A Tale of 2 Data Science platforms - ODH & RHODS

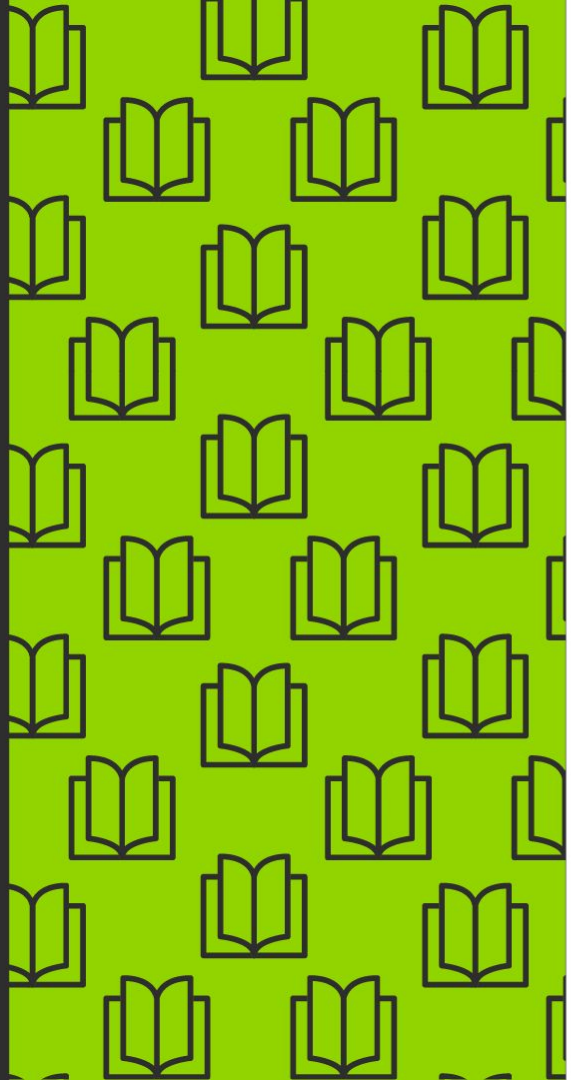
Audrey Reznik
Sr. Principal Software Engineer,
Red Hat OpenShift Data Science team

Speaker

Audrey Reznik
Sr. Principal Software Engineer

Agenda

- Intelligent Applications
  - Personas involved in developing an Intelligent App
- What is ODH?
  - Where can I find ODH?
  - The end-to-end AI/ML Platform
  - ODH Tools & Components
- What is RHODS?
  - The Red Hat Managed Cloud Platform
  - Red Hat Managed Cloud Services
  - ISV Managed Cloud Services
  - Customer Managed ISV Software
- Red Hat Partner Ecosystem

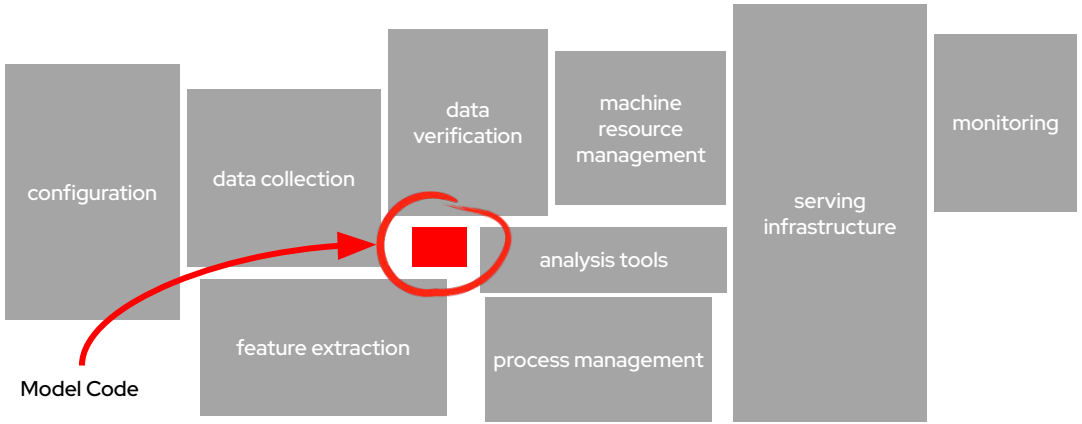


Intelligent Applications

Red Hat OpenShift

A Model's role in Intelligent Applications

The model is a tiny but essential part ...

