



Prepare for the future of gen AI with Red Hat and HCLTech

*"Bringing AI into the enterprise is no longer an 'if,' it's a matter of 'when.'"*¹

Ashesh Badani,

Chief product officer and
senior vice president, Red Hat

The emergence of gen AI in the enterprise

The exponential rate of artificial intelligence (AI) and generative AI (gen AI) adoption seen in enterprises of all industries in recent years shows no signs of slowing down, as many organizations are realizing the wide-ranging benefits of this technology.

Research shows that 65% of organizations are already actively investing in gen AI,² while 80% of enterprises plan to incorporate gen AI into their business processes by 2026.³

Furthermore, IDC predicts that global spending on gen AI software and related infrastructure and IT services will grow at a compound annual rate of 73.3% between 2023 and 2027, reaching an estimated US\$143 billion.⁴

Ranging from simpler use cases such as document summarization and chatbots, to more complex applications such as text-to-code generation and risk prediction and management, many organizations are using gen AI to improve operational efficiency and reduce operational costs, among other key benefits.

Early adopters of gen AI are already seeing significant improvements, including a 35% lift in innovation and a 32% boost in employee and customer experiences.⁵ Perhaps more significantly, research also predicts gen AI will create 80% productivity gains across knowledge workers and creative tasks,⁶ and forecasts global economic benefits of US\$3–4 trillion annually.⁷

But without the agile and flexible infrastructure required to support their gen AI innovation, enterprises will struggle to realize this full range of benefits.

The common challenges of gen AI infrastructure

Red Hat and HCLTech help customers unlock the full potential of gen AI by helping them address some of the most common challenges, including:

Data security and sovereignty. As datasets in domains like AI grow exponentially, transferring that data to cloud-based or distant infrastructure resources introduces significant risks and complexities related to data privacy, data security, and compliance with regulatory standards.

¹ Red Hat press release. "[Red Hat OpenShift AI Expands Predictive and Generative AI Flexibility Across the Hybrid Cloud](#)," 7 May 2024.

² "[The state of AI in early 2024: Gen AI adoption spikes and starts to generate value](#)," McKinsey, 30 May 2024.

³ "[Gartner Says More Than 80% of Enterprises Will Have Used Generative AI APIs or Deployed Generative AI-Enabled Applications by 2026](#)," Gartner, 11 Oct. 2023.

⁴ IDC press release. "[IDC Forecasts Spending on GenAI Solutions Will Reach \\$143 Billion in 2027 with a Five-Year Compound Annual Growth Rate of 73.3%](#)," 16 Oct. 2023.

⁵ IDC White Paper, sponsored by Red Hat. "[Why Open Source Artificial Intelligence Platforms Help Enterprise Business Transformation](#)," Document #US51272823, Nov. 2023.

⁶ Gartner press release. "[Gartner Identifies the Top 10 Strategic Technology Trends for 2024](#)," 16 Oct. 2023.

⁷ "[The economic potential of generative AI: The next productivity frontier](#)," McKinsey, 14 June 2023.

Infrastructure costs. One of the most significant barriers that enterprises entering the gen AI space must contend with is the underlying cost of the required infrastructure and computational resources. Costs can become difficult to control without the right approach to managing resource consumption and allocation.

Out-of-the-box solutions. With the surge of popularity in the space, many organizations have tried to shortcut their way to gen AI success by looking for out-of-the-box gen AI solutions that they believe will deliver everything they need. But too often, these gen AI solutions are incapable of meeting the specific requirements of an organization. This is especially true for security and compliance needs, which are vital to the ongoing success of gen AI innovation, and can vary significantly depending on the industries or countries the enterprise operates in.

By working with Red Hat and HCLTech, customers can find the technology, expertise, and professional services they need to navigate these challenges, among many others, and build gen AI solutions that meet their specific security and compliance requirements, with a prudent approach to cost management.

Unlock the power of gen AI with Red Hat and HCLTech

The core of Red Hat and HCLTech's joint service offering is the convergence of Red Hat's end-to-end AI application platform—Red Hat® OpenShift® AI—and the technical expertise and professional services HCLTech has built over its years as a leader in the gen AI space.

Bringing these together, Red Hat and HCLTech work with enterprises to find solutions to their critical business challenges with gen AI technology.

Red Hat OpenShift AI—which is available as a fully managed cloud service or as a self-managed platform on-premise or in a cloud environment—provides enterprises with a modern application platform to build, train, deploy, manage, and scale AI-enabled applications, with a common platform for all teams to collaborate on artificial intelligence/machine learning (AI/ML) models.

These capabilities are further supported by:

- ▶ Hybrid and multicloud portability.
- ▶ Reduced management complexity and simplified collaboration.
- ▶ An extensive graphic user interface that simplifies how enterprises integrate AI.
- ▶ The flexibility of Red Hat's open source strategy that allows organizations to innovate and test new AI technologies with minimal resources or disruption.
- ▶ Prebuilt AI and automation components and tools that reduce development time and simplify AI integration.
- ▶ Access to Red Hat's extensive partner ecosystem of solutions and services to streamline the entire AI lifecycle.

