

Red Hat OpenShift AI

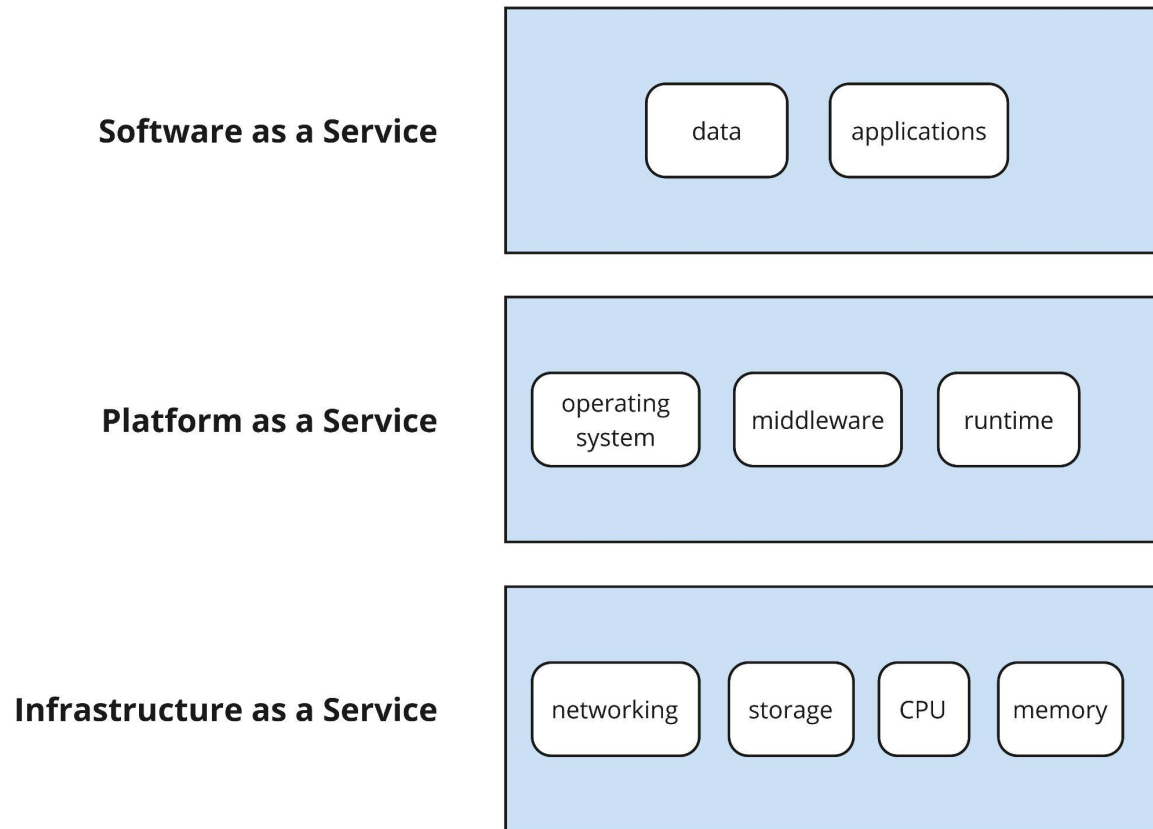
A Hands-On Workshop

Max Murakami
Specialist Solution Architect
mmurakam@redhat.com

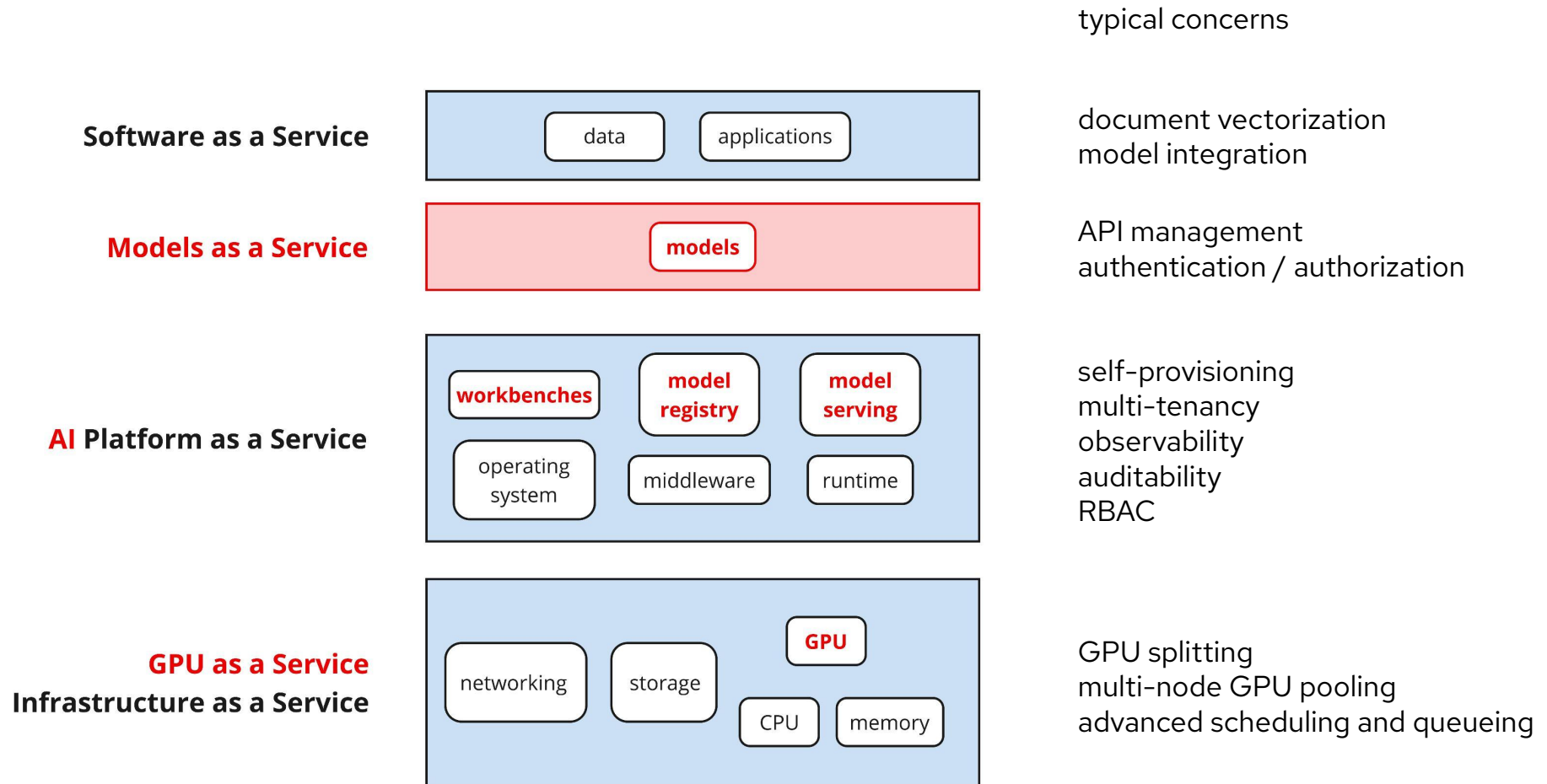


Red Hat has been
building platforms
so you don't have to
for
30 years

Service Layers



AI Service Layers





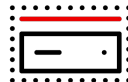
Trusted, Consistent and Comprehensive foundation



