large language models and multimodal models mark a significant advancement in AI. The ability of machine systems to generate controlled outputs that are natively human-readable dramatically simplifies their use and accelerates game-changing value for your business.

With GenAI, the pace of change is breakneck. Foundation models can be customized with enterprise data to provide accurate and timely information to boost productivity, improve customer experience, and automate business processes. However, deploying GenAI into production is a complex task. Adapting foundation models to a specific business case requires a combination of fine-tuning the models with private data and/or retrieval augmented generation (RAG), which enables models to leverage data from external sources to deliver more accurate results. Accomplishing this requires adjusting your operating model and training your team to use AI technologies effectively.
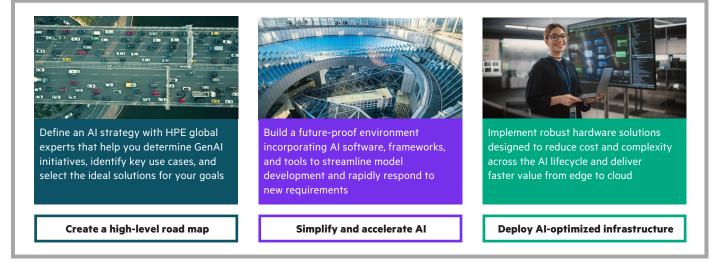
Pioneer the next era of AI with a trusted partner

GenAI is considered the ultimate data-driven hybrid cloud workload. It creates unique and unprecedented demands on IT and business — and introduces new requirements that most enterprises have never encountered before. What worked yesterday won't move you forward tomorrow.

Hewlett Packard Enterprise is a trusted partner for the next era of AI. We offer proven technology and expertise, built from decades of experience as the global leader in supercomputing. Our AI strategy is hybrid by design, recognizing that data can live anywhere and that inferencing increasingly must happen at the edge.

To drive your business transformation, HPE and NVIDIA® introduced an open, full-stack AI-native architecture that accelerates AI innovation at the edge, core, and cloud. These solutions bring HPE leadership in hybrid cloud, computing, and AI / machine learning (ML) with industry-leading NVIDIA GPUs, NVIDIA AI Enterprise software, and NVIDIA networking platforms to unlock the power of GenAI. The solution is preconfigured to help you innovate quickly and gain value from day 1.

Define an AI strategy with HPE global experts that help you determine GenAI initiatives, identify key use cases, and select the ideal solutions for your goals

Build a future-proof environment incorporating AI software, frameworks, and tools to streamline model development and rapidly respond to new requirements

Implement robust hardware solutions designed to reduce cost and complexity across the AI lifecycle and deliver faster value from edge to cloud

Create a high-level road map

Simplify and accelerate AI

Deploy AI-optimized infrastructure

Figure 1. The stages of AI innovation

Why GenAI solutions from HPE?

Each stage of the AI lifecycle has its own unique demands on data and IT. Training and tuning AI models at scale demands innovative software and powerful computing technologies to improve accuracy and speed results. To realize these benefits, enterprises are on a mission to balance sustainable solutions with exceptionally high performance. They must extend their cloud-native environment to include an AI-native approach if they hope to effectively utilize GenAI.

HPE and NVIDIA have introduced a unique full-stack proposition[1] that combines industry-leading AI software and infrastructure to drive impact at any scale. Our AI-native architecture is simple to deploy and manage, making it the right fit for any stage of your AI journey. This breakthrough solution enables you to adopt AI across training, tuning, and inferencing — from our proven expertise in building systems designed for extreme scale and performance, to implementing and integrating models into existing environments, all the way to high-throughput edge solutions.

HPE's enterprise computing solution for generative AI is the ideal entry point for enterprises of all sizes. Each component is tuned for edge or data center deployments, powered by an ultra-scalable architecture that features cutting-edge AI software, industry-leading compute, pretrained foundation models, and comprehensive services to help you quickly develop production applications. GPU-enabled applications enable faster AI outcomes and greater inferencing performance to maximize the impact of GenAI.

What to expect

HPE delivers a curated mix of software and hardware offerings that are purposefully engineered to accelerate the AI lifecycle and expand the capabilities of AI for future inferencing requirements. This technology evolution is powering common GenAI use cases (like conversational search, business process automation, and content creation).

