

HPE Discover 2025: The Executive Recap



■ The Era of Sovereign AI Factories

The defining message of HPE Discover 2025 was the shift from "AI experimentation" to "industrial-scale production." This theme was anchored by the expanded partnership between HPE and NVIDIA.

In the press announcement accompanying the event, NVIDIA CEO Jensen Huang framed the strategy clearly: "We're transforming the data center into an AI factory, a manufacturing plant for the new industrial revolution."

However, the barrier to this revolution is no longer just compute power, it is **sovereignty**. Governments and regulated industries (Financial Services, Healthcare, Public Sector) cannot rely on public clouds for their most sensitive intelligence. They need to own the production.

Key Announcement: The Grenoble AI Factory Lab

To address this, HPE and NVIDIA announced a massive expansion of their partnership centered on the European market.

- **The Facility:** A new **AI Factory Lab in Grenoble, France**, launching in Q2 2026.
- **The Purpose:** This facility allows customers to validate performance on infrastructure that is physically located and running entirely within the EU.
- **The Tech Stack:** It is equipped with the latest **NVIDIA AI Enterprise** government-ready software, HPE servers, and **HPE Juniper Networking PTX/MX routers**.
- **The Strategic Value:** This creates a "safe harbor" for EU enterprises to test sovereign AI factories before deploying them, ensuring that data residency and regulatory compliance needs are met from Day 1.

"Helios": The Infrastructure of Scale

While sovereignty solves the legal barrier, the **AMD "Helios" AI rack-scale architecture** solves the physics barrier.

- **The Problem:** Training trillion-parameter models creates traffic jams that traditional networks cannot handle.
- **The Solution:** HPE unveiled the "Helios" rack, a turnkey solution connecting **72 AMD Instinct™ MI455X GPUs** per rack.
- **The Specs:**
 - **260 terabytes/second** of aggregated scale-up bandwidth.
 - **2.9 AI exaflops** of FP4 performance.
 - **Industry-First Ethernet:** Developed with Broadcom, this system uses standard Ethernet for scale-up networking, breaking the reliance on proprietary interconnects and lowering TCO.

"Self-Driving" Networks & The Production Gap

The "Agentic" Network (Juniper + Mist Integration)

Following the acquisition of Juniper Networks, HPE demonstrated how AI is moving from a passive tool to an active agent in IT operations. The integration of **HPE Aruba Networking** and **HPE Juniper Networking** Mist has created a "self-driving" network.

- **Agentic Mesh Technology:** Now available for Mist, this technology enhances anomaly detection with advanced reasoning. It doesn't just alert IT to a problem; it performs autonomous root-cause analysis.
- **Large Experience Model (LEM):** By using billions of data points from apps like Zoom and Teams, the network can predict and fix video quality issues before users even report a ticket.
- **Implication for Business:** As our infrastructure becomes "agentic," the need for automated testing and guardrails becomes existential. You cannot manually oversee a network that makes thousands of decisions per second.

The Virtualization Pivot (HPE Morpheus)

HPE is also aggressively targeting the "VMware tax" with expanded capabilities for HPE Morpheus Software.

- **Cost Reduction:** HPE claims Morpheus VM Essentials can reduce VM license costs by up to 90%.
- **Tech Stack:** It offers a KVM-based alternative with a unified control plane for both VMs and Containers (Kubernetes).
- **Resilience:** Integration with **HPE Zerto** ensures continuous data protection and "always-on" availability.

The "10% Cliff"

Despite these advancements in hardware and networking, a critical industry statistic highlighted in the "**Trustworthy AI at Scale**" session remains the primary bottleneck: **Less than 10% of GenAI projects successfully reach production.**

The barrier isn't a lack of tools; it's a lack of an integrated ecosystem. No single vendor can deliver infrastructure, security, orchestration, and governance in isolation. This is why HPE launched "**Unleash AI**".

The "Unleash AI" Ecosystem

A Curated Ecosystem for Production

Recognizing that enterprise AI requires a diverse "village" of specialized vendors, HPE introduced the Unleash AI partner program. This is a comprehensive ecosystem of independent software vendors (ISVs) validated to run on HPE Private Cloud AI, allowing customers to choose the best tools for their specific data and model needs.

The keynote highlighted three of these critical partners that illustrate how this ecosystem solves the "Pilot to Production" challenge across different layers of the stack.

1. Orchestration & Scale: Dataiku

- **The Role:** Providing a "Universal AI Platform" to manage the lifecycle of AI development.
- **The Value:** Dataiku gives teams a central control plane to build and orchestrate AI applications, from simple models to complex Agents, using NVIDIA's industry-specific AI Blueprints. It ensures that data science teams can collaborate and deploy workloads efficiently across the HPE stack.

2. Security & Sovereignty: Fortanix

- **The Role:** Protecting sensitive data in use.
- **The Value:** For highly regulated industries, Fortanix brings **Confidential Computing** to the HPE Private Cloud. Their platform, **Armet AI**, ensures that data remains encrypted even during processing (training and inference). This allows enterprises to run secure AI workloads on-premises without fear of data exposure or insider threats.

3. AI Quality Assurance & Testing: QuantPi

- **The Role:** The automated Quality Assurance (QA) engine for model reliability.
- **The Value:** QuantPi provides a scalable, model-agnostic testing engine that serves as the "Trust Layer" for the entire AI portfolio. Unlike fragmented testing approaches, QuantPi's engine applies a standardized testing framework across **GenAI, Agents, Computer Vision, Voice, and Video**. It leverages **NVIDIA technology** for smart data augmentation to rigorously stress-test models against diverse real-world conditions, preventing **unreliable performance** and **unintended bias**.

Resources & Next Steps: Kickstarting Your 2026 AI Strategy

As you kickstart 2026 and move from planning to execution, use the resources below as a starting guide to validate key components of your production stack.

The "Pilot to Production" Checklist before deploying your first agentic workload

- ☐ **Infrastructure:** Is the environment sovereign and air-gapped?
- ☐ **Data Intelligence:** Is data accessible and "active" for high-speed training?
- ☐ **Orchestration:** Do developers have a centralized "Workshop" to build models?
- ☐ **Security:** Is data encrypted during inference (in-use)?
- ☐ **Quality Assurance:** Have the models been stress-tested on bias, performance etc.?
- ☐ **Operations:** Can you monitor full-stack observability and agent behavior post-deployment?

Further readings

- [HPE Private Cloud AI](#) (including list of verified "Unleash AI" partner overview)
- [Trustworthy AI at Scale](#) (Technical Whitepaper by HPE, NVIDIA and QuantPi)

When you are ready to validate your roadmap, we are available to demonstrate how QuantPi can automate the quality assurance or "Trust Layer" for your agentic workloads.

[Book a Demo](#)

△ QUANTPI