By bringing this foundation together with HCLTech's expertise and professional services—built on the experience of successfully implementing over 150 gen AI use cases in a range of industries across clients—enterprises can maximize their chances of successfully building valuable gen AI initiatives.

Through a combination of specialized technical knowledge, intellectual property (including solutions for gen AI-powered chatbots, code generation, candidate interviewing, and more), an ecosystem of collaborators, and global delivery capabilities, HCLTech helps their customers accelerate gen AI adoption with core infrastructure, intelligent operations, and digital services.

This is delivered through services that address the critical stages of an organization's gen AI journey, including:

- ▶ AI advisory services.
- ▶ Operation and management services.
- ▶ Platform and infrastructure building services.

Enterprises that work with HCLTech on their gen AI projects also benefit from:

- ▶ Consulting services to support adoption and scaling of gen AI.
- ▶ An end-to-end solution with tooling, DevOps, and automation capabilities.
- ▶ Reference architectures and frameworks to support adoption of gen AI vertical and horizontal use cases.
- ▶ Business case and total cost of ownership (TCO) analysis.
- ▶ Faster time-to-market through cloud-native gen AI labs, IPs, and frameworks.
- ▶ The expertise of a highly trained and certified pool of data scientists, AI engineers, and platform architects and developers.
- ▶ Access to a diverse partner ecosystem that includes Dell, HPE, NVIDIA, IBM watsonX, Rubrik, Commvault, and hyperscalers such as Microsoft Azure, Amazon Web Services (AWS), and Google.

How Red Hat and HCLTech help enterprises tap into the full potential of gen AI

Working with Red Hat and HCLTech provides customers with the technology base and consulting services needed to overcome the barrier to entry, maximize their AI resources, and create successful gen AI solutions that provide business value far into the future.

The convergence of Red Hat's AI platforms and solutions and HCLTech's AI framework and service delivery model provides enterprises with a complete offering that includes:

- ▶ An end-to-end gen AI technology stack.
- ▶ A gen AI infrastructure-as-a-service offering.
- ▶ Comprehensive managed services including consulting, design, model training, fine tuning, implementation, monitoring, and operations services.

These collaborative offerings and services help enterprises:

- ▶ Streamline model development and deployment lifecycle.
- ▶ Implement robust governance in compliance with industry standards and data privacy regulations.
- ▶ Continuously monitor AI workloads, especially for industry-specific use cases that call for proactive maintenance.

By guiding enterprises through the most common challenges of managing gen AI projects and the infrastructure they require, Red Hat and HCLTech are able to help customers achieve significant operational benefits, including:

Faster time-to-value. Red Hat and HCLTech deliver bespoke offerings that meet any enterprise’s needs right away to decrease implementation friction and improve time-to-value. This is made possible through the combination of working closely with their customers to understand their specific challenges and goals, a vast range of customized industry-specific solutions and frameworks, and the extensive industry and gen AI expertise needed to customize models and infrastructure to any organization’s specific requirements.

Increased focus on data security and governance. Organizations require a framework that focuses on security and compliance to help them protect their data against unauthorized access and data breaches. Red Hat and HCLTech help customers create that framework with a platform that allows data processing and AI capabilities to be moved closer to the data’s origin, or refine data management practices to reduce transfer volumes, while also providing the capabilities needed to enforce robust encryption, access controls, and compliance measures throughout the data lifecycle.

Optimized resource usage and costs. With a focus on high resource utilization and enhanced operational efficiency of the underlying gen AI infrastructure, organizations can streamline operational costs and reduce the TCO of their AI investments. Red Hat and HCLTech help customers achieve this with Red Hat OpenShift AI providing the capabilities, and HCLTech providing the needed expert guidance, to efficiently manage and distribute their gen AI workloads to make sure that resources, such as graphic processing units (GPUs), are used to their fullest potential.

Prepare for the future of gen AI with Red Hat and HCLTech

The emergence of gen AI has created many opportunities for enterprises to drastically overhaul and improve both how they approach internal operations and how they deliver value to their end customers, and the future of the technology is only expected to open up even more possibilities.

Red Hat and HCLTech can help your enterprise prepare for that future now with the technology, expertise, and services you need to successfully develop, deploy, and manage gen AI to enhance your core business operations.

[Contact HCLTech](#) to book a free discovery session with a gen AI expert and discuss your organization’s gen AI objectives and challenges, how you can maximize the value of your infrastructure to implement gen AI use cases, and how Red Hat and HCLTech can help.



About Red Hat

Red Hat is the world’s leading provider of enterprise open source software solutions, using a community-powered approach to deliver reliable and high-performing Linux, hybrid cloud, container, and Kubernetes technologies. Red Hat helps customers develop cloud-native applications, integrate existing and new IT applications, and automate and manage complex environments. [A trusted adviser to the Fortune 500](#), Red Hat provides [award-winning](#) support, training, and consulting services that bring the benefits of open innovation to any industry. Red Hat is a connective hub in a global network of enterprises, partners, and communities, helping organizations grow, transform, and prepare for the digital future.

North America

1 888 REDHAT1
www.redhat.com

Europe, Middle East,
and Africa

00800 7334 2835
europe@redhat.com

Asia Pacific

+65 6490 4200
apac@redhat.com

Latin America

+54 11 4329 7300
info-latam@redhat.com