Red Hat OpenShift AI

Introduction

Agenda

- What it is and advantages
- User and Admin features
- Roadmap
- Extend beyond OpenShift AI (MLOps, ISVs, Partners)
- Example Architectures
- Workshop setup

What is it?

What is OpenShift AI?

- OpenShift AI is an addon to OpenShift
- RHOAI provides AI tooling while OpenShift provides the underlying, kubernetes-based, development platform

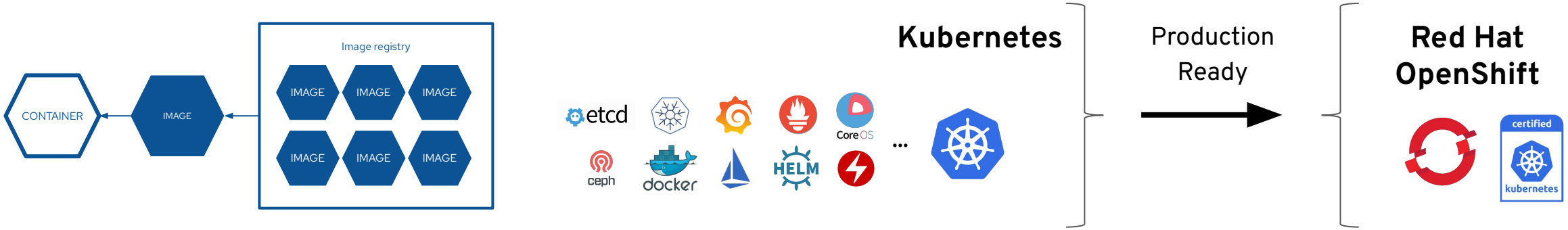
What tooling do you need for AI?

A lot like normal software development:

- Environments for experimentation
- Pipelines for automation
- Version control and artifact tracking for reproducibility and collaboration
- Deployment management for getting things running
- Monitoring for keeping track of how it's running

But you need to deal with data in addition to code, so it all has its own flavor

Containers, K8s, OpenShift for Data Science



Why containers?

- Fewer resources
- Environment isolation
- Quick deployment
- Quick startup/shutdown
- Encapsulation and portability
- Reusability
- Reproducibility

Why Kubernetes?

- Automated rollouts and rollbacks
- Self-healing
- Service discovery and load balancing
- Horizontal scaling
- Designed for extensibility

Why OpenShift?

- Cloud and Infra Agnostic
- GPU Support
- Multi-tenancy
- Zero Trust Security Model
- Metrics and Monitoring
- IAM integration
- Web UI based Workflows

Some AI tools you can run on OpenShift

- Development environments such as Jupyter, VS Code, and R Studio

- Machine learning libraries

- Distributed model training
- Parallelize workloads across nodes and GPUs

Elyra

- AI pipeline editor
- Define workflows through Jupyter

Kubeflow Pipelines

- Machine learning workflow orchestration
- Experiment tracking

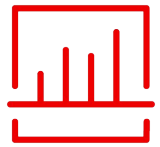
Kserve ModelMesh

- Deploying machine learning models as micro-services
- Pre-built inference servers

