

# Turn AI vision into measurable value with HPE AI Factory and Red Hat AI

## Is the complexity of AI standing in the way of innovation?

As organizations embrace the promise of artificial intelligence, many discover that turning AI from experimentation into production is difficult.

[Gartner](#) predicts that more than 40% of agentic AI initiatives will be cancelled by 2027 due to cost, risk, and unclear business value.

AI workloads are diverse and resource intensive, requiring highly specialized infrastructure. Multitenancy demands, cloud costs, operational complexities and strict security requirements further complicate progress. Teams often struggle with data management, talent gaps, model drift, and limited internal expertise. Without a structured way to industrialize AI, organizations face delays, increased costs, and inconsistent outcomes across projects.

## Technology Used

- Red Hat AI Enterprise
- Red Hat OpenShift

## Key benefits

- Faster AI deployment with an end-to-end factory approach
- Lower operational complexity through integrated infrastructure and tooling
- Reduced AI delivery costs with efficient resource utilization and on-premises control
- Improved scalability with enterprise-grade AI lifecycle management
- Stronger governance and compliance for secure AI development

## HPE AI Factory with Red Hat AI enables scalable, production-ready AI

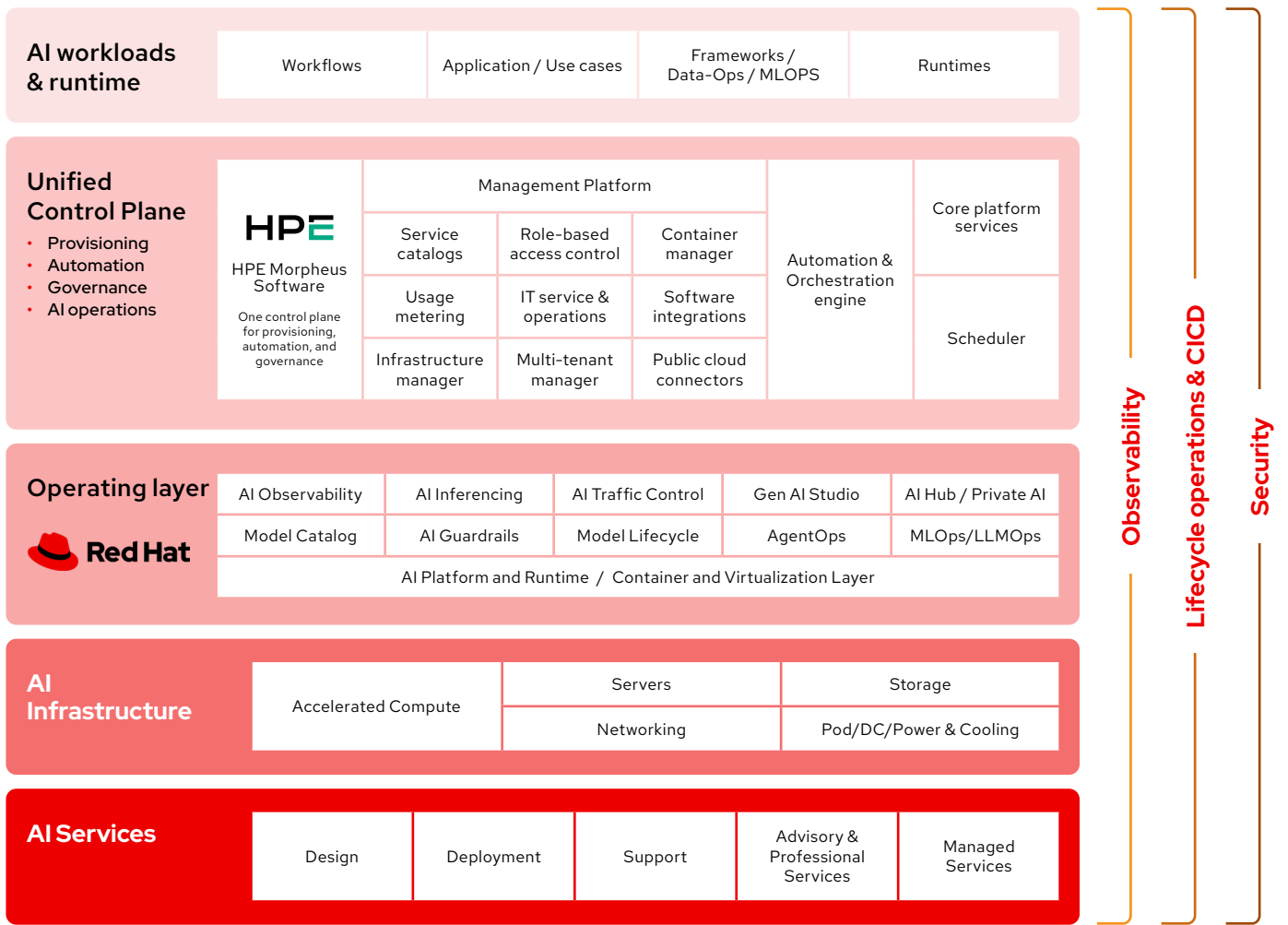
At the core of the HPE AI Factory is HPE Morpheus Software—a unified control plane that turns hybrid IT and AI operational complexity into control. HPE Morpheus simplifies provisioning, automation, and governance across infrastructure, data, and AI workloads, enabling organizations to operationalize AI consistently across on-premises and cloud environments.