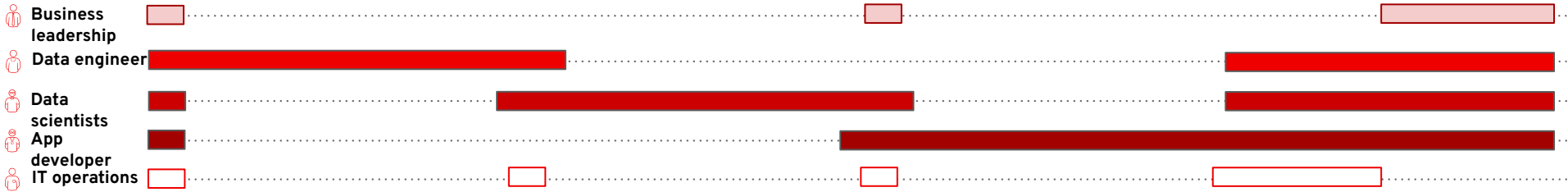


Source:
Adapted from Sculley et al, "Hidden Technical Debt in Machine Learning Systems", NIPS 2015

Red Hat OpenShift

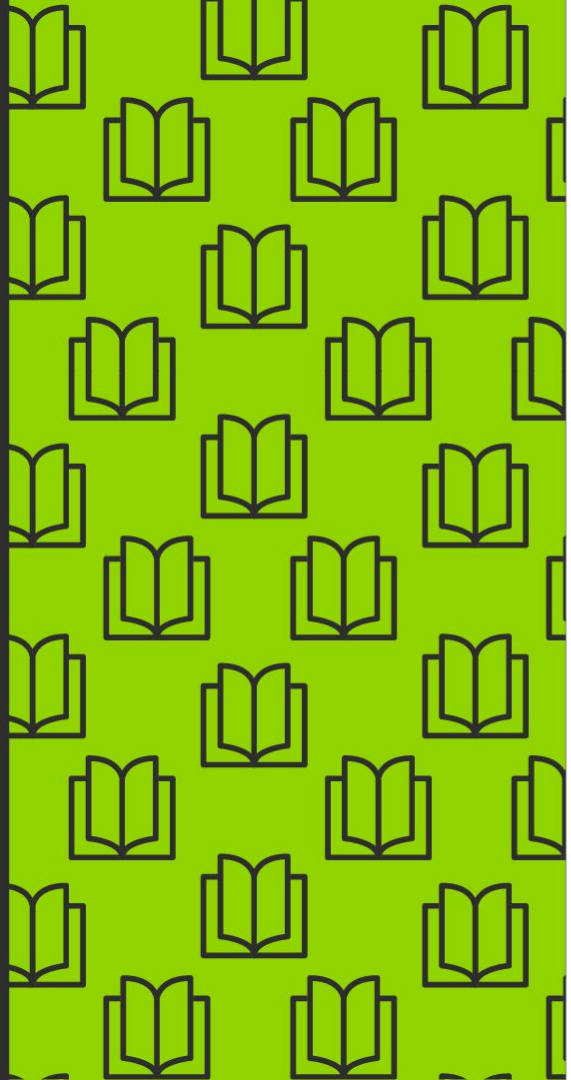
Primary Personas and Tasks across MLOps

Scoping / Objective	Data	Training	Integration	Monitoring
	Data Engineering	**Data Science**	**Continuous Integration & Deployment**	**Monitor / alerts**
	Data Ingestion	Data Splitting	Data Preprocessing	Consumption & optimization metrics
	Data Cleansing	Feature Engineering	App Dev / Heuristics	Satisficing (Gating) metric
	Data Analysis	Model Development	Inferencing Pipeline	Logging & Visualization
	Data Transformation	Model Training	Deployment Targets	Explainability, Interpolation
	Data Validation	Training Optimization	Deployment Patterns	Drift, Decay, Skew, Shift
		Model Validation		Improvements



Disclaimer: All roles may be active, but not primary in phases. Roles may differ between organizations, match roles to tasks above (e.g. data engineer to data).

Red Hat

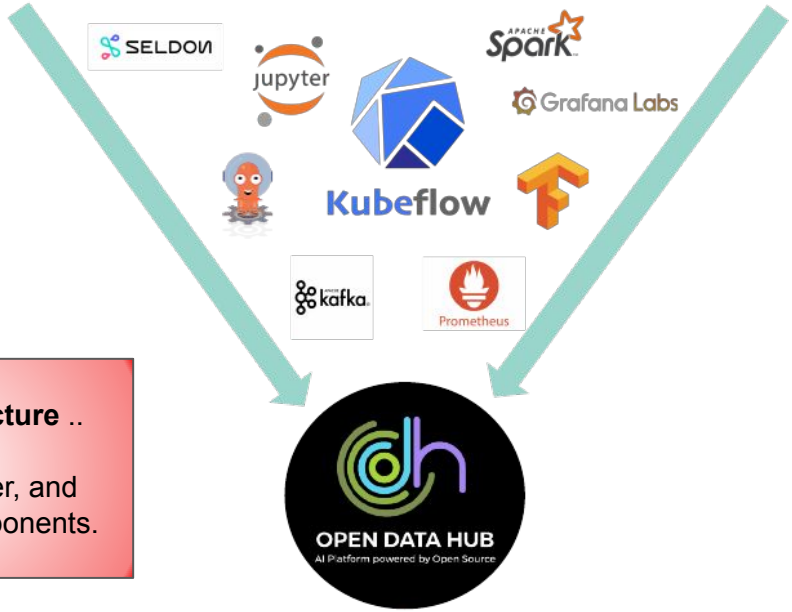


ODH

Open Data Hub

Red Hat OpenShift

Open Data Hub (ODH) is the **Fedora** of Red Hat's AI/ML strategy



Reference architecture .. can be a mix of open-source, partner, and supported RH components.