The tools that are baked into OpenShift AI

We provide the tools, the use case is up to you

Structured Data Problems



Red Hat OpenShift AI

Starburst

Integrating and processing data at scale

Deep Learning

Red Hat OpenShift AI

NVIDIA

Supported and certified GPU acceleration

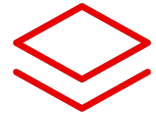
Edge Inference

Red Hat OpenShift AI

Red Hat Device Edge

Deploying machine learning models on small devices

Foundation Models



Red Hat OpenShift AI

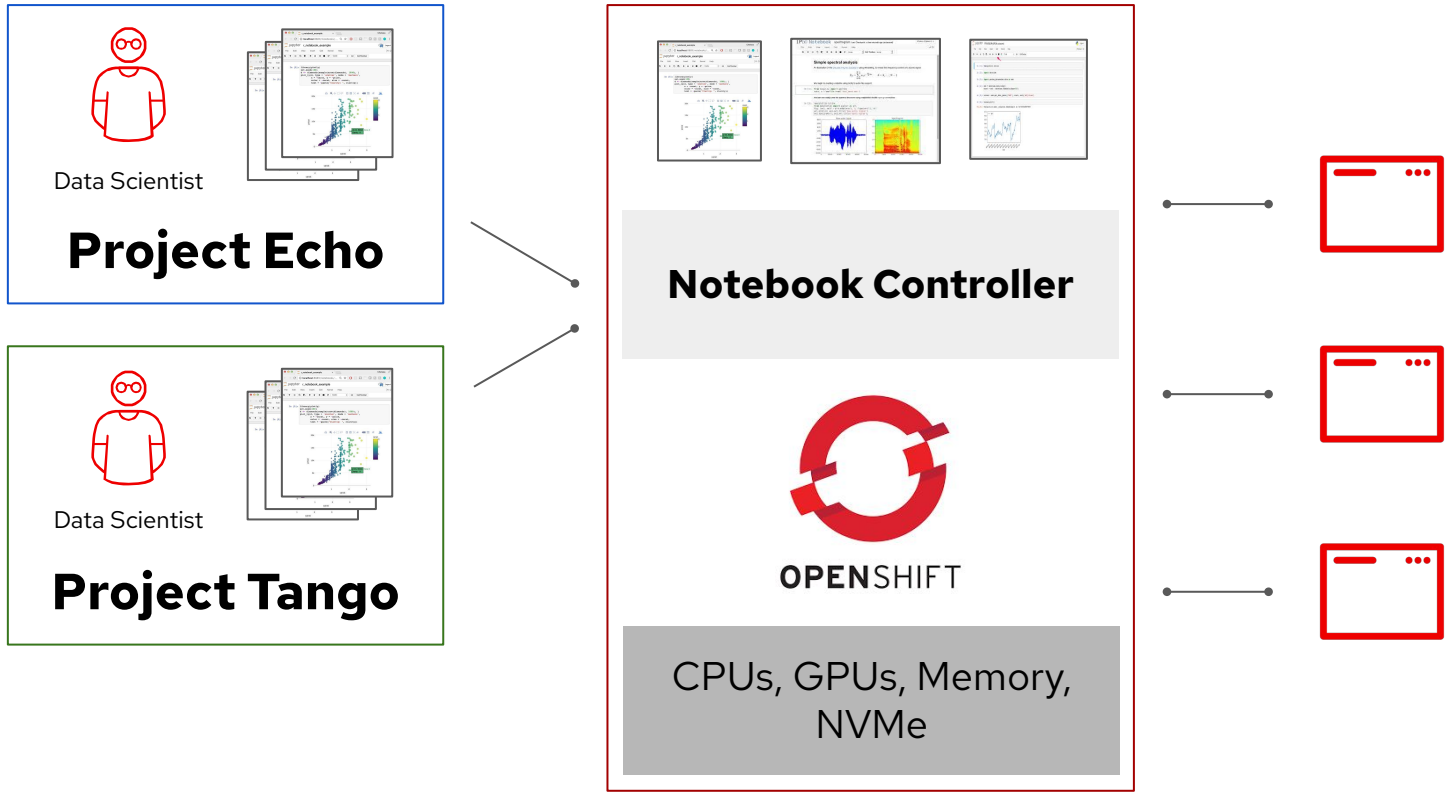
Hugging Face

Distributed training and finetuning of large models

Red Hat

Advantages

A Multi-Tenant Data Science Platform



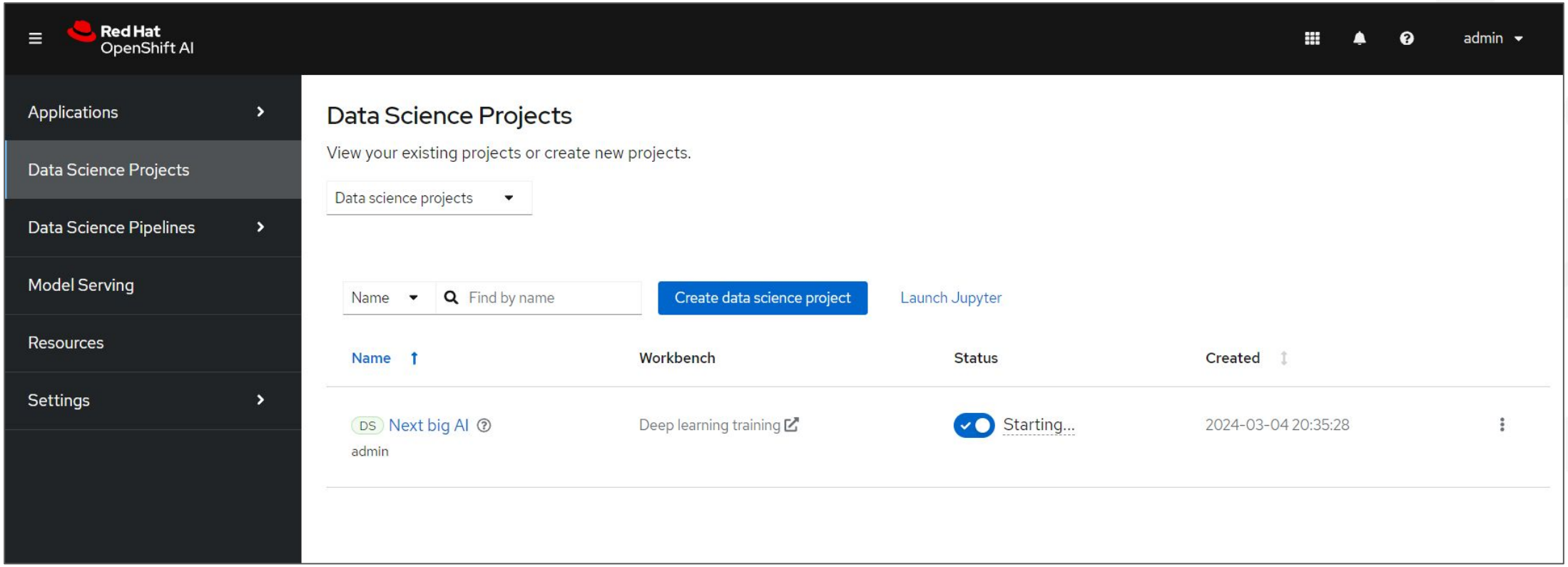
- Identity and RBAC
- Project isolation (data, networks)
- Sharing of resources with protection
- Resource isolation
- Resource quotas
- Priority and pre-emption

