Making every engineer a machine learning expert

Openshift Commons, Barcelona, KubeCon 2019

Itay Gabbay & Guy Menachem, MOD, Israel
Tushar Katarki, Red Hat, USA

About Us

Itay Gabbay

Machine Learning Team Leader

Guy Menahem

Head Of Private Cloud R&D

Tushar Katarki

AI/ML Product Manager, OpenShift

IDF - Development acceleration using private cloud

WHAT IF I TOLD YOU

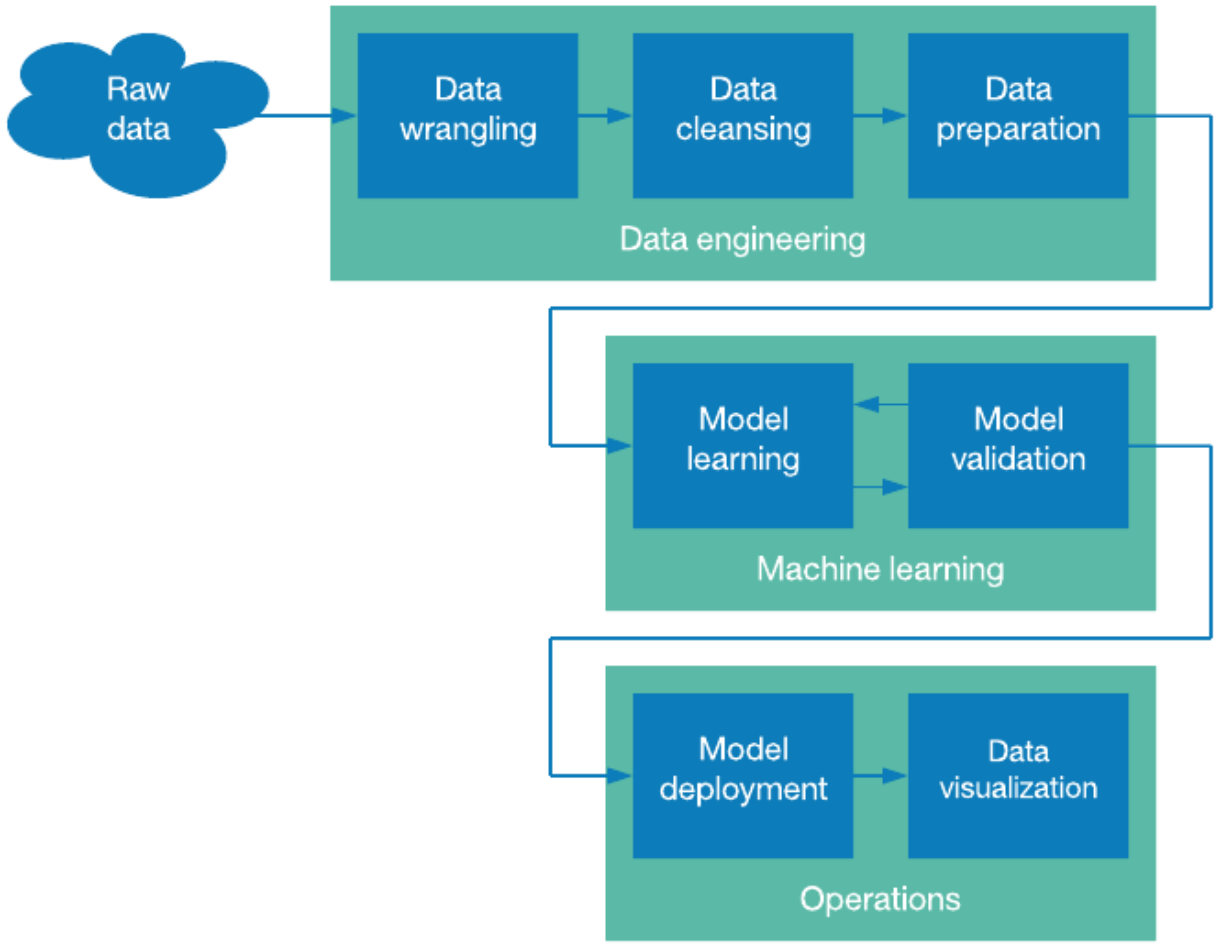
YOU CAN BE A ML EXPERT

Machine learning for dummies

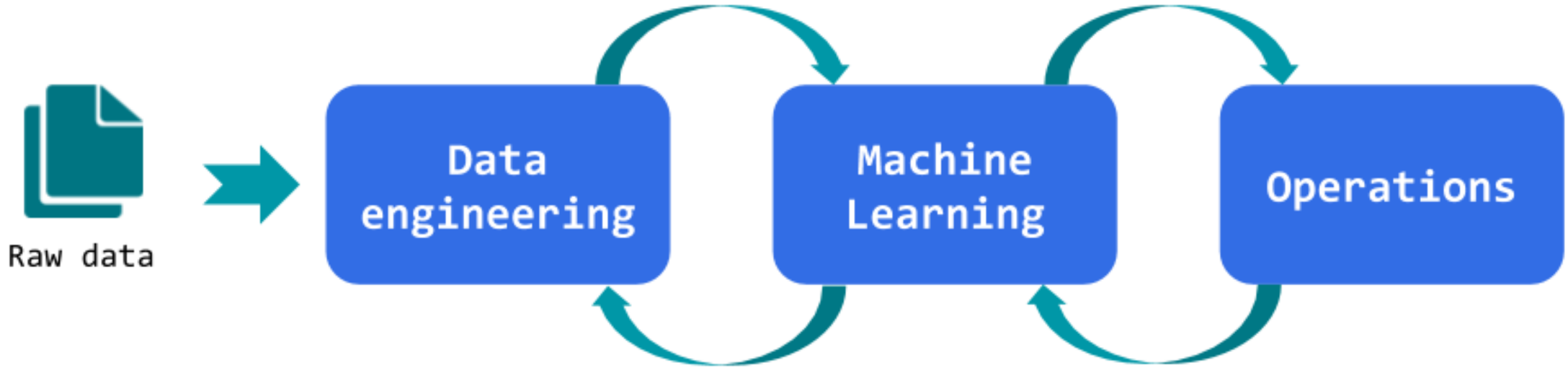
Name	Age	Date	Running nose?	Fever	Got flu?
John Smith	37	27-02-2018	Yes	37.8	YES
Alice Jones	25	14-04-2017	No	39	NO
...	...	...	...	...	...
James Gonzales	42	27-08-2017	No	36.0	???


Machine Learning Algorithms
Deep Learning
Boosting
Ensemble
Elastic Net
Regularization
Cubist
Rule system
Regression
Linear Regression
clustering
k-Means
k-Medians
intance Based
k-Nearest Neighbour
Dimensionality Reduction
Sammon Mapping
Projetion Pursuit
Decision Tree
C5.0
C4.5
M5
Bayesian
Naive Bayes
Bayesian Network