Hardware Acceleration



Physical



Virtual



Private
Cloud



Public
Cloud

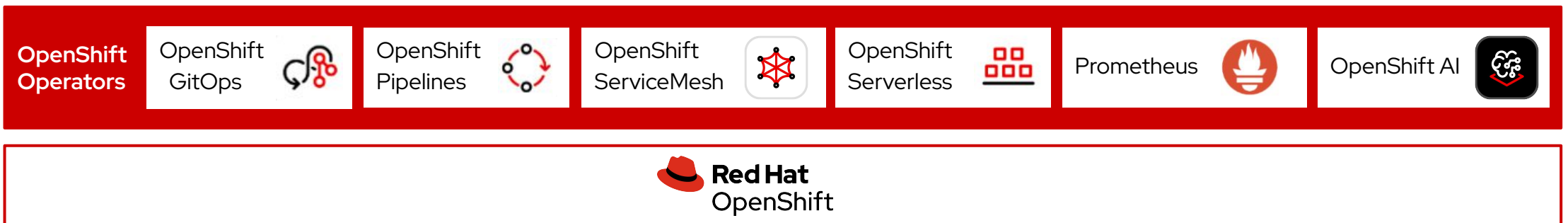


Edge

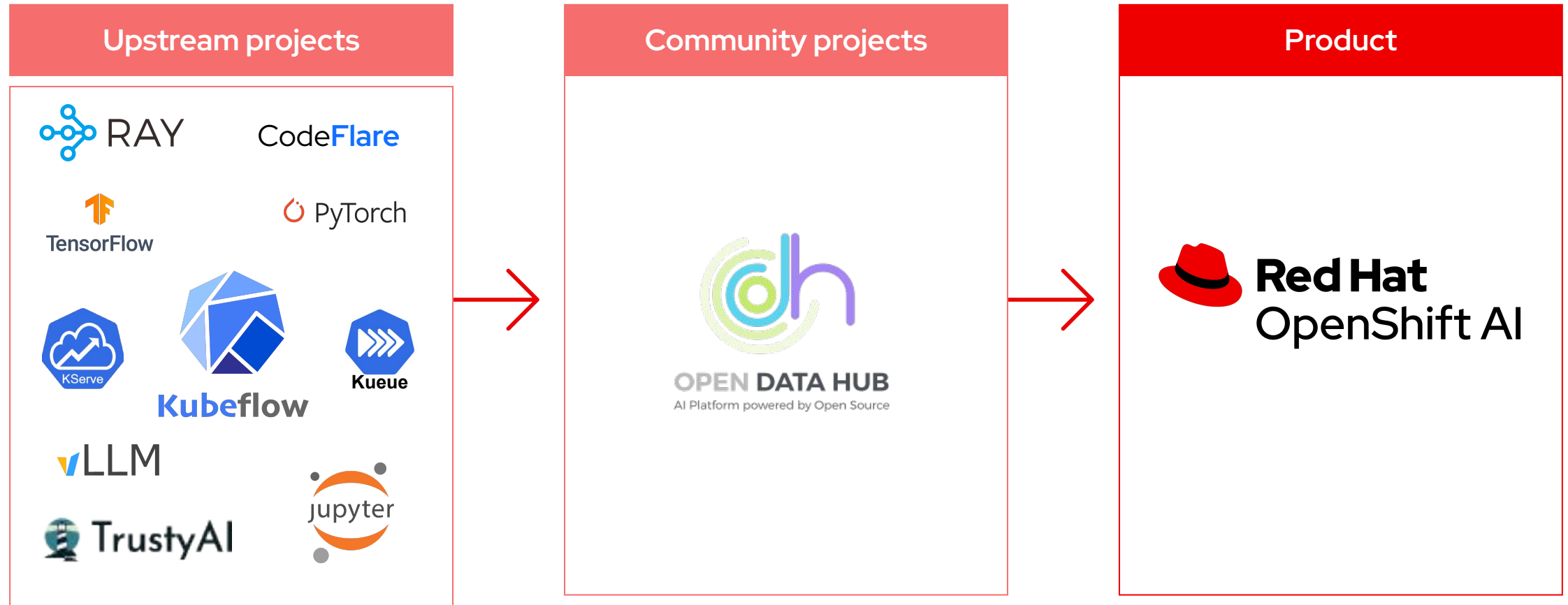
Red Hat OpenShift

A platform for continuous development, integration, and deployment for AI/ML models

- ▶ **One platform for all.** Collaborative environments for dev, data engineers, data scientists and DevOps.
- ▶ **Extend capabilities with operators.** Operators allow bringing AI/ML capabilities to OpenShift.
- ▶ **Hybrid-cloud support.** On-premise support for model development, delivery and deployment.
- ▶ **Support for AI/ML workloads.** Model test, iteration, integration, promotion and serving into production as services.