Backed by Red Hat AI Enterprise, the solution delivers a full-stack environment that transforms the AI lifecycle from fragmented and unpredictable into streamlined, repeatable, and reliable.

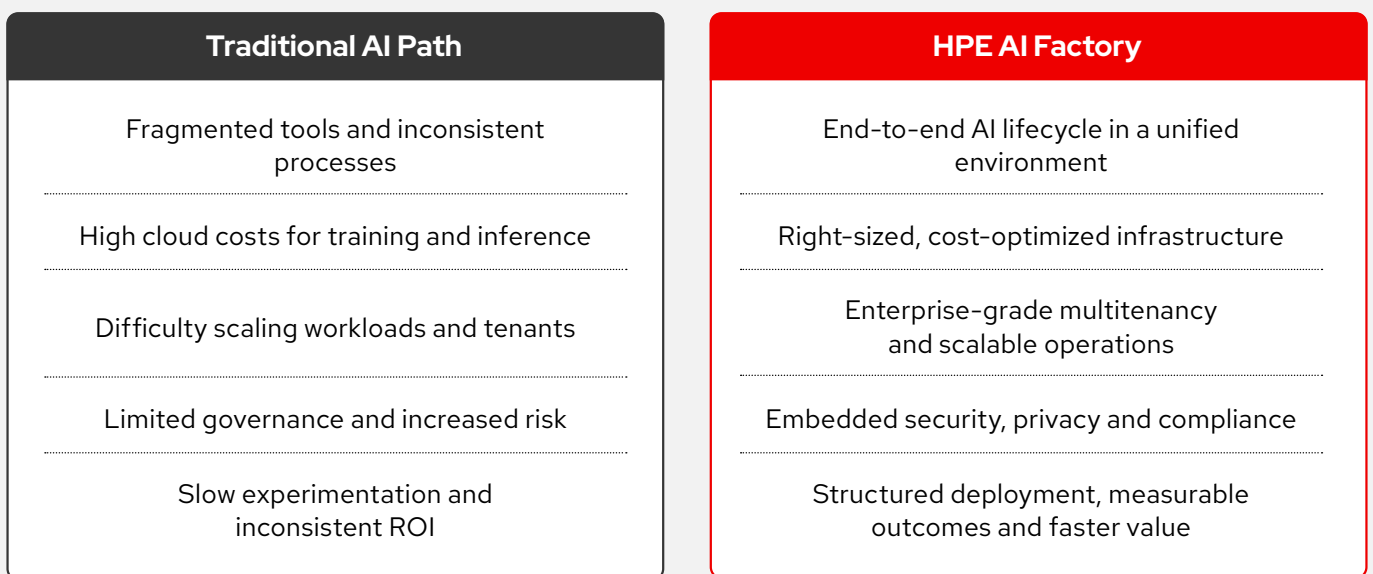
With the HPE AI Factory portfolio, organizations can industrialize AI with confidence. This control plane delivers simplicity through consistent operations, freedom of choice across hybrid environments, and predictable economics with built-in governance and cost control. The approach unifies data access, model development, deployment, retraining, and governance into a single consistent framework.

HPE experts work side by side with teams across IT, data, and business functions, ensuring that AI programs align with strategic outcomes while reducing risk and accelerating time to value.

