

# Introduction to the Tools

## Getting Started

Humair Khan  
Senior Software Engineer

# Everything is Open Source!

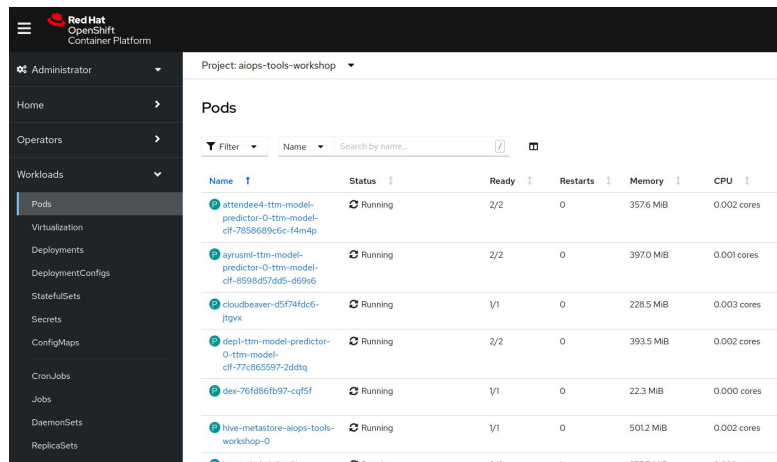
---

- All applications and services use open source software
- You too can deploy this workshop
- You will need a running instance of OpenShift
- Operate First Community Cloud also as long running OpenShift clusters that run all these applications for the general public to experiment with.

Everything is deployed on  
Openshift

# OpenShift

- An enterprise-ready Kubernetes container platform built for an open hybrid cloud strategy
- Sits on top of Kubernetes
- Interactions with OCP is limited in this Workshop



The screenshot displays the Red Hat OpenShift Container Platform console interface. The left sidebar shows the navigation menu with options: Administrator, Home, Operators, Workloads, Pods (selected), Virtualization, Deployments, DeploymentConfigs, StatefulSets, Secrets, ConfigMaps, CronJobs, Jobs, DaemonSets, and ReplicaSets. The main content area shows the 'Pods' page for the 'Project: aiops-tools-workshop'. It includes a search bar and a table of running pods.

Name	Status	Ready	Restarts	Memory	CPU
attende4-ttm-model-predictor-0-ttm-model-cif-7858689dc6-14m4p	Running	2/2	0	3576 MiB	0.002 cores
ayrusml-ttm-model-predictor-0-ttm-model-cif-8598d57ed5-d6996	Running	2/2	0	3970 MiB	0.001 cores
cloudbeaver-d5f74dc6-jtgvx	Running	1/1	0	228.5 MiB	0.003 cores
depl-1-ttm-model-predictor-0-ttm-model-cif-77c865597-2ddtq	Running	2/2	0	393.5 MiB	0.002 cores
dek-76fd86fb97-cqf5f	Running	1/1	0	22.3 MiB	0.000 cores
hive-metastore-aiops-tools-workshop-0	Running	1/1	0	501.2 MiB	0.002 cores

# Open Data Hub

open source AI tools for running large and distributed AI workloads on  
OpenShift Container Platform

# ODH Components

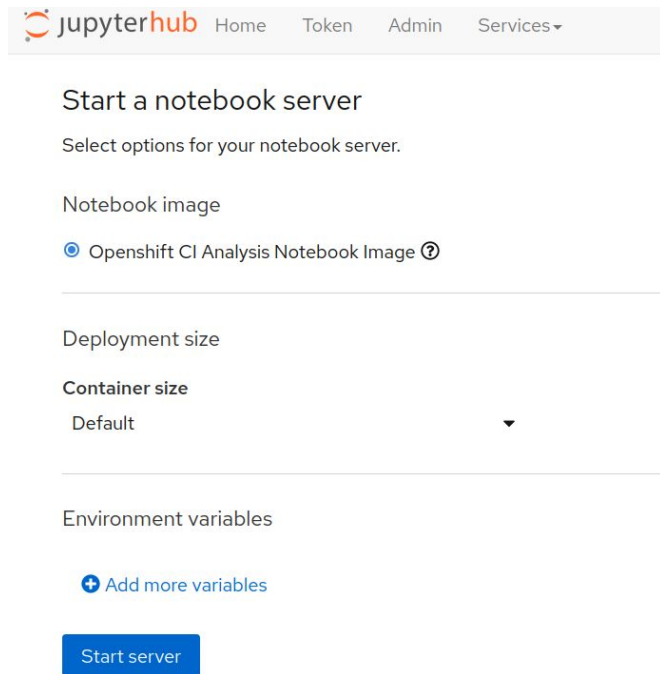
---

- JupyterHub
- Superset
- Trino
- Seldon
- KubeFlow Pipelines



# JupyterHub

- A central hub to manage Jupyter Notebooks on the Cloud
- ODH Integrates JupyterHub with OCP
- Notebooks as Kubernetes Pods
- Admin / User interface
  - Access Control
- Easily supports adding new Notebook images and environments



The screenshot shows the JupyterHub web interface. At the top is a navigation bar with the JupyterHub logo and links for Home, Token, Admin, and Services. The main heading is 'Start a notebook server', followed by the instruction 'Select options for your notebook server.' Below this are three sections: 'Notebook image' with a radio button selected for 'Openshift CI Analysis Notebook Image', 'Deployment size' (empty), and 'Container size' with a dropdown menu showing 'Default'. There is also an 'Environment variables' section with a link to 'Add more variables'. At the bottom is a blue 'Start server' button.

jupyterhub Home Token Admin Services▼

## Start a notebook server

Select options for your notebook server.