- User friendly for quick start and fast prototyping
- Flexibility and depth to cover your use cases as you grow into it
- AI platform that also handles everything you need around AI

User-level features

The Dashboard

Adds an interface for the data science components



Red Hat

Red Hat OpenShift AI - Key features

Model development

Interactive, collaborative UI for **seamless access** AI/ML tooling, libraries, frameworks, etc.

Model serving

Model serving routing for **deploying models to production** environments

Model monitoring

Centralized monitoring for **tracking models performance and accuracy**

Data & model pipelines

Visual editor for **creating and automating** data science pipelines

Distributed workloads

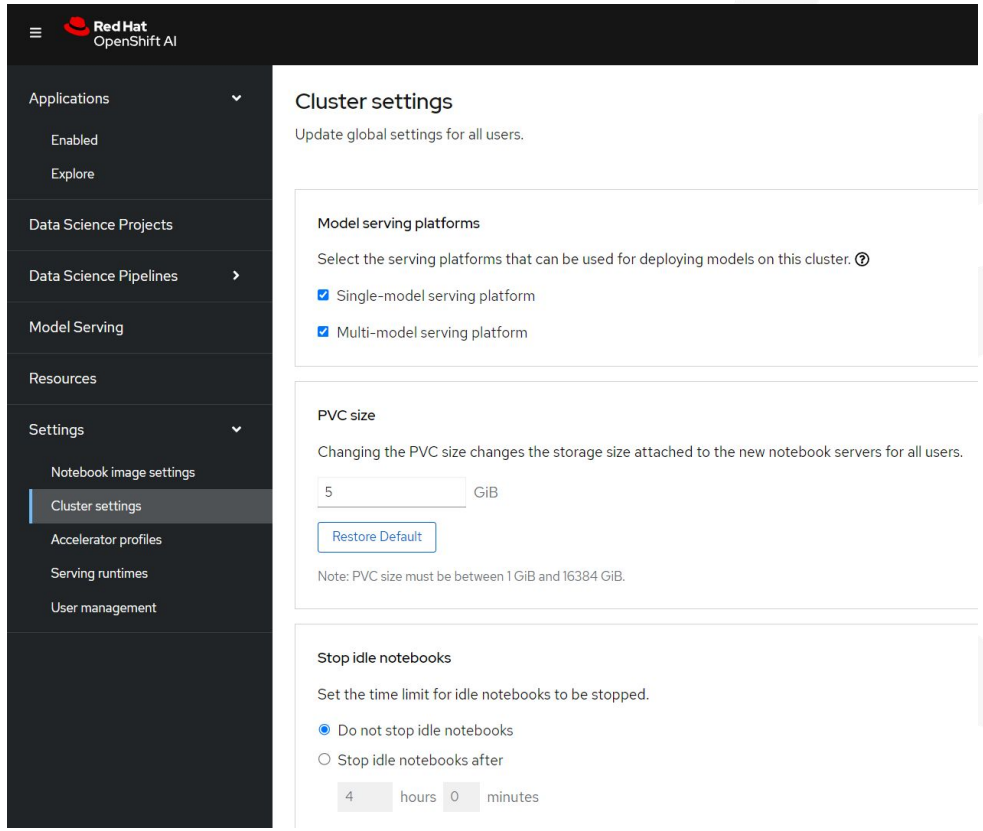
Seamless experience for **efficient data processing, model training, and tuning**