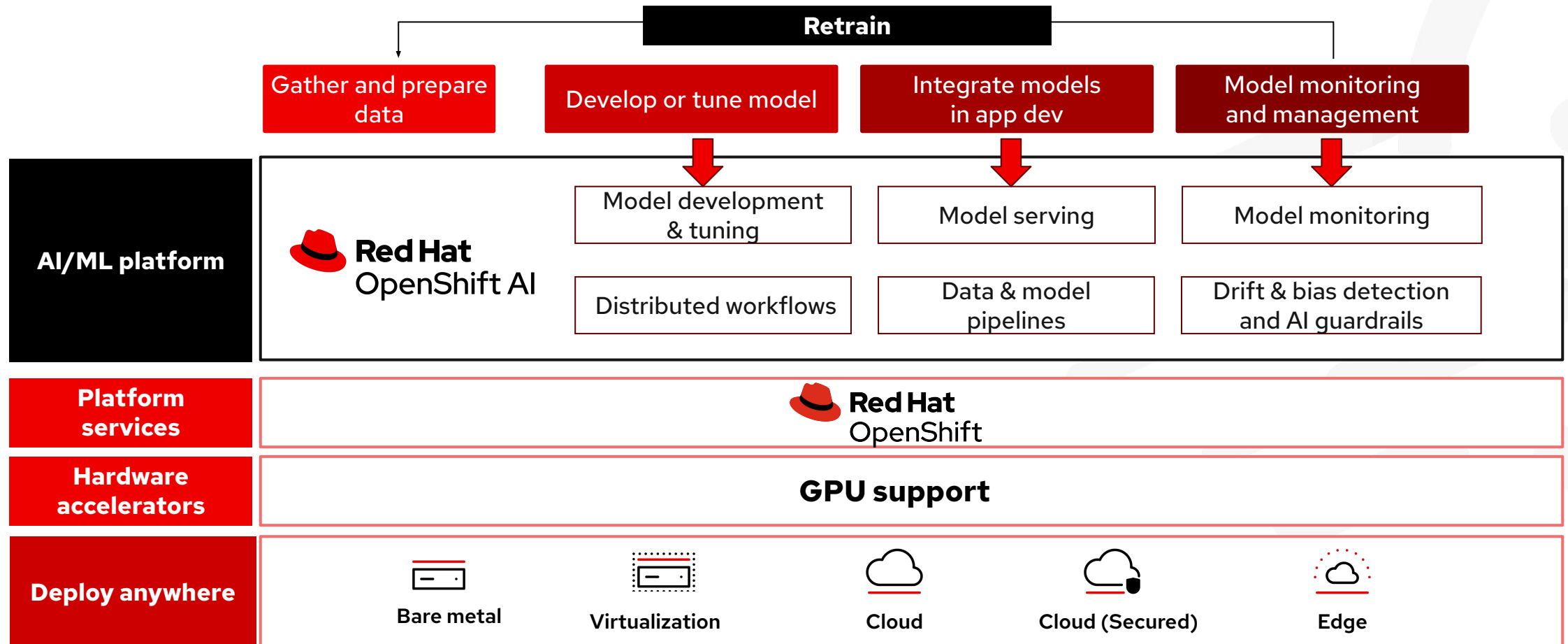


Bringing AI open source to the enterprise

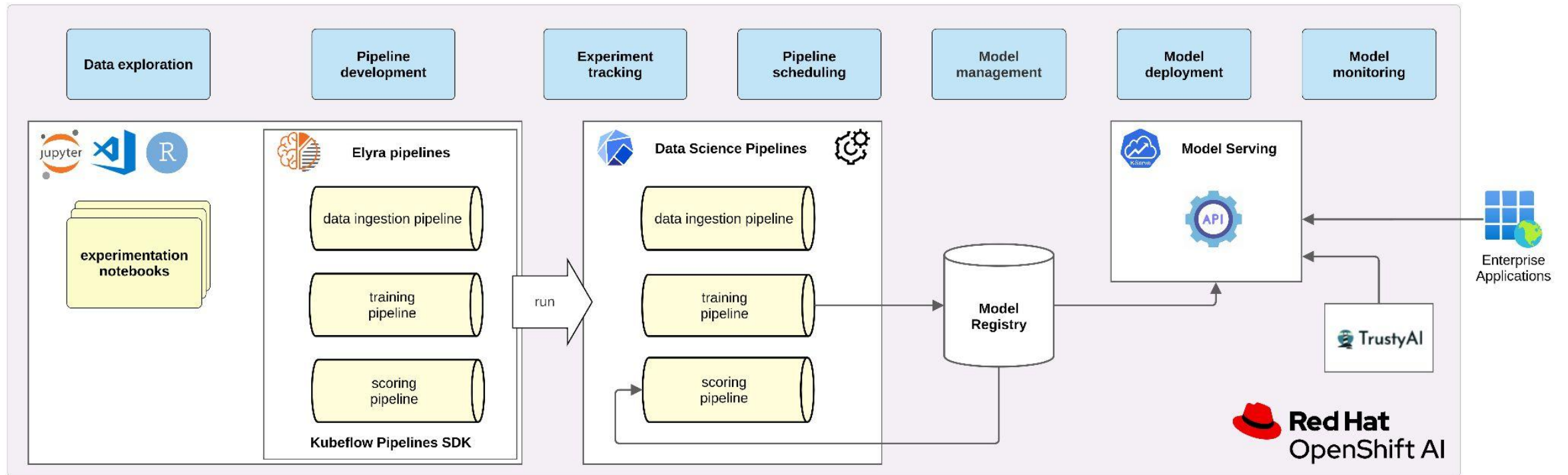