Platform is supported ->

Platform

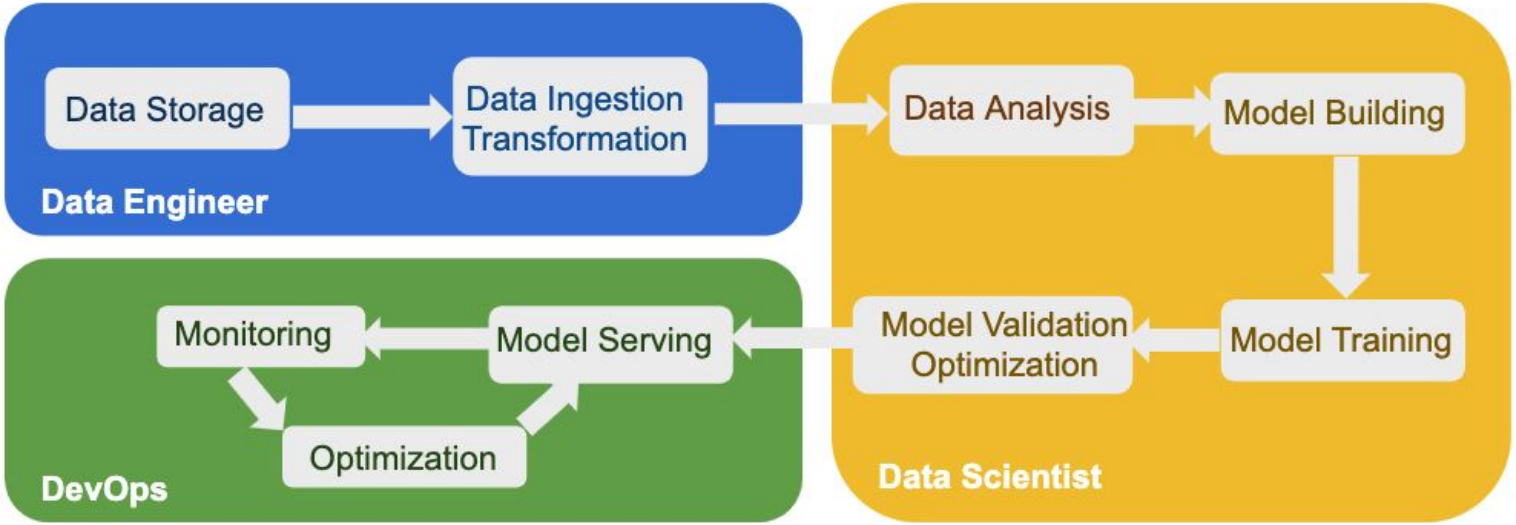
RED HAT® OPENSHIFT Container Platform

Goals

- Provide an end-to-end AI/ML blueprint on OpenShift
- One stop easy operator deployment for the platform on OCP
- Provide tools for each stage in the MLOps lifecycle
- Provide monitoring tools for model and services used by MLOps
- Provide development tools for Data Scientists
- Provide ETL tools used by Data Engineers
- AI/ML pipelines and long processing tasks

Red Hat

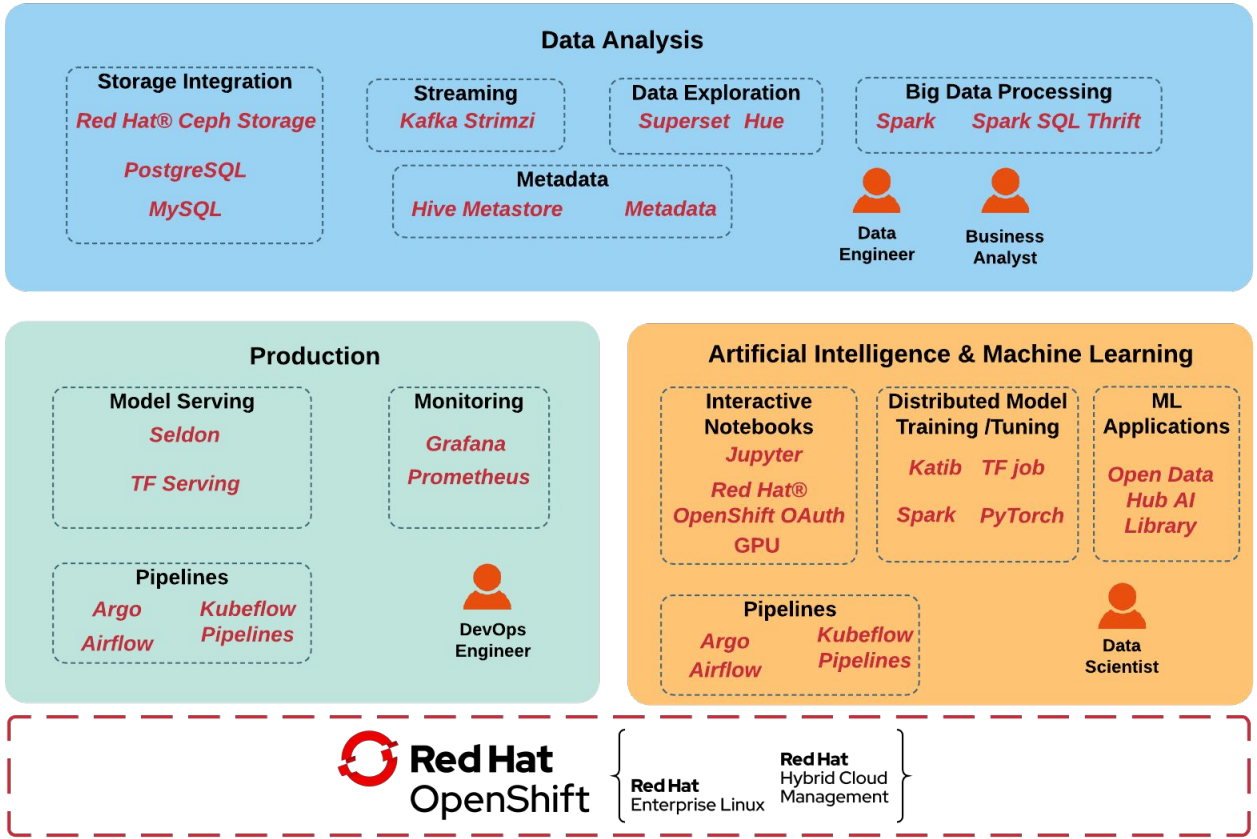
Open Data Hub is an open source community AI/ML platform on Openshift