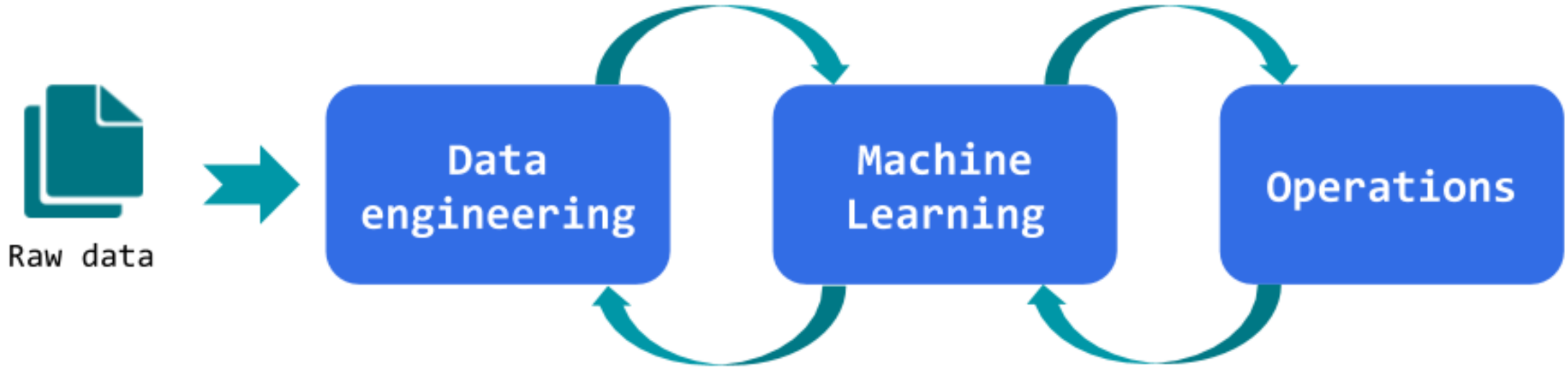
Common Data Science Pipeline



Common Data Science Pipeline



Common Data Science Pipeline



Top 4 Challenges of Data Science

Environment

Usually a machine learning model needs unique resources like GPUs or huge memory

History

What is the best model we've got?

Optimization

Choosing the best algorithm and model

Deployment

Deploy it to Production

Our Solution

Research Environment

Using JupyterHub to dynamically allocate resources and enable collaboration between data scientists

Our Solution

Experiments Reproduction

Keeping track of machine learning experiments using a uniquely developed python package

Our Solution

Hyper-Parameters Optimization

Choosing the best model, using complex optimization algorithms and distributed evaluations

Our Solution

Model Deployment

Deploy the best model to production with a one line of code, using custom containers compatible with the most popular machine learning frameworks

Live Demo

Cluster Architecture

Openshift

Openshift Stack

GPU Compute nodes

Self written GPU scheduler

V100 V100

Multi GPU node

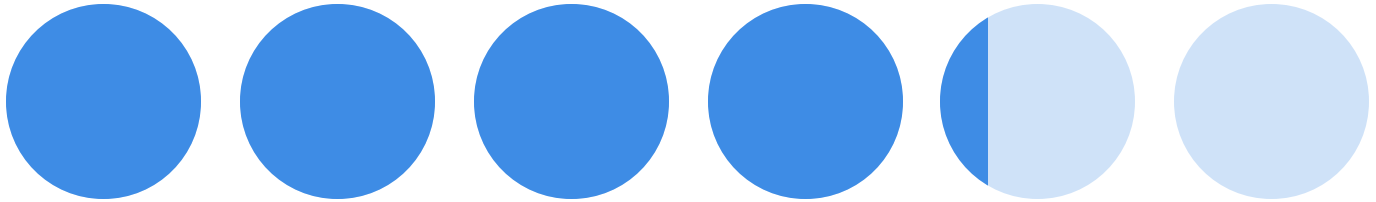
Nvidia DGX-1

Auto ML

30%

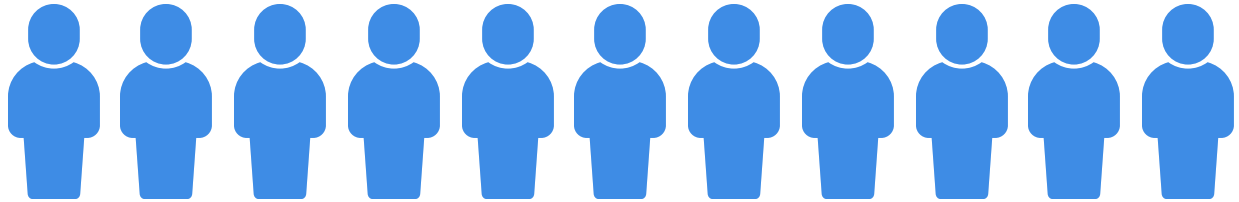
Average improvement in machine learning model performance

