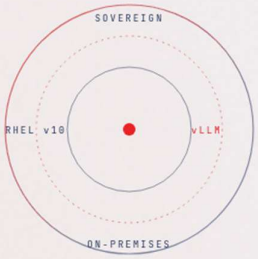
TWO-DAY HANDS-ON WORKSHOP · EDITION 2026

Hybrid Cloud, *Data Sovereignty*



A two-day, instructor-led deep dive into **OpenShift AI with vLLM**, **Red Hat Enterprise Linux v10**, **Ansible automation**, and **IBM HCI** — engineered for teams that keep their data on-premises and their inference local. Build, secure, and operate an end-to-end RHEL AI stack running large language models in containers, fully automated by Ansible.

DOWNLOAD AGENDA | RESERVE A SEAT →

FORMAT	AUDIENCE	STYLE	CAPACITY
2 Days · On-site	IT, DevOps, AI Eng.	Labs & Demos	Limited Seating

// THE STACK • OpenShift AI • vLLM • RHEL v10 • Ansible • IBM HCI • Containers

02 / AGENDA

Two days. *One sovereign stack.*

Day 1 grounds the foundations — Ansible, OpenShift, secure APIs, and containers. Day 2 advances into the RHEL AI stack: deploying vLLM-powered LLMs in containers, orchestrated by OpenShift and driven by Ansible automation.

DAY 01
Foundations & Hands-On Labs

DAY 02
RHEL AI Stack · LLMs in Containers

9:00 – 9:30 AM

WELCOME

Welcome & Orientation

- Workshop objectives, structure, and ground rules
- Overview of technologies: OpenShift AI, RHEL v10, Ansible, IBM HCI, containerized workloads
- Safety, logistics, and shared resources

9:30 – 10:30 AM

SESSION 01 · AUTOMATION

Introduction to Ansible

- What is Ansible? Key concepts and terminology
- Role of Ansible in automation and orchestration
- Setting up Ansible on RHEL v10

10:45 AM – 12:00 PM

SESSION 02 · PLATFORM

OpenShift & CI/CD Pipeline Fundamentals

- OpenShift architecture and deployment scenarios
- Continuous Integration / Continuous Deployment pipeline basics
- Integrating Ansible with OpenShift for automated deployments

---

12:00 – 1:00 PM

SESSION 03 · SECURITY

Building & Securing APIs

- API design and development for cloud-native, on-premises environments
- Security best practices: authentication, authorization, and encryption
- Hands-on API protection using OpenShift and Ansible

1:00 – 2:00 PM — Lunch Break —

2:00 – 3:00 PM

SESSION 04 · CONTAINERS

Introduction to Containers

- Container concepts: images, registries, orchestration
- Why containers? Benefits for scalability, portability, and sovereignty
- Overview of containerized workloads on IBM HCI with GPU acceleration

---

3:15 – 4:45 PM

HANDS-ON LABS

Guided Lab Block

- Lab 1 · Setting up Ansible playbooks for automation
- Lab 2 · Deploying and configuring OpenShift clusters
- Lab 3 · Building secure APIs and integrating with OpenShift
- Lab 4 · Creating and running your first container

4:45 – 5:30 PM

LIVE DEMOS

Instructor-Led Demonstrations

- Live demo: OpenShift deployment workflow
- Live demo: CI/CD pipeline integration with Ansible
- Live demo: API security implementation
- Live demo: Container deployment and management

5:30 – 6:00 PM

WRAP-UP

Day 1 Recap & Q&A

- Recap of key concepts and activities
- Open Q&A session
- Preview of Day 2 — the RHEL AI stack with LLMs in containers

DAY 01

Foundations & Hands-On Labs

DAY 02

RHEL AI Stack · LLMs in Containers

9:00 – 9:45 AM

SESSION 01 · ARCHITECTURE

RHEL AI Stack — Architecture & Data Sovereignty

- → The RHEL AI stack: kernel, drivers, container runtime, model server, and orchestration layer
- → Data sovereignty principles: keeping inference, embeddings, and prompts on-premises
- → IBM HCI as a sovereign substrate — security boundaries, GPU passthrough, tenant isolation
- → Reference architecture walk-through and design trade-offs

10:00 – 11:00 AM

SESSION 02 · MODEL SERVING

vLLM on RHEL v10 — Running LLMs in Containers

- → Why vLLM: PagedAttention, continuous batching, and high-throughput inference
- → Packaging vLLM as a container image for RHEL v10 with NVIDIA GPU support
- → Loading open models (Granite, Llama, Mistral) from a local, air-gapped registry
- → OpenAI-compatible API endpoints — exposed securely behind OpenShift routes

11:15 AM – 12:15 PM

SESSION 03 · ANSIBLE AUTOMATION

Ansible-Driven LLM Deployment

- → Authoring Ansible roles for the full LLM lifecycle: pull, deploy, warm, scale, retire
- → Inventory design for hybrid on-premises clusters (RHEL v10 + IBM HCI nodes)
- → Idempotent playbooks: GPU driver installation, container runtime config, vLLM rollout
- → Hands-on: deploying a vLLM-served LLM end-to-end with a single playbook

12:15 – 1:15 PM — Lunch Break —

1:15 – 2:15 PM

SESSION 04 · OPENSHIFT AI

OpenShift AI — Orchestrating Containerized LLMs

- → OpenShift AI overview: KServe, model registry, serving runtimes, and workbenches
- → Deploying vLLM as a serving runtime on OpenShift AI with GPU node selectors
- → Autoscaling LLM workloads — replicas, queue depth, and GPU utilization signals
- → Multi-model serving and tenant routing for shared on-premises clusters

2:30 – 3:30 PM

SESSION 05 · DATA & RAG

Local Data, Local RAG — Keeping Knowledge On-Premises

- → Building a Retrieval-Augmented Generation pipeline that never leaves the data center
- → On-premises vector stores running in containers (PostgreSQL/pgvector, Milvus)
- → Embedding model serving alongside vLLM — Ansible-managed model registry
- → Compliance patterns: audit logs, prompt redaction, and PII boundaries

3:45 – 4:30 PM

SESSION 06 · CI/CD FOR AI

CI/CD Pipelines for LLM & Container Lifecycle

- → Model promotion pipelines: dev → staging → sovereign production
- → Tekton + Ansible pipelines for image builds, security scans, and rollout
- → Automated regression and evaluation gates for model updates
- → GitOps for declarative LLM deployments on OpenShift

4:30 – 5:15 PM

SESSION 07 · OPERATIONS

Day-2 Operations — Monitor, Secure, Govern

- → Observability for LLMs: latency, tokens/sec, GPU saturation, and cost-per-request
- → Security hardening — sealed secrets, mTLS between services, signed container images
- → Governance: model cards, lineage, approval workflows, and data residency controls
- → Capstone exercise: production-ready RHEL AI stack, fully automated by Ansible

5:15 – 6:00 PM

WRAP-UP

Day 2 Recap, Q&A & Next Steps

- → Summary of the sovereign RHEL AI stack — from foundations to production
- → Feedback session and open participant discussion
- → Resources, certifications, and continued-learning pathways
- → Networking reception with instructors and peers