Red Hat

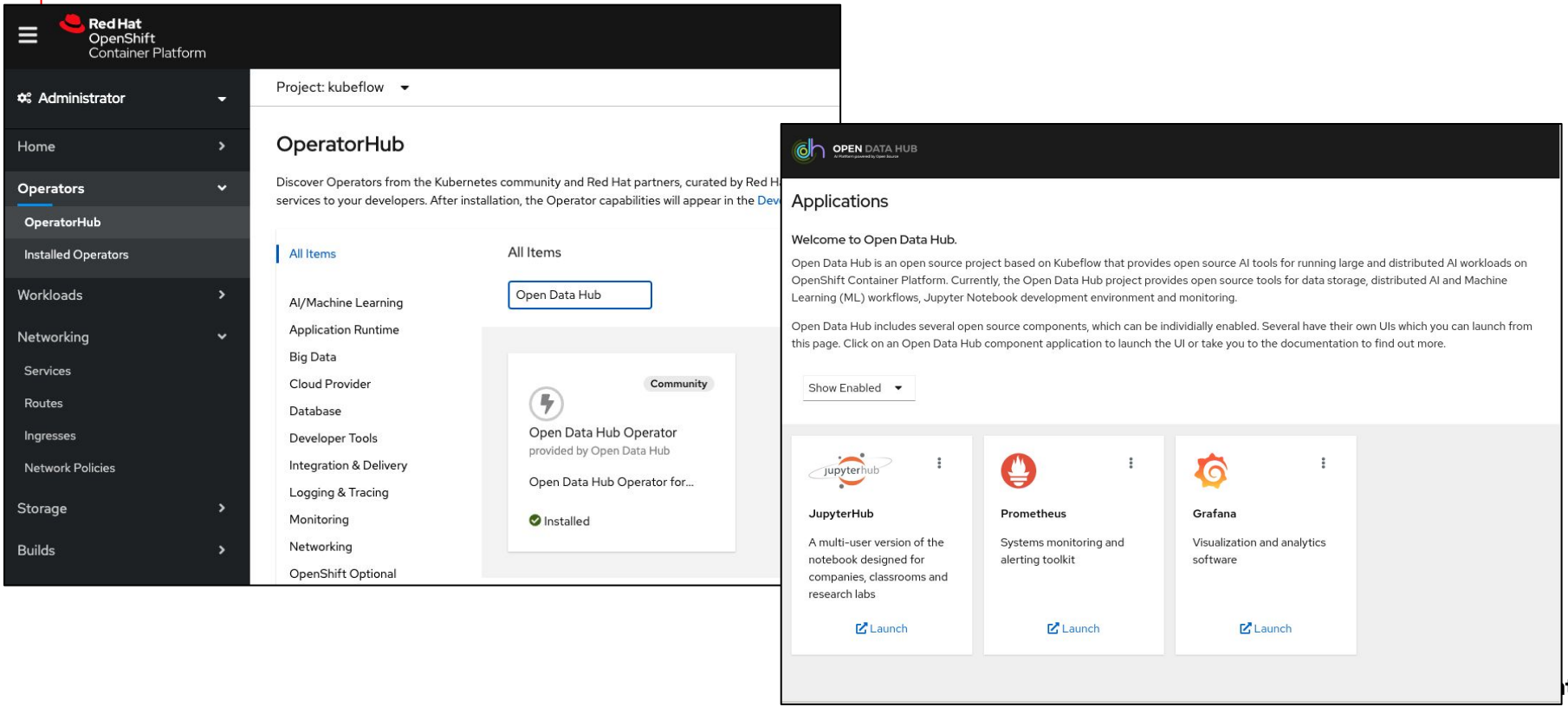
OPEN DATA HUB

AI Platform powered by Open Source

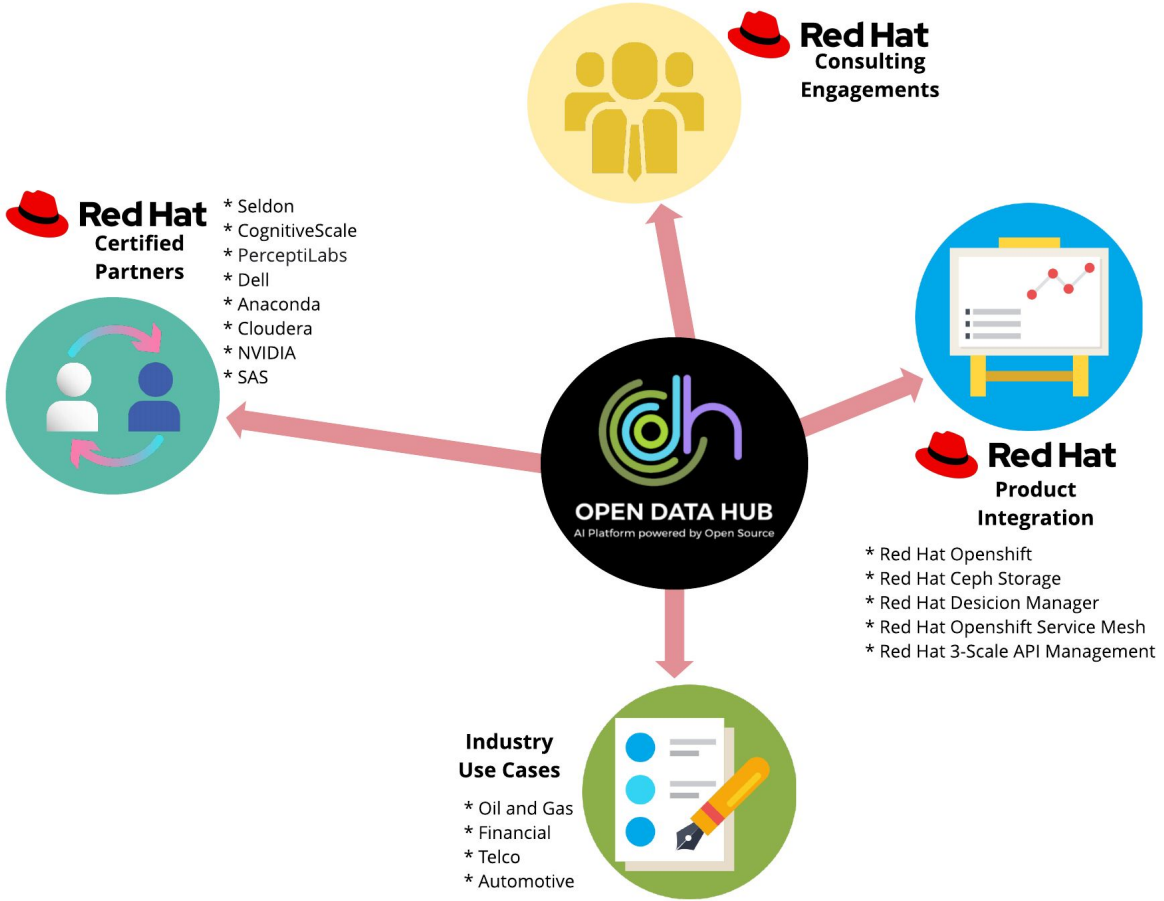