Red Hat OpenShift AI

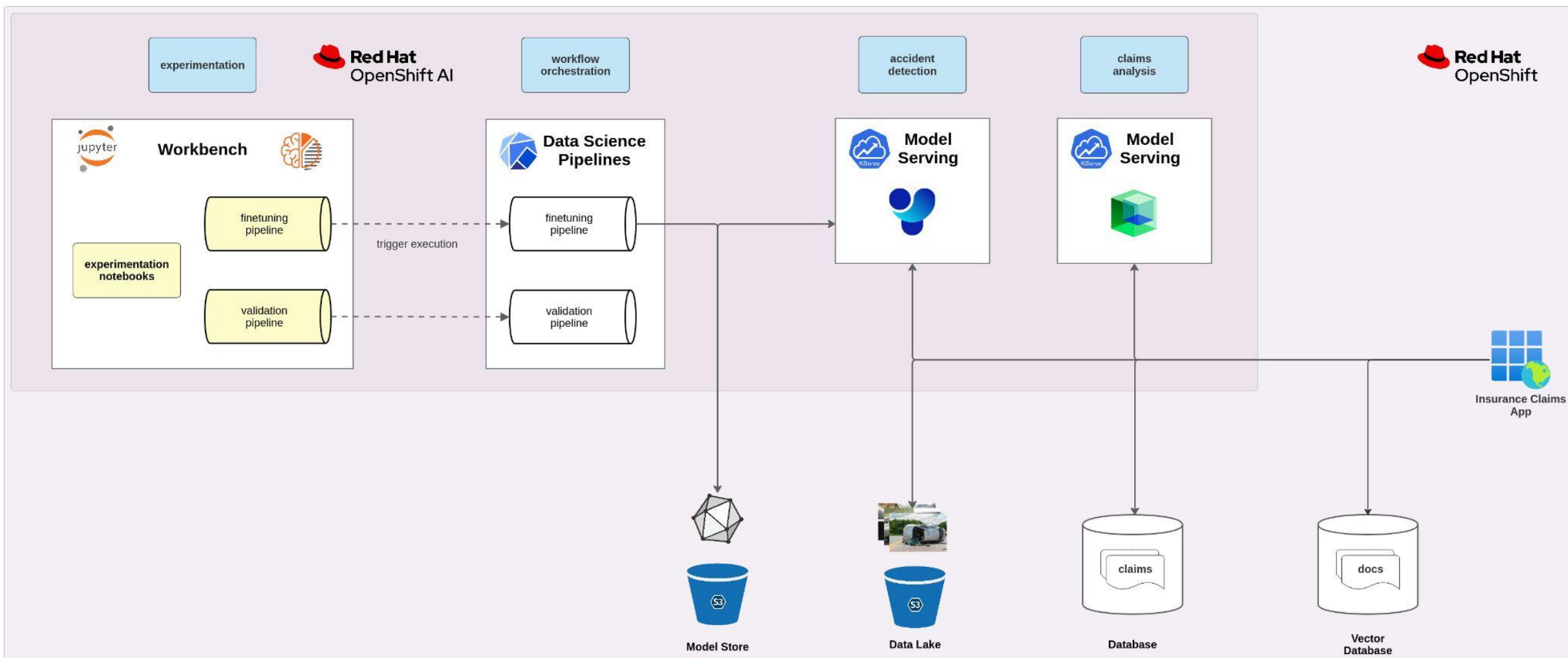
One platform for generative and predictive AI