Admin-level features

Easy customization

RHOAI admin panel allows easy access to:

- Custom environments
- Cluster resource usage management
- Hardware acceleration
- Custom deployments
- User management



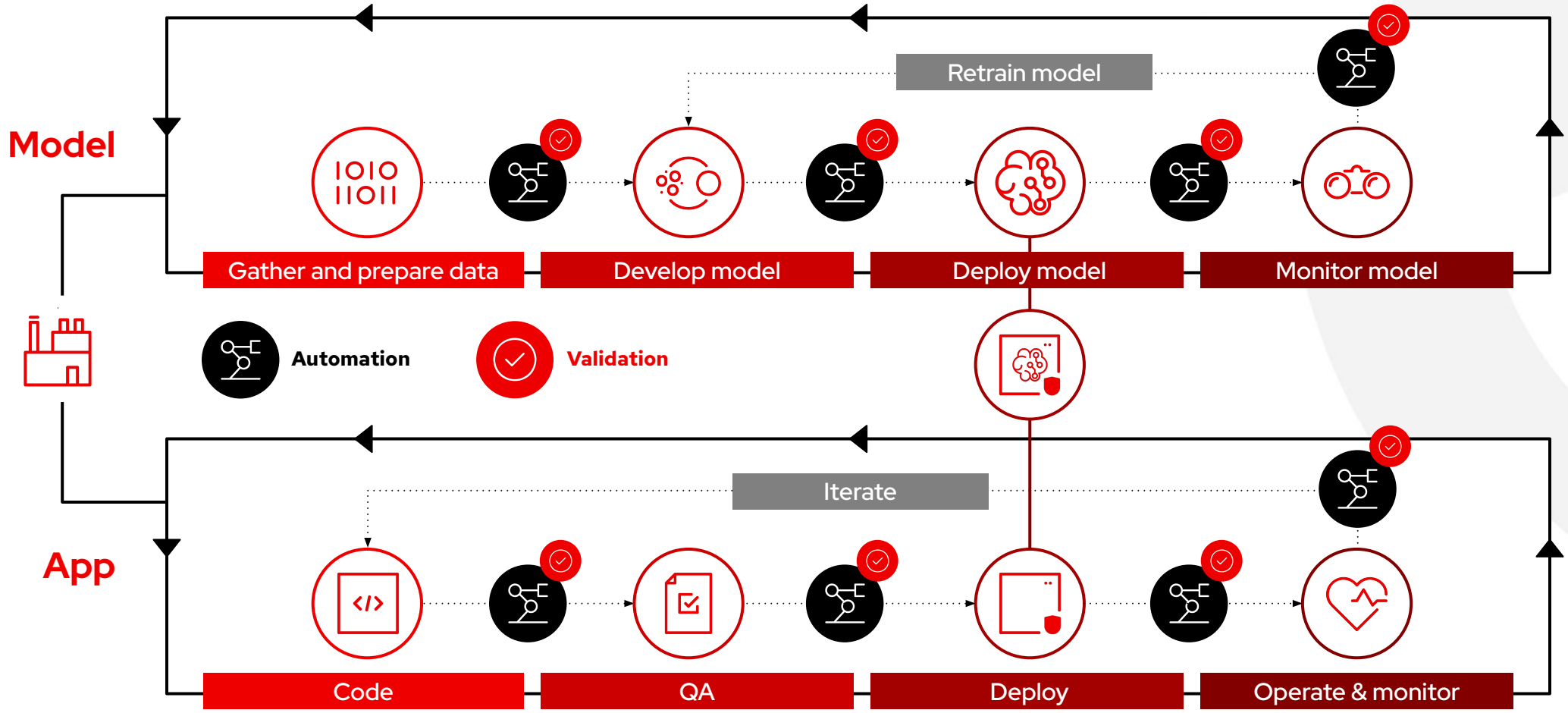
Red Hat

Depth of customizability

Every component in RHOAI is reflected in OpenShift and can be managed through yaml configs and be automated.