Red Hat

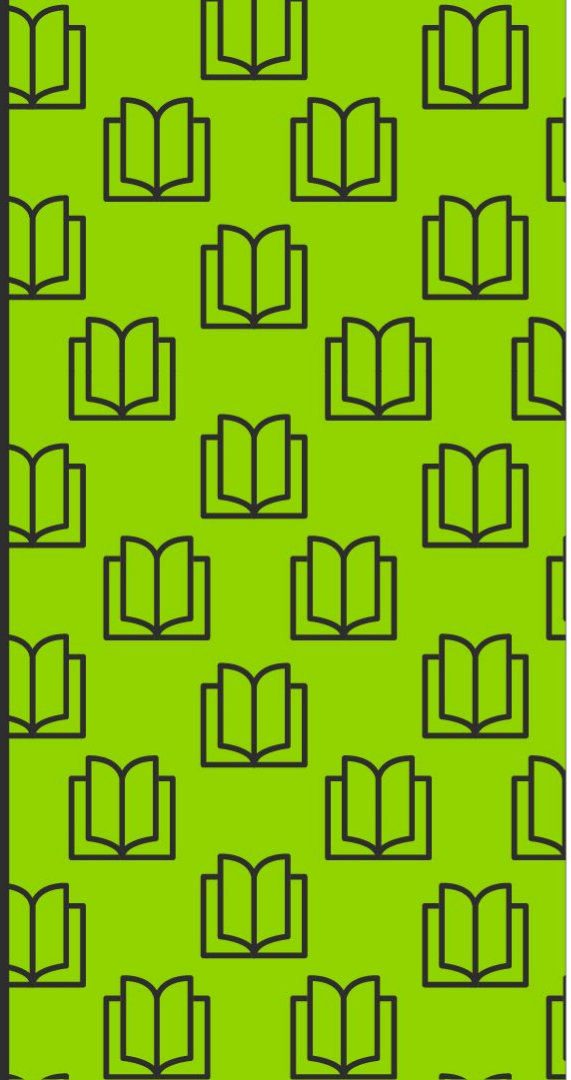
Open Data Hub is an **Operator** installed from OpenShift **OperatorHub**