ML workflow with OpenShift AI



Insurance Claims Processing



Hands-on Lab: AI at Parasol Insurance

The screenshot displays the Parasol Insurance web application interface. On the left is a dark sidebar with navigation links: Dashboard, Policies, Claims, Coverages, Annuities, Subscriptions, Reports, Admin, Settings, and Original App. The main content area shows a claim summary for CLM202401, which has been processed by Sarah Turner. The claim details include the date and time (January 2nd, 2024, at approximately 3:30 PM) and the location of the event (Intersection of Birch Street and Willow Avenue in the city of Evergreen). The summary section describes a car accident where Sarah Turner's black Toyota Camry (DEF-456) collided with Daniel Reynolds' blue Chevrolet Traverse (GHI-789) at a red light. The customer sentiment is noted as polite and professional. On the right side of the claim details, there is a 'Parasol Assistant' chat window. It features a photo of a car accident scene and a conversation where the assistant provides legal advice regarding the right of way at a red light. The chat window includes a text input field and a send button.

URL: red.ht/rhoai-wien-mai25
password: rhoai

Q&A – References

- [OpenShift AI Produktdokumentation](#)
- [ai-on-openshift.io](#) Community Website
- Workshop source code: [Github repository](#) ([lab guide](#))
- llm-d: [article](#)
- verteiltes Training mit Ray: [article](#)
- GPU enablement: [article](#)
- explainable AI mit Trusty AI: [Github](#)

Model Registry

List models

Applications

Enabled

Explore

Data Science Projects

Data Science Pipelines

Experiments

Experiments and runs

Artifacts

Executions

Model Registry

Model Serving

Resources

Settings

Cluster settings

Serving runtimes

Model Registry settings

User management

Registered models

View and manage your registered models.

finance-team-registry Keyword Filter by keyword Register model 1 - 10 of 12 1 of 2

Model name	Labels	Last modified	Owner
Fraud Detection Model A machine learning model trained to detect fraudulent transactions in financial data.	Fraud Detection Machine Learning Financial	Just now	Alice Smith
Customer Churn Prediction Model Predicts the likelihood of a customer churning based on historical data and customer behavior.	Strategic Customer Churn Prediction and Retent... Predictive Analytics	5 minutes ago	Bob Johnson
Credit Risk Assessment Model Assesses the credit risk of loan applicants using machine learning algorithms.	Credit Risk Machine Learning Financial	3 days ago	David Lee
Stock Price Prediction Model Predicts future stock prices based on historical stock data and market trends.	Time Series Analysis Financial	2 months ago	Charlie Brown
Credit Scoring Predicts the creditworthiness of individuals or businesses based on their financial history and other re...	Portfolio Management Credit Score Predictor Risk Assessment 2 more	2 months ago	Michael Johnson
Product Recommendation Recommends products to customers based on past purchases and preferences.	Product Recommendation	3 months ago	Hannah Liu
Investment Portfolio Optimization Model Optimizes investment portfolios to maximize returns and minimize risks.	Portfolio Optimization Investment Strategy Financial	1 year ago	Charlie Brown

Fraud Detection Model

A machine learning model trained to detect fraudulent transactions in financial data.