HPE brings you the advantages of purpose-built AI architecture for today and the future:

• Data-first pipeline to manage public and proprietary data across multigen IT

• AI lifecycle management software to accelerate workflows for training, tuning, and inferencing

• Hybrid approach to run AI anywhere with advanced data protection

• Supercomputing DNA built into the entire portfolio to train the largest models

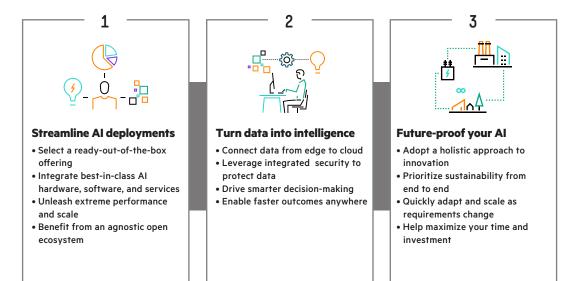• Open ecosystem for freedom of choice with no lock-in

1

Streamline AI deployments
• Select a ready-out-of-the-box offering
• Integrate best-in-class AI hardware, software, and services
• Unleash extreme performance and scale
• Benefit from an agnostic open ecosystem

2

Turn data into intelligence
• Connect data from edge to cloud
• Leverage integrated security to protect data
• Drive smarter decision-making
• Enable faster outcomes anywhere

3

Future-proof your AI
• Adopt a holistic approach to innovation
• Prioritize sustainability from end to end
• Quickly adapt and scale as requirements change
• Help maximize your time and investment

Figure 2. How to unlock the power of GenAI

[1] "Introducing an AI-native architecture for AI-driven transformation," HPE Newsroom, 2023

HPE's enterprise computing solution for Generative AI

HPE Machine Learning Development Environment

Build more accurate ML models faster with an out-of-the-box AI software solution that reduces complexity and cost

HPE Ezmeral Software

Create a future-proof SaaS foundation for AI that scales with your business, helping you manage AI apps across hybrid multicloud environments

NVIDIA AI Enterprise with NVIDIA NeMo™

Simplify and accelerate AI with a cloud-native software platform featuring frameworks like NVIDIA NeMo for custom model development

HPE ProLiant Gen11 servers with NVIDIA® L40S GPUs

Adopt easy-to-upgrade servers with extreme acceleration that support parallel processing and scale rapidly to meet AI needs from edge to cloud.

NVIDIA Spectrum-X with NVIDIA Spectrum-4 and NVIDIA® BlueField®-3

Create an optimized AI network with accelerated Ethernet and gain superior performance, security, and cost-savings for cloud computing platforms.

HPE AI Services

Simplify each stage of the AI journey with new AI and hybrid cloud services, consulting, workforce training, and deployment solutions.
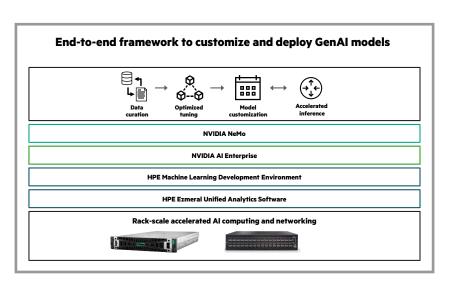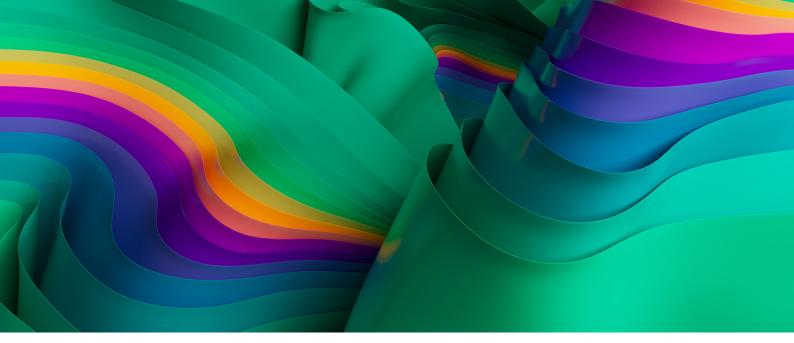

Figure 3. HPE's enterprise computing solution for generative

Figure 4. End-to-end framework to customize and deploy GenAI models

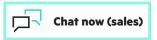# Discover the unlimited possibilities of AI

AI can turn questions into discovery, insights into action, and imagination into reality — but harnessing massive amounts of varied, distributed data for AI requires a new breed of architecture.

HPE is committed to defining and optimizing infrastructure for the next era of AI. We bring together market-leading hybrid cloud, computing, and AI capabilities to drive AI-powered transformation. Our full-stack AI-native solutions are designed to accelerate time to value. With purpose-built architecture, you possess the critical tools, technologies, and resources to develop high-quality AI models quickly and securely.

If you are unsure how to get started, we can help. HPE AI Services — Generative AI Discovery is a two-hour session conducted by HPE AI experts where you learn about best practices, strategic guidelines, and developing a high-level implementation plan.

Learn more at

HPE.com/HPE-enterprise-computing-solution-gen-ai

Chat now (sales)

Hewlett Packard Enterprise