

Revolutionize your business with the transformational potential of AI

HPE Private Cloud AI, together with HCLTech



Impacting every industry, every day

The market for artificial intelligence (AI) technology is expanding at an unprecedented rate, driven by advancements across the AI spectrum. According to IDC:

- Generative AI (GenAI) is the fastest-growing AI segment.
 - More than 80% of enterprises are experimenting with GenAI initiatives.¹
 - GenAI is growing at 2x the rate of AI overall.
 - GenAI will be a \$143 billion market by 2027.²
- GenAI will create huge value for successful adopters. By 2034, IDC forecasts that GenAI will add nearly \$10 trillion to global GDP.³
- AI training and inference need diverse computing resources. Training a large model often requires GPU-accelerated clusters with hundreds of petaflops of compute power over several weeks.
- Enterprises prefer a hybrid cloud approach to GenAI; 68% of enterprises see hybrid and multicloud as key to their GenAI strategy,⁴ and 50% plan to use their own or dedicated infrastructure, not public cloud.⁵



Releasing the power of AI

Artificial intelligence is shaping the future of business, but adopting and managing these technologies can be difficult. Enterprises need partners with deep AI expertise to release the power of AI and evolve with new requirements.

Even as the financial benefits of AI continue to grow, statistics reveal that only 14% of customers have fully realized their AI strategy, and that only one in 10 AI projects reach a live production environment.⁶ There could be many reasons for this disconnect:

- **Time to productivity:** Enterprises struggle to enable AI and IT operations teams to successfully transform from AI pilots to production and deliver business value from AI projects.
- **Complexity, security, and privacy:** AI environments are complex to configure and operate securely, with multiple elements needing to work together.
- **Ability to scale:** Many organizations want to start small with their AI environment and grow with their maturity and demand, so they need to deploy a future-proof solution.
- **Cost and compliance:** AI may be seen as costly—both financially and in the ability to comply with governance, regulation, and sustainability, creating risks for an organization's brand and reputation.

AI is profoundly changing how we live and work—affecting every industry sector. For example, AI is revolutionizing:

- **Financial services** by accelerating the review of complicated and sometimes subjective financial documents such as loan applications, while keeping money more secure with advanced fraud detection and prevention capabilities
- **Healthcare** by providing personalized treatments and speeding diagnosis; by offering powerful virtual assistants that can leverage expansive medical knowledge combined with private patient data to remove administrative burdens from doctors
- **Retail and the public sector** by enabling these organizations to improve customer experiences and efficiency through automation and dynamic forecasting, helping to reduce risk and drive increased customer satisfaction

¹, ⁴ "Future Enterprise Resiliency and Spending, Wave 1," IDC, January 2024.

² "IDC Forecasts Spending on GenAI Solutions Will Reach \$143 Billion in 2027 with a Five-Year Compound Annual Growth Rate of 73.3%," IDC, October 2023.

³ "Get Productive with AI," HPE conversation guide.

⁵ "Preference for Flexible Consumption of Dedicated GenAI Infrastructure," IDC, May 2024.

⁶ "Reasons Why Generative AI Pilots Fail To Move into Production," Forbes, January 2024.