Actions

VersionsDetails

KeywordFilter by keywordRegister new version

Version name	Last modified	Owner	Labels
v8.0 - Cross-domain trained on data from multiple domains for enhanced generalization	1 minute ago	Joe Doe	-
v7.0 - Adaptive Learning Version of the fraud detection model with adaptive learning capabilities to adapt to changing fraud patter...	3 days ago	Joe Doe	-
v6.0 - Explainable AI using explainable AI techniques to provide insights into model predictions.	5 days ago	Bob Anderson	-
v5.0 - Ensemble Ensemble version of the fraud detection model combining multiple base models for improved accuracy.	1 week ago	Joe Doe	Custom label 1 very very...2 more
v4.0 - Advanced Features Version of Fraud Detection incorporating advanced features and machine learning algorithms	2 months ago	Bob Anderson	-
v3.0 - Real-time optimized for low-latency processing of transactions	2 months ago	Jack Smith	Real-time version
v2.0 - Enhanced improve accuracy and performance	2 months ago	Jack Smith	-

Model details and versions

Fraud Detection Model

Our Advanced Fraud Detection Model represents the pinnacle of modern fraud detection technology, meticulously designed to safeguard businesses and financial institutions against the ever-evolving threat of fraudulent activities. Leveraging cutting-edge machine learning algorithm...

Actions

VersionsDetails

Description

Our Advanced Fraud Detection Model represents the pinnacle of modern fraud detection technology, meticulously designed to safeguard businesses and financial institutions against the ever-evolving threat of fraudulent activities. Leveraging cutting-edge machine learning algorithms, statistical analysis, and behavioral analytics, our model offers unparalleled accuracy and efficiency in identifying fraudulent transactions, activities, and patterns.

At its core, our model utilizes a sophisticated ensemble approach, combining the strengths of multiple machine learning techniques to achieve superior performance. By aggregating insights from various algorithms, including decision trees, random forests, logistic regression, and neural networks, our model can effectively detect and mitigate a wide r...[Show more](#)

Model ID

124dsdk-jlaskw-dl32oa-2kjfd-sjkwer2

Owner

Haley Wang

Last modified at

1 minute ago

Created at

1 minute ago

Labels

ClassificationTransformersPyTorchONNXdistilbertgenerated_from_trainereEnglishTransformers.jsEval Results24 papersPredictive AnalyticsReal-time DetectionEnsemble LearningFeature EngineeringEnterprise SolutionCloud DeploymentDocumentationAnomaly DetectionSupervised Learning3 more

Properties

Key	Value
team	finance
code_format	def print_prime(n): ... Print all primes between 1 and n Show more

Deploy and keep track

Applications

Enabled

Explore

Data Science Projects

Data Science Pipelines

Experiments

Experiments and runs

Artifacts

Executions

Model Registry

Model Serving

Resources

Settings

Cluster settings

Serving runtimes

Model Registry settings

User management

Registered models - Haley-private > my-model > v0.0.2

v0.0.2

improve accuracy and performance

DetailsRegistered deployments

Only partial of the model deployments showing in the list


This list only shows the model deployments that are deployed through Model Registry. To view the full list of deployments of this model and manage deployments, please navigate to the [Model Serving](#) section.

Model deployment name	Project	Serving runtime	Inference endpoint	Status
Invest-Portfolio-CD-4	finance-prod	ONNX Runtime	http://invest-portfolio-cd-3.example.com/pr...	✓
Invest-Portfolio-CD-3	finance-prod	ONNX Runtime	http://invest-portfolio-cd-3.example.com/pr...	✓
Invest-Portfolio-CD-2	finance-dev	ONNX Runtime	http://invest-portfolio-cd-2.example.com/pr...	✓
Invest-Portfolio-CD-1	finance-dev	OpenVINO Model Server (Supports GPUs)	http://invest-portfolio-cd-1.example.com/pr...	!


A vertical red bar on the left side of the slide contains a complex, stylized graphic. It features various icons: a cloud with a keyhole, a database cylinder, a server rack, a computer monitor, and several arrows pointing in different directions. There are also 'X' and 'O' symbols scattered throughout the design.

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](https://www.linkedin.com/company/red-hat)

 [facebook.com/redhatinc](https://www.facebook.com/redhatinc)

 [youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)

 twitter.com/RedHat