Notebook image

☒ Openshift CI Analysis Notebook Image ?

Deployment size

Container size

Default ▼

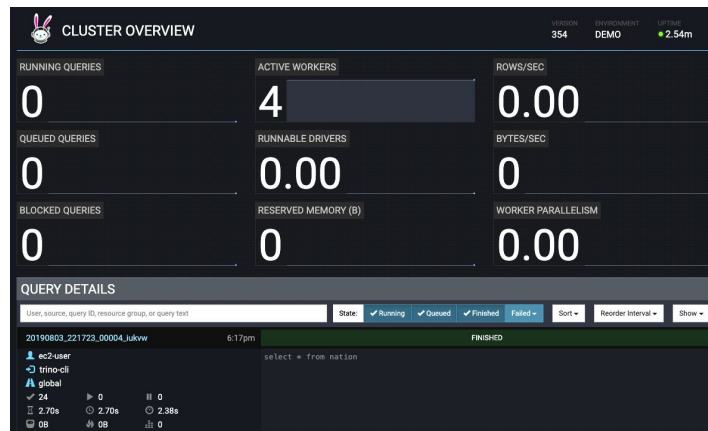
Environment variables

[+ Add more variables](#)

Start server

# Trino

- A distributed SQL query engine designed to query large data sets distributed over one or more heterogeneous data sources.
- It is **not** a general-purpose relational database
- It **is** a tool designed to efficiently query vast amounts of data using distributed queries.





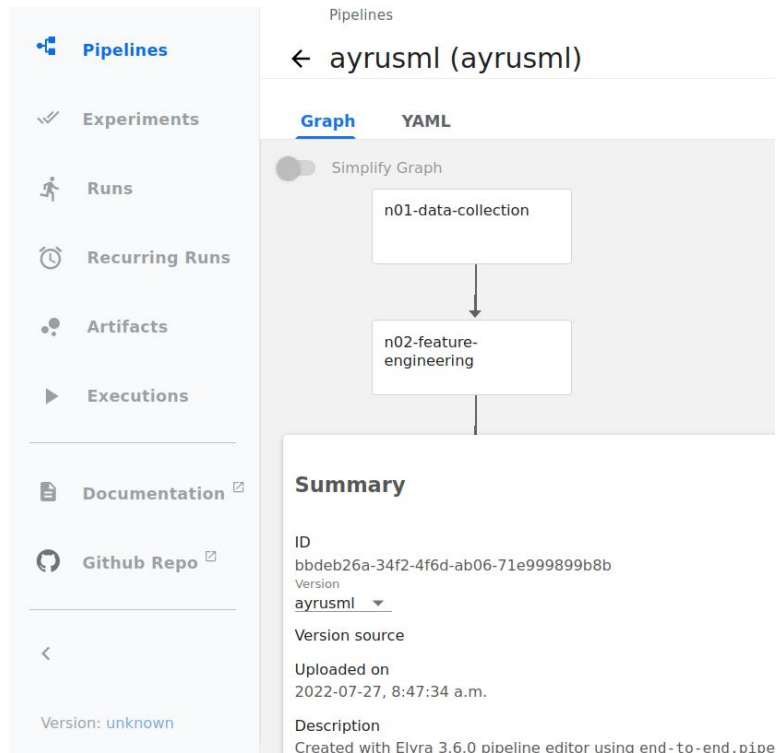
# Seldon

---

- Seldon provides a set of tools for deploying machine learning models at scale.
- This is the one time you will be interacting with Open Shift Directly

# Kubeflow pipelines via Tekton

- A platform for building and deploying portable, scalable machine learning (ML) workflows based on Docker containers.
  - In our case containers within K8s Pods
- Provisioned via Kubeflow directly, but there is incoming support for it via ODH
- We will use kfp to run the training pipeline of the ML model

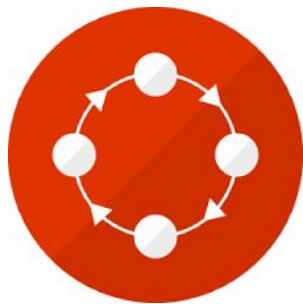


# Superset

- Superset is a modern data exploration and data visualization platform
- A powerful, web-based **SQL Editor** for advanced querying
- ODH integrates Superset with OCP
  - Access control, governed by OCP/K8s rbac
- Out of the box support for nearly any SQL database or data engine
  - Including Trino



## Other Applications



OpenShift Pipelines



MINIO

# Re-create the environment

Instructions can be found here:

[github.com/operate-first/apps/tree/master/workshops/aiops-tools](https://github.com/operate-first/apps/tree/master/workshops/aiops-tools)