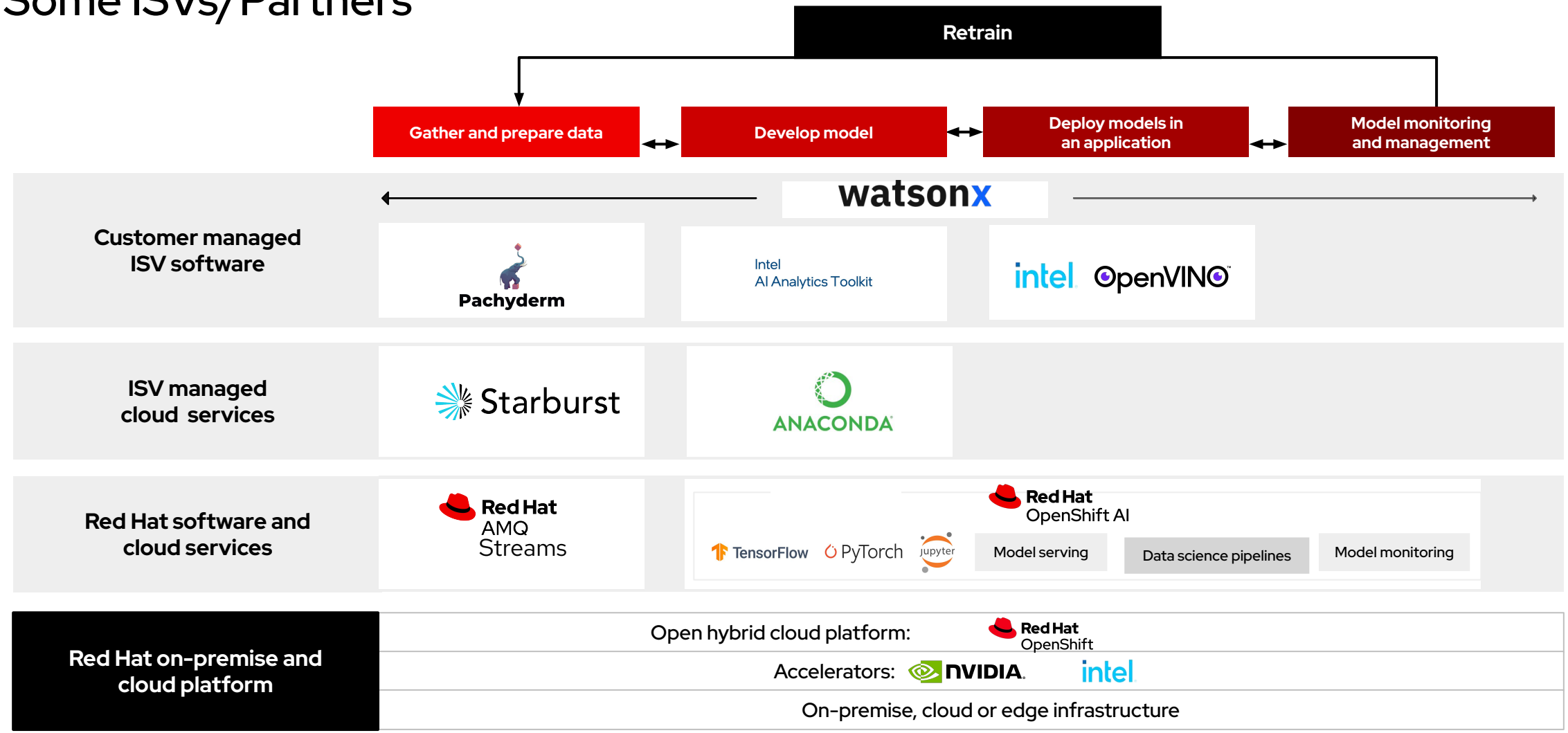
AI in the larger OpenShift picture

Lifecycle for operationalizing models



ISVs and Partners

Some ISVs/Partners

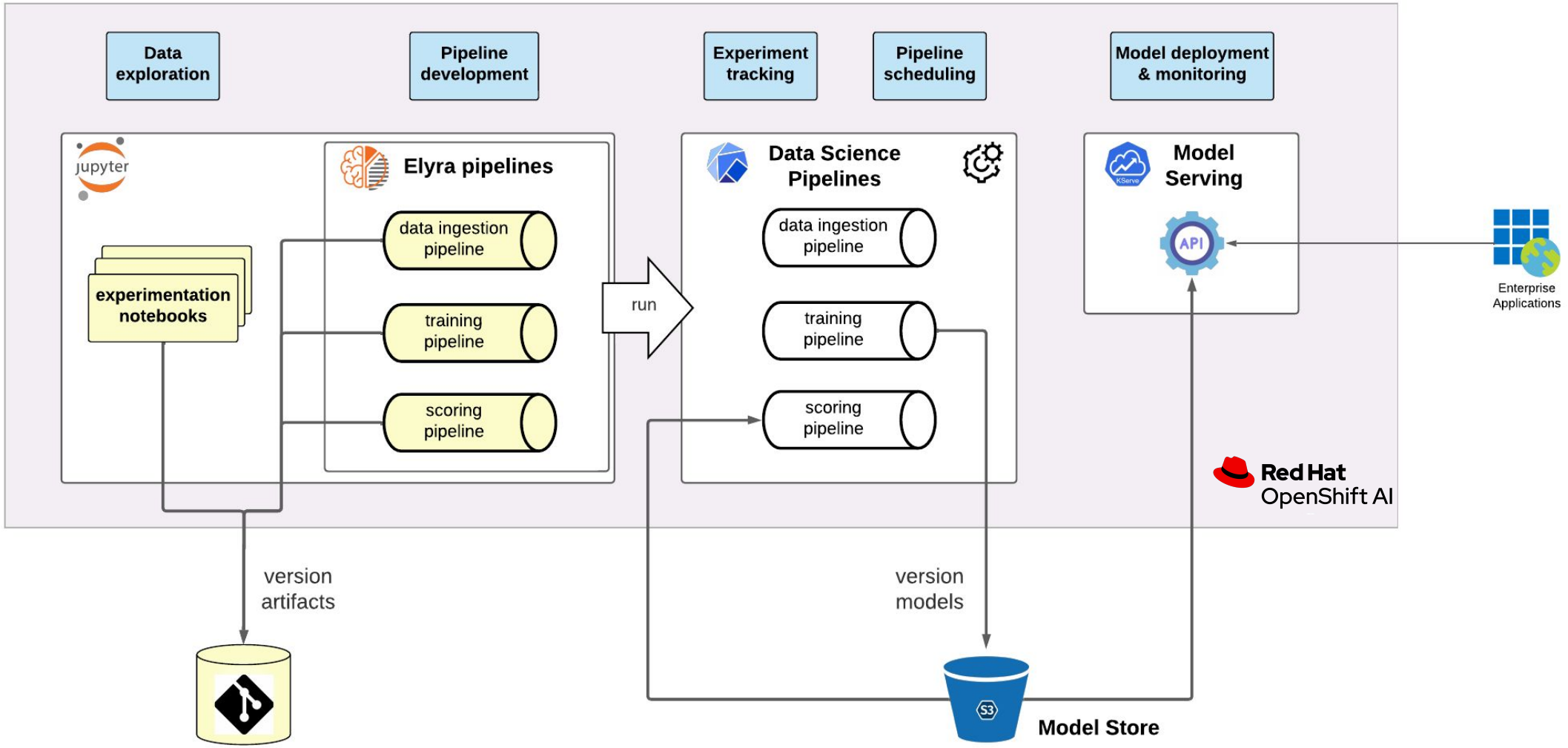