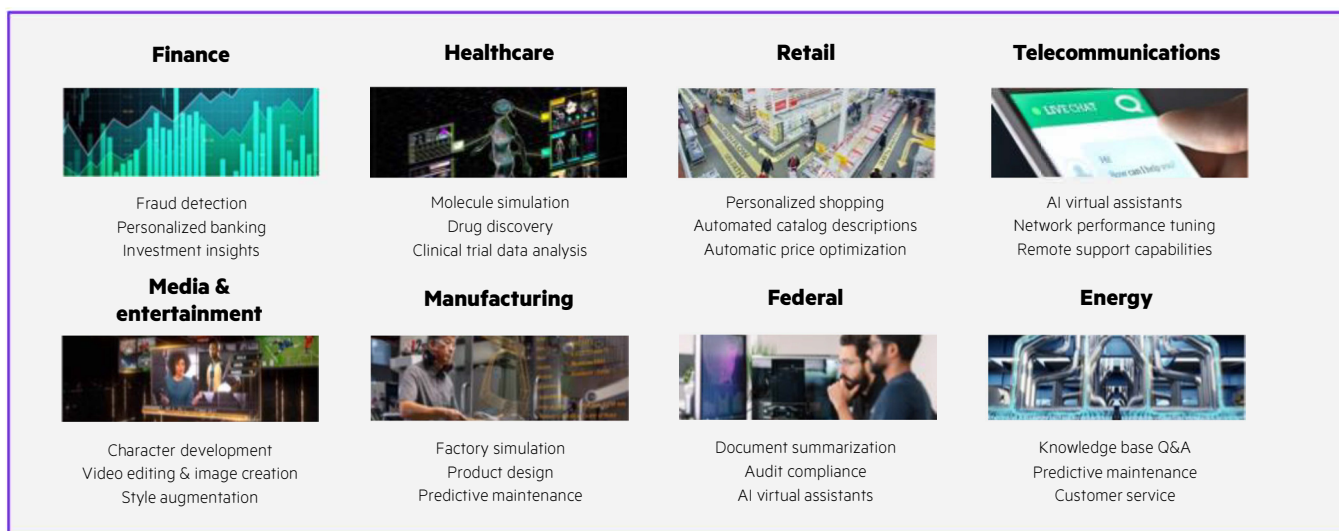


Figure 1. AI use cases across vertical industries

Today, Hewlett Packard Enterprise (HPE) and HCLTech can help your organization capitalize on the transformational potential of AI. When you work with HCLTech and HPE to deploy a tailored solution built on HPE Private Cloud AI and HCLTech's U4X (Utility for Everything) and cognitive Services, you can:

- Turn your AI strategy into successful projects.
- Speed the time to value for your AI initiatives and beyond.
- Extend AI across your organization.
- Help make your data AI ready.
- Solve your toughest business problems at scale.

Delivering solutions for every stage of your AI journey

AI is opening unlimited possibilities for organizations of every size. HCLTech and HPE help you capitalize on those opportunities by delivering comprehensive AI solutions that can turn questions into discoveries, insights into action, and imagination into reality.

Turnkey private AI cloud

Co-developed with NVIDIA—and part of the NVIDIA AI Computing by HPE portfolio—HPE Private Cloud AI is a turnkey private cloud designed to accelerate your path from AI pilot to production. Offering a public cloud experience in a private environment, this end-to-end AI platform combines NVIDIA AI Enterprise software and powerful GPUs with leading HPE compute, storage, software, and services.

HPE Private Cloud AI also includes HPE AI Essentials—a robust set of open-source tools that has been validated and integrated to address the needs of anyone involved in developing GenAI applications.

With HPE Private Cloud AI, you can use the latest technologies to gain secure access to all your data from data center to edge—helping you achieve instant AI productivity and drive greater value across AI use cases.



AI training

Using a very large data set to train an AI model to make accurate predictions or decisions

AI tuning

Optimizing an AI model for a specific purpose; using a smaller data set to further train the model

AI inferencing

Allowing a trained AI model to draw conclusions from brand-new data



Workload-optimized configurations

Available in purpose-driven architecture sizes and expandable starting points, HPE Private Cloud AI includes everything you need to deploy and run the most common AI workloads. Depending on the workloads you need to support, you can choose from the following workload-optimized configurations.

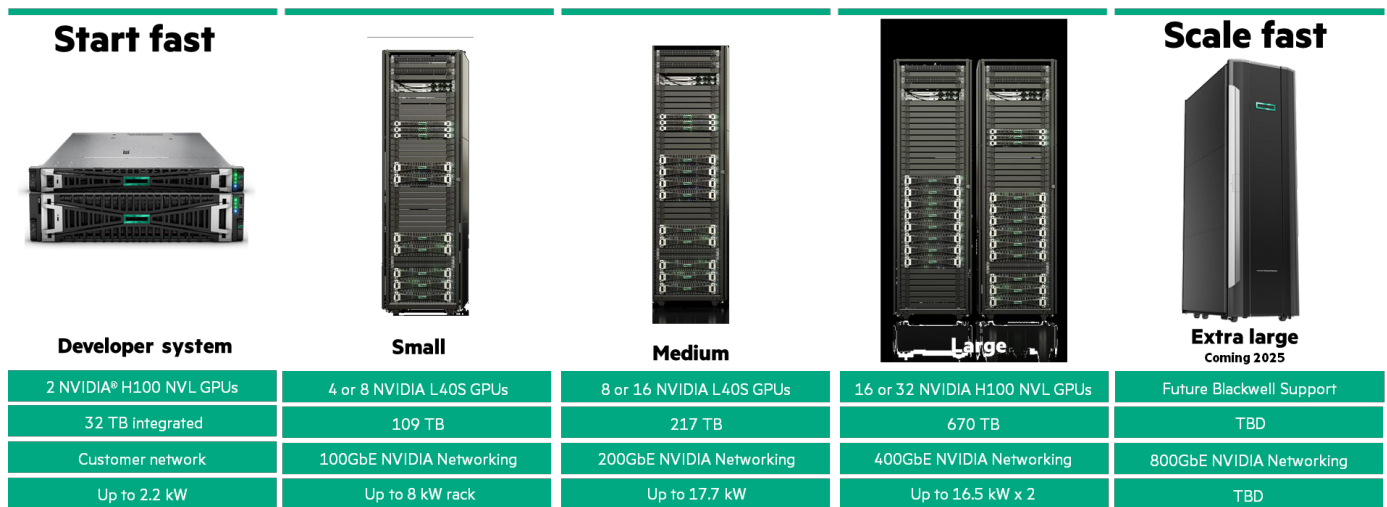


Figure 2. HPE Private Cloud AI architectures in small, medium, large, and extra-large configurations

Enhancing efficiency and saving costs by delivering everything as a service

HCLTech provides a wide range of services via their U4X and Cognitive Infrastructure services offerings to ensure organizations can realize the intended value from their AI investments.