- **Accelerated deployment:** A pre-integrated stack and a structured methodology shorten the path from concept to production.
- **Scalable growth:** Flexible infrastructure supports evolving AI workloads, from model training to complex simulations.
- **Operational efficiency:** Automated MLOps and integrated tools reduce manual work and maintain performance as models evolve.
- **Security and governance:** Built-in controls ensure responsible, compliant and auditable AI across all environments.



## Traditional AI attempts versus HPE AI Factory



## Key features of the solution

- **Full-stack AI infrastructure:**  
HPE delivers integrated compute, networking, storage, software and services to power demanding AI workloads.
- **End-to-end AI lifecycle management:**  
Support for data pipelines, model development, training, inference and continuous improvement.
- **Support for diverse AI teams:**  
Suitable for enterprises, model builders, service providers and sovereign AI initiatives.
- **Performance and reliability:**  
Engineered for retraining, simulations, inference quality monitoring and rapid iteration.
- **Automated operations and MLOps:**  
Simplifies deployment, monitoring and retraining as data and use cases evolve.
- **Expert-guided delivery:**  
HPE AI specialists work alongside internal teams to design, operate and scale AI programs effectively.
- **HPE and Red Hat** have a long-standing partnership, particularly in mission-critical and regulated industries. We have extended this to AI factories.
- **The HPE AI Factory solution** supports tight integration with the Red Hat AI Enterprise offering (which includes Red Hat OpenShift, Red Hat OpenShift AI, and Red Hat AI Accelerators), which provides the benefit of running efficient, cost-effective AI models, agentic AI workflows, and AI-powered applications on HPE infrastructure.
- **HPE AI Factory** control plane supports automated deployment and monitoring of Red Hat SW inside a tenant.
- **The self-service catalogs** provisioned inside a tenant can be configured to host Red Hat provided applications and tools such as OpenShift AI for users to consume.
- **As part of the HPE AI Factory solution,** we provide compliance and security for Sovereign AI deployments through a combination of hardened/validated/certified SW and services to deploy and maintain it.

---

## Getting started is simple

HPE AI Factory provides a structured, high-value pathway to operational AI.

- **Define strategy and assess readiness:**  
Identify AI priorities, evaluate data requirements and assess organizational capabilities.
- **Develop data foundations:**  
Address silos, clean and integrate data, and establish governance models for reliable access.
- **Design the AI architecture:**  
Build a tailored environment based on the HPE AI Factory solution and Red Hat AI Enterprise.
- **Build and deploy AI solutions:**  
Implement pipelines using pre-trained models, RAG, fine-tuning or training as required.
- **Operate, govern and refine:**  
Use automated MLOps and monitoring to maintain model relevance and continuous improvement.

This structured path ensures organizations can move from vision to an operational AI Accelerator in a predictable timeframe, supported by expert guidance.

## Success story

A major North American telecommunications company that operates a nationwide network leveraged the HPE AI Factory approach to accelerate AI adoption across multiple business units. By aligning data, infrastructure and processes under a unified factory model, the organization reduced operational AI costs by up to 40 percent and increased developer efficiency by up to 60 percent for documentation-driven use cases.

Teams saw improved collaboration and clearer ROI, enabling faster scaling of AI initiatives. The result was a streamlined, measurable and repeatable AI delivery model that supported long-term growth.



### Day 1/ Day 0 services help you:

- Assess technical readiness.
- Check and update data foundation.
- Determine data center strategy.
- Discover and align use cases.
- Design platform ecosystem.
- Address governance and cyber resilience.

### Day 1 services help you:

- Implement solution architecture.
- Customize to target use cases.
- Automate MLOps.
- Optimize carbon footprint and energy efficiency.

### Day 2 services help you:

- Operate the solution ongoing.
- Augment your staff to support the solution.
- Focus on business outcomes, leaving managing operations to HPE.

← **Cybersecurity, networking, education, financing, upcycling** →

## Let HPE help you confidently transform your business with AI

As your strategic partner, HPE helps your organization confidently harness the transformative power of AI. Through comprehensive services, covering ideation, development, deployment, and ongoing optimization, we help accelerate innovation, achieve measurable results, and future-proof your business.

### 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.

f facebook.com/RedHatinc

X @RedHat

in linkedin.com/company/red-hat

**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