Red Hat

Example Architectures

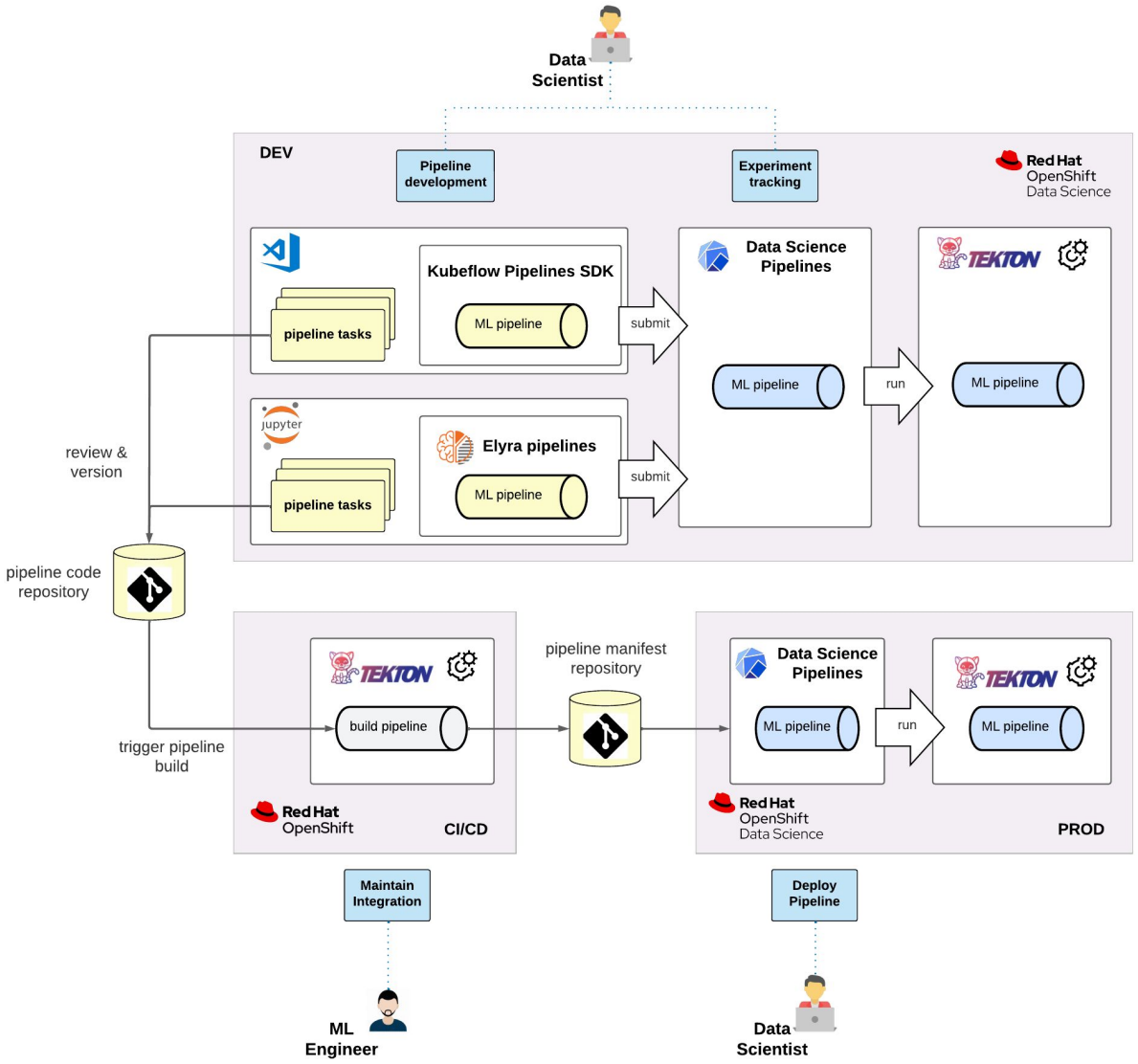
Red Hat

ML workflow with RHOAI



Red Hat

ML Pipelines at Insurance Company



Red Hat

Workshop setup

Red Hat

About the training

- 1 environment with 60 users
- Created for training purposes - not production
- For every hands-on section, we have some slides to explain the theory