AI consulting & advisory services	Platform and infrastructure build services	Operate and govern services
<ul style="list-style-type: none"> Joint offering with HPE with the goal of creating roadmaps or understanding specific use cases and building an approach and methodology to evaluate/implement it at scale Business case analysis, existing infrastructure analysis and how to transition towards the AI enabled Infra/Infra for AI in various consumption models help evaluate and decide buy vs. build options or the best fit options with DPU, TPU & GPU models Training and inference platform roadmaps 	<ul style="list-style-type: none"> Create, test and deploy greenfield or brownfield Cognitive Infrastructure options in various consumption models Extending an existing landscape based on specific use cases and requirements. Core infra and platform build services Software stack configuration, automation via accelerators and IPS Proof of Concept implementations Pre-integrated and validated stack options with HPE with flexible commercial models Fully tested and validated third-party integrations as part of build can be utilized 	<ul style="list-style-type: none"> Various support and operate models combining on-site/off-site support for Cognitive Infrastructure lifecycle management GenAI platform management Availability and performance monitoring Unified IT infrastructure and cloud monitoring and observability Governance models to ensure compliance is always maintained Various IPS and accelerators to support operate and govern services

Figure 3. HCLTech Cognitive Infrastructure service overview



Productivity use cases





Analyze and summarize queries of care functions	Storage support bot	Migration support bot	Database migration support bot
<div></div> <p>Utilize advanced search techniques and generative technologies to retrieve relevant information from HCLTech knowledge repositories. Doing so enables users to discover related topics and drill down into subtopics, with citations and supporting documents provided for accuracy.</p>	<div></div> <p>Quickly resolve storage-related issues by summarizing solution steps, creating knowledge bases, and generating root cause analyses. Streamline data center operations with virtual assistance to manage and optimize storage resources, providing instant support and information to enhance the overall experience and satisfaction for support engineers.</p>	<div></div> <p>Retrieve data from a central document repository and generate coherent summaries based on input text. The platform helps architects and engineers quickly resolve migration issues using historical and internet data; it can be extended to various products and business functions, and it is deployable on-premises or in the public cloud.</p>	<div></div> <p>Employ advanced inference and summaries for user queries related to migration tools, historical data, and knowledge bases, including database migration and documentation; the bot uses a RAG-based implementation.</p>

Figure 4. Productivity use cases for HCLTech Cognitive Infrastructure services

HCLTech Cognitive Infrastructure services as an offering is differentiated in the market basis multiple reasons:

- End to end services including workshop based Advisory services, build, migrate, operate (dedicated or shared) and continuous feedback
- Wide array of ecosystem-based partners for full stack/use case based best option deployment
- Shared and dedicated LLMOps to suit hybrid enterprise needs
- One of the industry’s largest pool of GenAI certified resources
- Dedicated GenAI labs and CoEs
- Wide variety and range of simulation and scenario building exercises
- Pre-validated architectures and blueprints with OEMs and CSPs to ensure enterprises can focus on their business objectives
- A robust security/governance and privacy framework to ensure GenAI investments can move from pocket to scale
- Vertical specific use cases ex: Clinical assistant for Pharma
- Helping customers navigate the buy vs. build or third party vs. in-house/native dilemmas
- Sustainability enabled outcomes



Better together—HPE and HCLTech for the next era of AI

For more than 30 years, HPE and HCLTech have been driving innovation together. Our combined experience in AI, HPC, converged infrastructure, virtualization, and hybrid cloud allows our customers to build on an integrated, scalable, and automated IT platform for the future.

HPE and HCLTech work together to make IT environments more efficient and secure from edge to cloud—ensuring fast responses to a rapidly changing landscape without sacrificing privacy, performance, or control. Whatever the size of your organization, HPE and HCLTech can help you transform from traditional technology platforms to a dynamic environment designed for the next era of AI and beyond.

Why HCLTech

- Positioned as a leader in the 2023 Gartner Magic Quadrant for Data Center Outsourcing and Hybrid Infrastructure Managed Services Worldwide
- Rated as a Leader in ISG Provider Lens—Analytics Services, U.S. 2023
- Recognized with a Customers' Choice distinction in the 2023 Gartner Voice of the Customer report for Data Center Outsourcing and Hybrid Infrastructure Managed Services, Worldwide
- Expertise in 100+ GenAI use cases
- Speeding time to market through cloud-native GenAI labs, LPs, and frameworks
- Backed by a pool of AI engineers, data scientists, and infrastructure platform subject matter experts

Learn more at

[HPE.com/us/en/alliance/hcl.html](https://hpe.com/us/en/alliance/hcl.html)

[HPE.com/ai](https://hpe.com/ai)

Visit **HPE.com**



Chat now (sales)

© Copyright 2025 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

NVIDIA is a trademark and/or registered trademark of NVIDIA Corporation in the U.S. and other countries. All third-party marks are property of their respective owners.

a50013030ENW