Open Data Hub comes with an ecosystem provided by **Red Hat** and **Certified Partners**



Red Hat

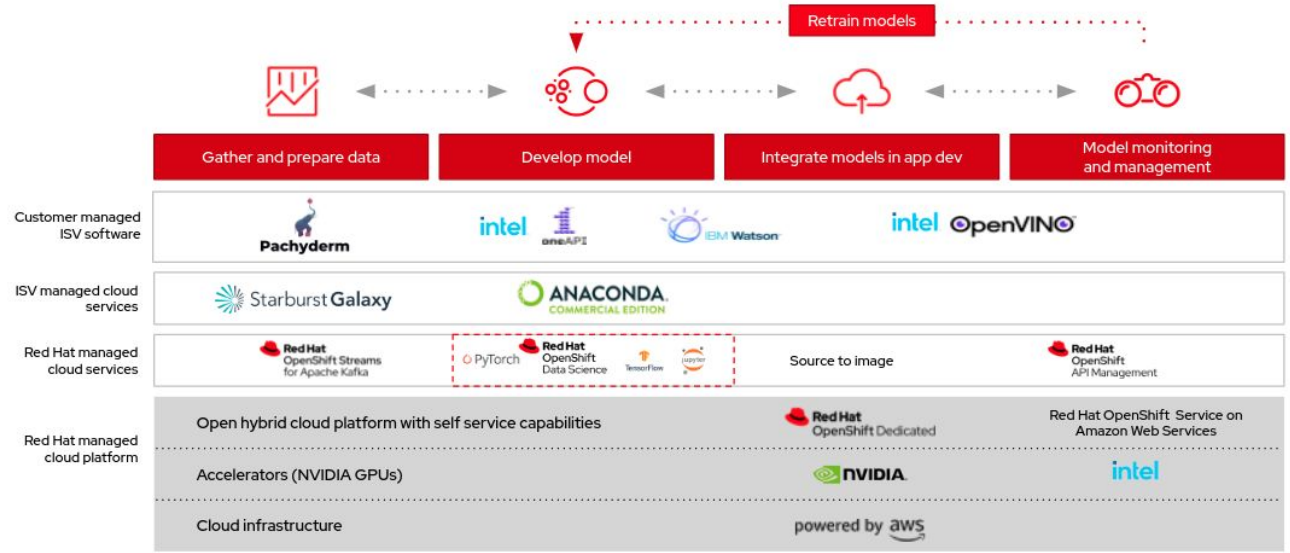


RHODS

Red Hat OpenShift Data Science

Red Hat OpenShift

Red Hat OpenShift Data Science - managed Cloud Service



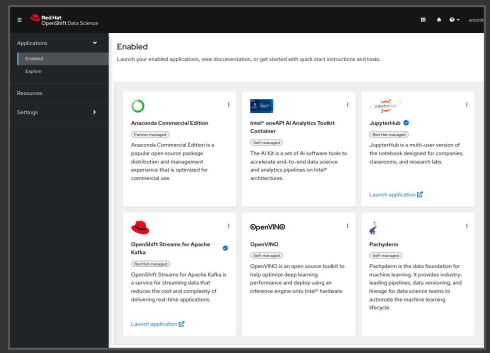
Accelerate Data Science

1. Fully managed, efficient cloud service
2. Tested, supported AI/ML tooling
3. Choose your technology partners
4. Operationalize AI/ML models

Try RHODS: https://developers.redhat.com/products/red-hat-openshift-data-science/overview