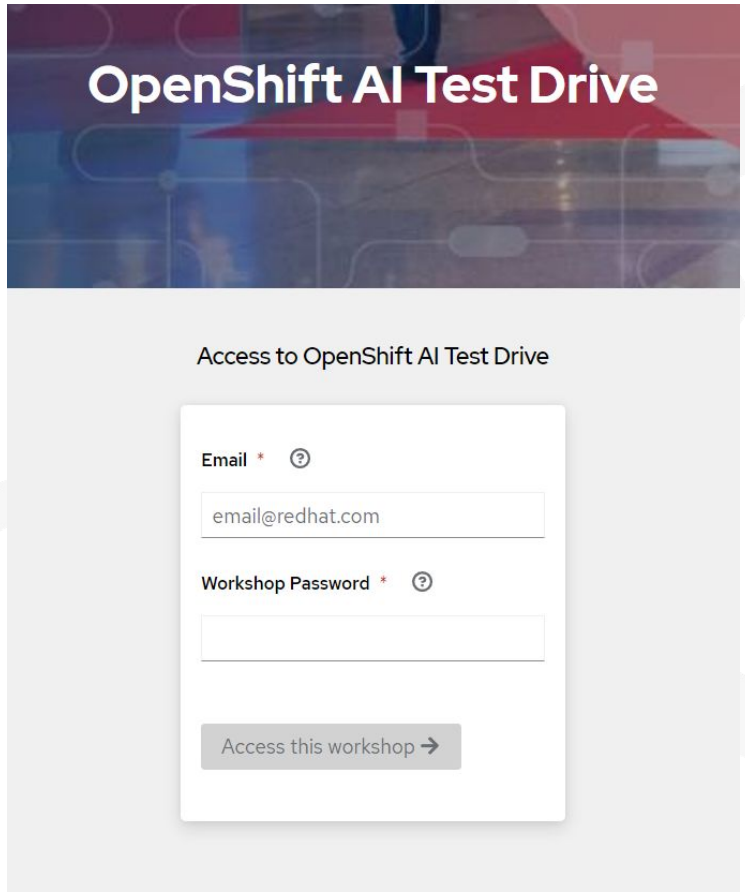
Agenda

Day	Time	Section	Content / topics	Attendees Profile
1	9-10 am	RHOAI Introduction	• Why AI/ML on OpenShift • RHOAI features • Partners • Roadmap	RHOAI Users + Platform Admins
1	10-12 am	RHOAI Getting Started	• Create projects • Workbenches • Data connections • Serve models	RHOAI Users + Platform Admins
1	1-4 pm	RHOAI Advanced	• Using custom notebooks • Pipelines • LLM serving • Recommended Practices	RHOAI Users + Platform Admins
1	4-5 pm	Extra	Q&A or finishing up any outstanding items	RHOAI Users + Platform Admins
2	9-11 am	RHOAI Admin Introduction & Lifecycle	• RHOAI flavors / Integrations • Components and dependencies • Common patterns • Install RHOAI/Update/Lifecycle/Uninstall	Platform Admins
2	11am-12am 1-4 pm	RHOAI Admin Configuration	• Users and groups • Custom notebook images • Custom serving runtimes • GitOps and lifecycling • GPUs and Accelerator Profiles • Managing RHOAI resources	Platform Admins
2	4-5 pm	Extra	Q&A or finishing up any outstanding items	Platform Admins

Red Hat

How to get access to the environment

- Use this link: https://demo.redhat.com/workshop/5ck324
- Enter your email and the password we give you.



- You will be navigated to a new screen with a link that looks like this:

Lab User Interface https://showroom-showroom-[redacted]-user1.apps.[redacted].opentlc.com/

- Note down the user (user1 in this case), this is your user.
- Next, access the environment through the link we post in the chat.

- Raise your hand in Teams when you are done.

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat

Red Hat