

HCLTech | intel.

Cognitive infrastructure services

built on Intel Xeon 6 and optimized
with Red Hat® OpenShift AI



Overview

HCLTech offers a high-performance cognitive infrastructure solution built on the newest Intel Xeon 6 processors to meet the rigorous demands of GenAI. This powerful platform is optimized with Red Hat® OpenShift AI to provide a complete, enterprise-grade solution for building, deploying and managing AI applications.

This integrated stack—combining Red Hat® OpenShift AI, HCLTech's frameworks and Intel Xeon 6 hardware—empowers organizations to harness large-scale data, streamline ML lifecycle management and accelerate AI deployment while ensuring transparency, governance and cost-efficiency.

Solution highlights for Red Hat® OpenShift AI



Red Hat® OpenShift AI:

An enterprise-ready platform for managing the lifecycle of predictive and GenAI models, including model development, training and serving, at scale, across hybrid cloud environments



End-to-end

services: From AI strategy workshops to infrastructure design, deployment and ongoing operations



Scalable

architecture: Supports training and inference of foundation models across edge and datacenter



Model governance:

Built-in capabilities for lineage, transparency and access control

Intel Xeon 6 value proposition

- **Built-in AI acceleration:** AMX and AVX-512 and Vector Neural Network Instructions (VNNI) combined with vLLM's advanced inference engine features like paged attention and continuous batching for faster LLM inference
- **Composable infrastructure:** CXL 2.0 enables memory pooling and hardware disaggregation
- **Performance/power flexibility:** P-core for high-performance tasks, E-core for efficient background operations
- **Intel Confidential Computing:** Secure with Intel TDX for workload isolation
- **Intel enterprise AI:** Integrated software and hardware suite—including Intel oneAPI AI Analytics Toolkit and OpenVINO Toolkit—accelerating AI development and deployment across edge, data center and cloud



Key benefits



Accelerate time-to-value:

Faster AI experimentation to production using Red Hat® OpenShift AI



Enable multi-team collaboration:

One platform for data science, MLOps and app development



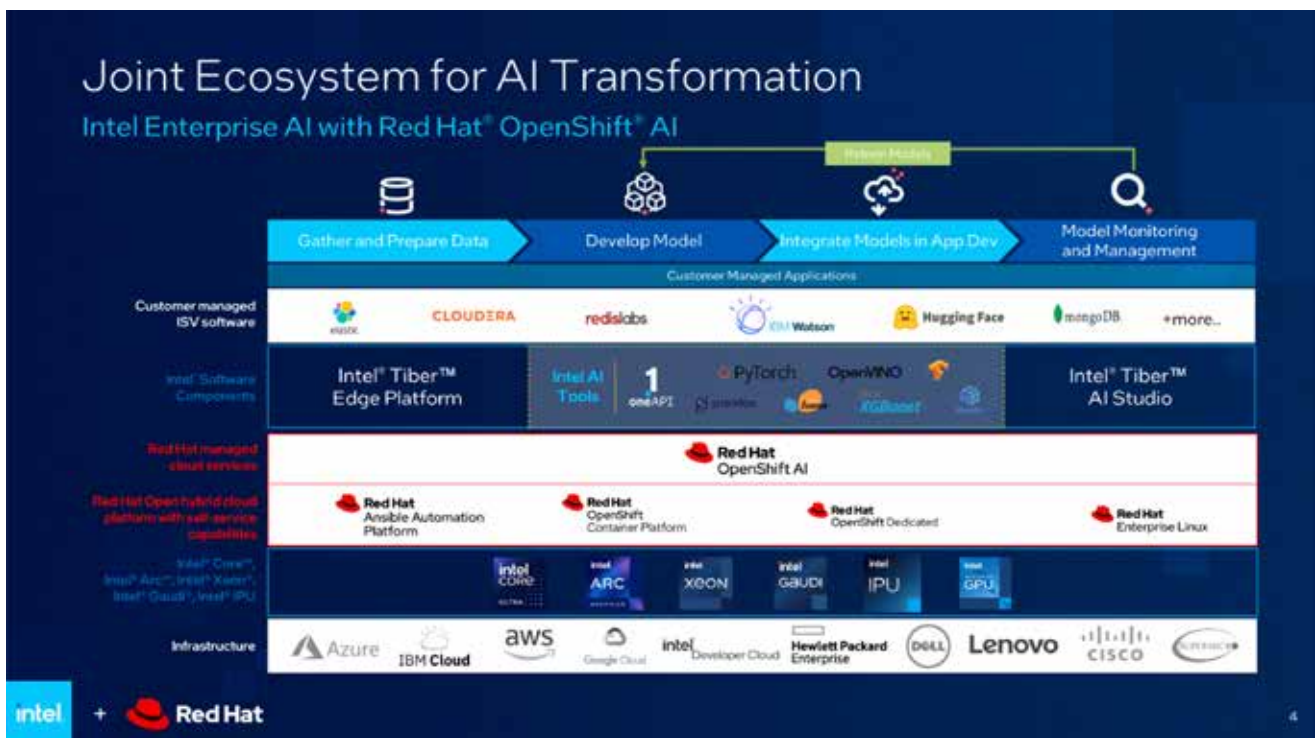
Reduce TCO:

Optimize resource usage with Xeon 6's flexible architecture and efficiently run GenAI workloads—often without extra GPUs—enabling significant server consolidation



Improve model outcomes:

Enhanced observability and performance tuning supported by hardware-level AI features



Why HCLTech, Intel and Red Hat®

The combined value of HCLTech, Intel and Red Hat® brings a powerful trifecta of hardware and software. We provide a trusted and robust platform with the Intel Xeon 6 stack, which is designed for modern AI workloads. This is complemented by our premier partnership with Red Hat® and deep expertise in OpenShift AI, creating a unified solution that addresses every stage of the AI lifecycle.

100+ GenAI use cases and frameworks delivered

Premier Red Hat® partner with deep OpenShift AI expertise

Dedicated Intel ecosystem unit and 30+ year partnership

End-to-end engagement from AI discovery to scaled deployment

HCLTech | Supercharging
Progress™

hcltech.com