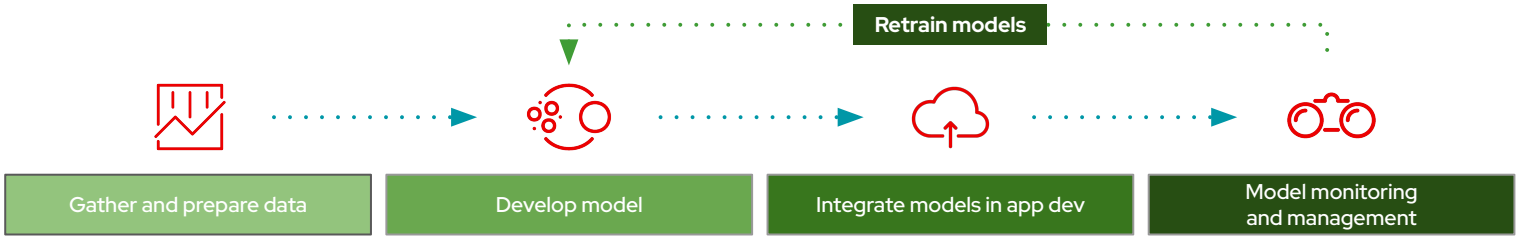
Red Hat OpenShift Data Science

Enabling rapid experimentation and model development on OpenShift allows developers to integrate models into their workflows with fewer obstacles.



Red Hat

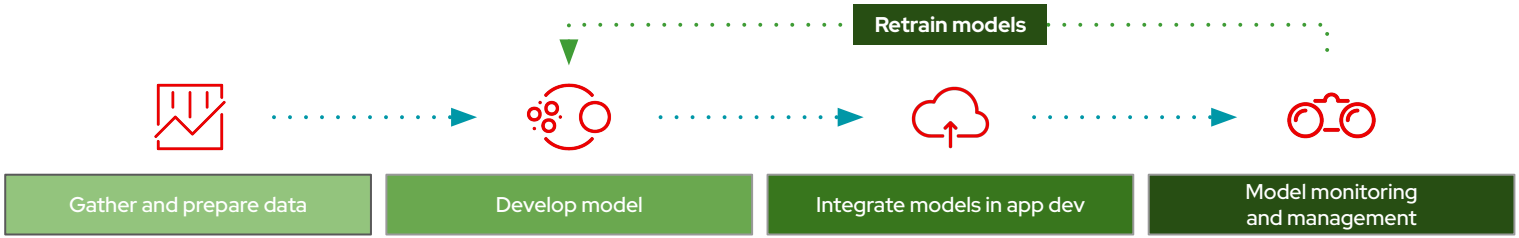
Red Hat Managed Cloud Services



Red Hat managed cloud platform			
	Open hybrid cloud platform with self service capabilities	Red Hat OpenShift Dedicated	Red Hat OpenShift Service on Amazon Web Services
	Accelerators (NVIDIA GPUs)	NVIDIA	intel
	Cloud infrastructure	powered by aws	

Red Hat OpenShift

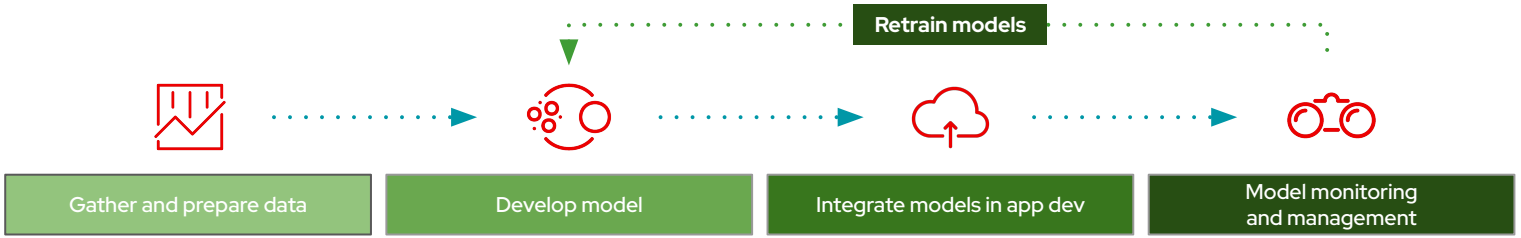
Red Hat Managed Cloud Services



Red Hat managed cloud services	Red Hat OpenShift Streams for Apache Kafka	Red Hat OpenShift Data Science (TensorFlow, jupyter, PyTorch)	Source to image	Red Hat OpenShift API Management
Red Hat managed cloud platform	Open hybrid cloud platform with self service capabilities		Red Hat OpenShift Dedicated	Red Hat OpenShift Service on Amazon Web Services
	Accelerators (NVIDIA GPUs)		NVIDIA	intel
	Cloud infrastructure		powered by aws	

Red Hat OpenShift

Red Hat Managed Cloud Services



ISV managed cloud services

Starburst Galaxy

ANACONDA. COMMERCIAL EDITION

Red Hat managed cloud services

Red Hat OpenShift Streams for Apache Kafka

Red Hat OpenShift Data Science

TensorFlow jupyter PyTorch

Source to image

Red Hat OpenShift API Management

Red Hat managed cloud platform

Open hybrid cloud platform with self service capabilities

Red Hat OpenShift Dedicated

Red Hat OpenShift Service on Amazon Web Services

Accelerators (NVIDIA GPUs)

NVIDIA.