70%



Increase in machine learning models number

600%

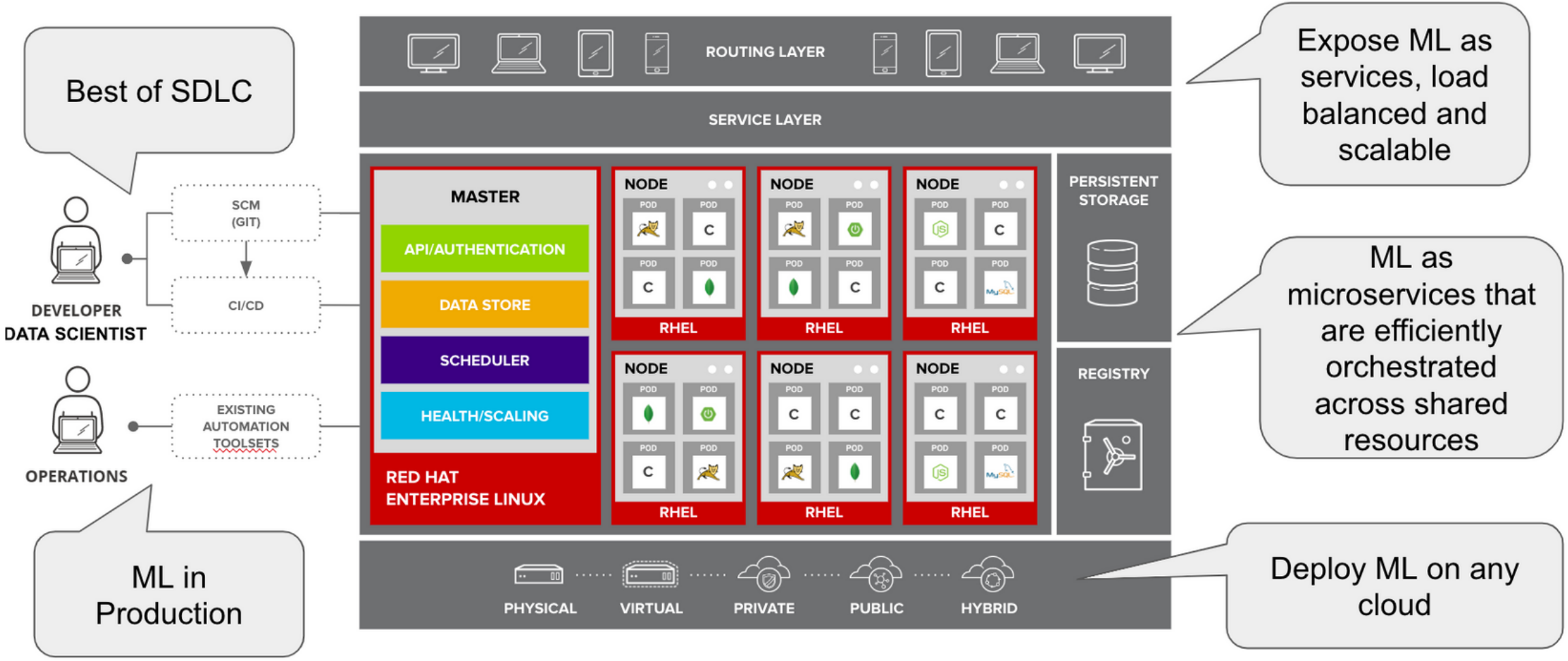


Growth of users in the previous 6 months

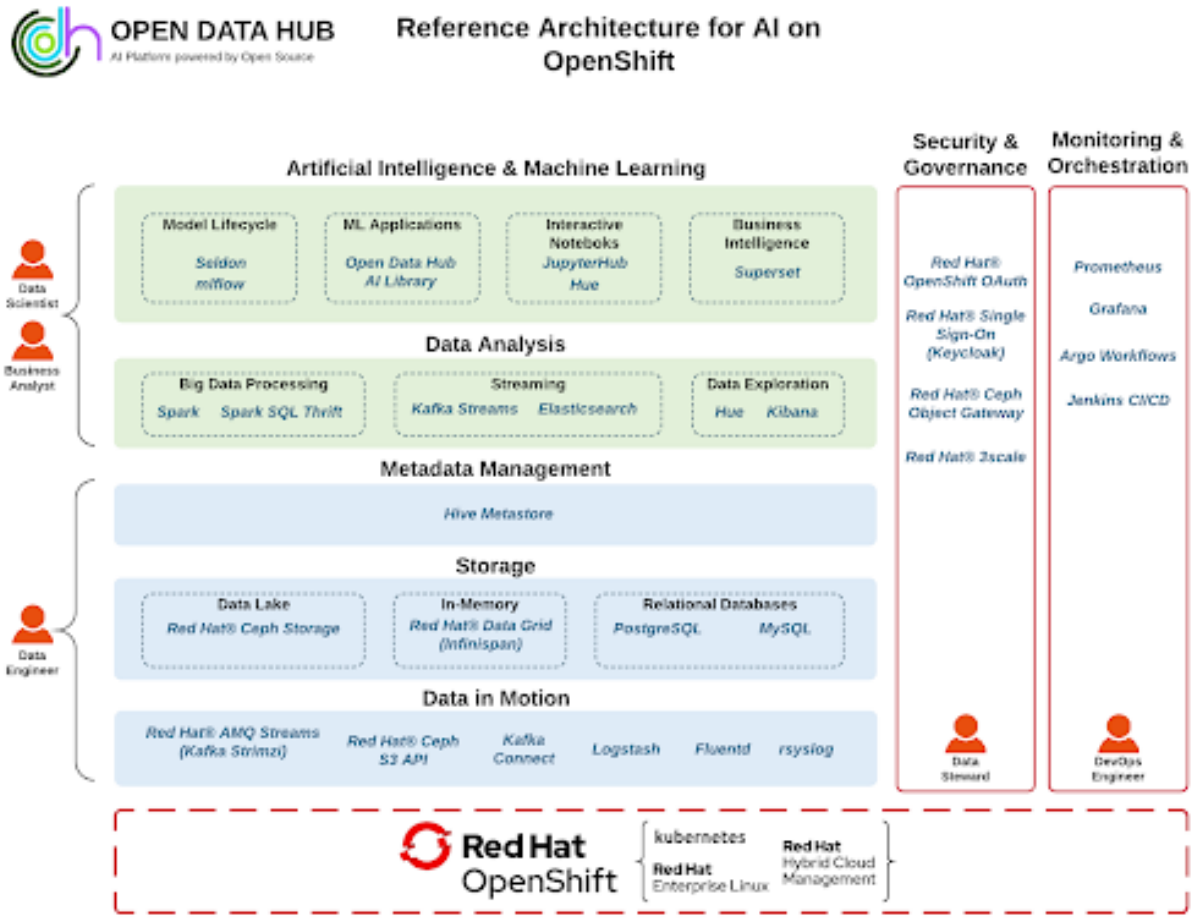
Current Challenges

- Code Remote Debugging
- Better way to save experiments history
- AutoML for unstructured data (images, audio)
- GPU split to multiple containers, no time scheduling
- Openshift multi cluster

OpenShift for AI/ML



Open Data Hub on OpenShift



Open source Reference Architecture for data and machine learning on OpenShift to create Intelligent Applications for hybrid cloud

Accelerated AI - Red Hat and NVIDIA

Easy Button for Accelerating AI in the Enterprise Datacenter

- OpenShift 4 on NVIDIA NGC-Ready Bare Metal servers with NVIDIA (datacenter) GPUs
- Full Automation with the new OpenShift 4 installer and GPU operator
- NVIDIA NGC Containers (AI/ML container images that are CUDA optimized) on OpenShift

Resources

- Offering: https://blogs.nvidia.com/blog/2019/05/09/red-hat-openshift-gpu-kubernetes/
- Overview: https://devblogs.nvidia.com/gpu-support-ai-workloads-openshift4/
- Early Access Customer Sign Up: https://www.openshift.com/accelerated-ai

Contact Us

Itay Gabbay & Guy Menachem

@ MLonIDF@gmail.com

Tushar Katarki

@ tkatarki@redhat.com