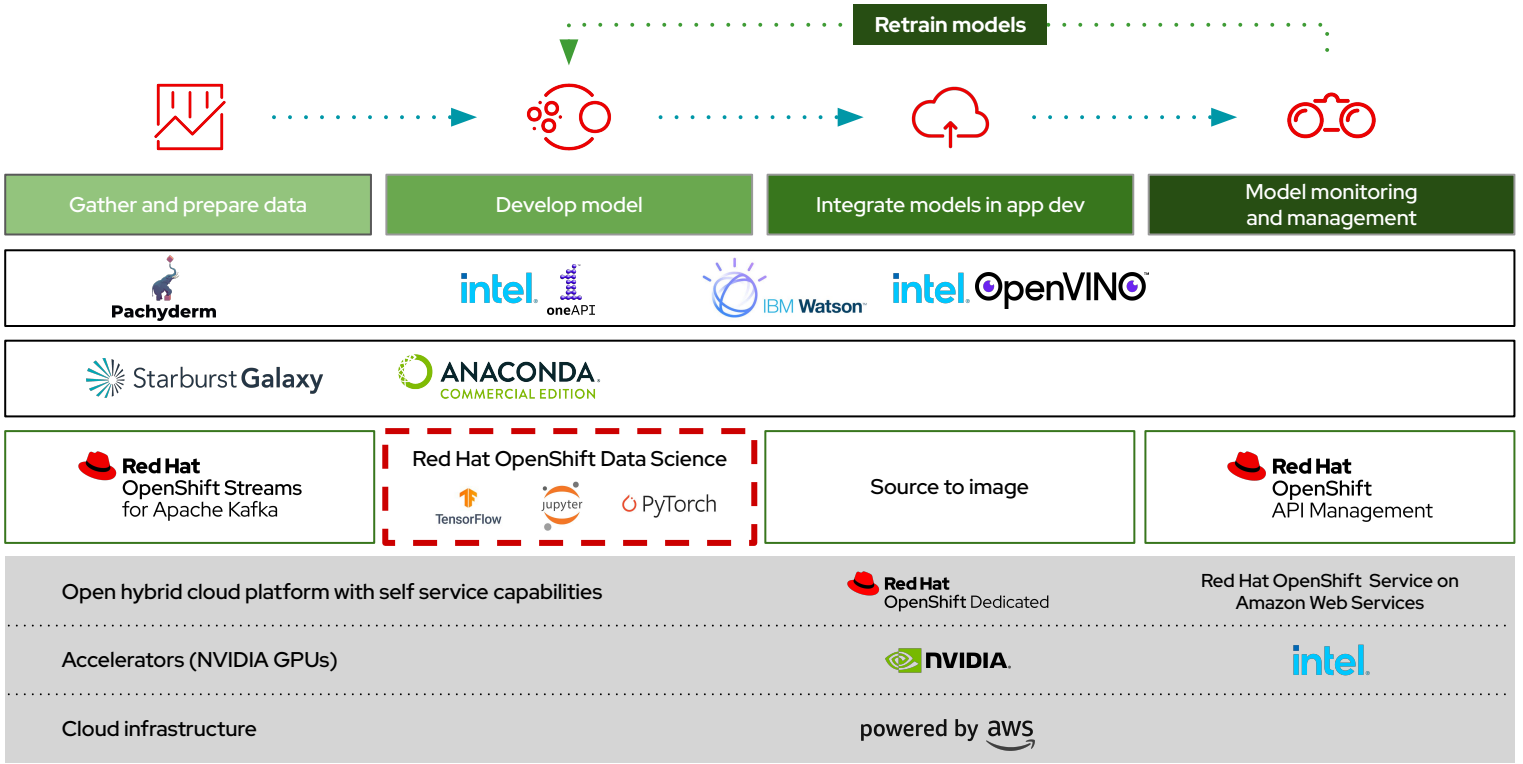
intel.

Cloud infrastructure

powered by aws

Red Hat OpenShift

Red Hat Managed Cloud Services & our Partner Ecosystem



Red Hat OpenShift

Red Hat's Partner Ecosystem

AI/ML life cycle

ANACONDA, CLOUDERA, cnvrg.io, CognitiveScale, C3.ai, DOMINO, GIGASPACES, H2O.ai, IBM, Labs.ai, NVIDIA, PerceptiLabs, SAS, UBIX

Data governance & security

ANACONDA, CLOUDERA, CognitiveScale, Lightbend

Data processing

CLOUDERA, C3.ai, H2O.ai, Lightbend, PerceptiLabs, Starburst

Data analytics

CLOUDERA, SAP, SAS

Databases

Couchbase, crunchy data, elastic, IBM, Microsoft, mongoDB., redislabs, SingleStore

AI Ops

avanseus, PROPHETSTOR, Lightbend

Infrastructure partners

DELL, Hewlett Packard Enterprise, intel, PENGUIN COMPUTING, SUPERMICRO

Hardware acceleration

NVIDIA, intel

Red Hat

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

youtube.com/openshift

facebook.com/OpenShift

twitter.com/openshift

RHODS Public Sandbox: https://developers.redhat.com/products/red-hat-openshift-data-science

RHODS Internal Sandbox: https://red.ht/rhods-internal

Red Hat OpenShift