Microsoft

Microsoft Cloud Workshop

MLOps with Azure Machine Learning

Jon Satchwell
Tiago Henriques

Welcome to this Virtual workshop!

Agenda

Time	Description
09:00 – 11:30	**Presentation and Demo** • Introduction to Azure Machine Learning • Brief overview, key features and MLOps • Architecture E2E for Machine Learning and MLOps **Hands-on Lab – setup guide overview**
11:30 – 13:30	Personal Time / Lunch Break
13:30 – 16:30	**Hands-on Lab (self-paced**) *Possibly ad-hoc meeting to discuss common problems during lab execution* *Questions can be asked in Teams chat* *Individual calls with the proctors if needed during this time*
16:30 – 17:00	Wrap-up and Q&A

Intro to Azure ML

Machine Learning on Azure

Domain specific pretrained models
To simplify solution development

Vision | Speech | Language | Web search | Decision

Familiar Data Science tools
To simplify model development

Visual Studio Code | Azure Notebooks | Jupyter | Command line

Popular frameworks
To build advanced deep learning solutions

PyTorch | TensorFlow | Scikit-Learn | ONNX

Productive services
To empower data science and development teams

Azure Machine Learning | Azure Databricks | Machine Learning VMs

Powerful infrastructure
To accelerate deep learning

CPU | GPU | FPGA

MLOps From the Intelligent Cloud to the Intelligent Edge

AzureML

For all skill levels

Industry leading MLOps

Open & Interoperable

Trusted

What is Azure Machine Learning?

Set of Azure Cloud Services

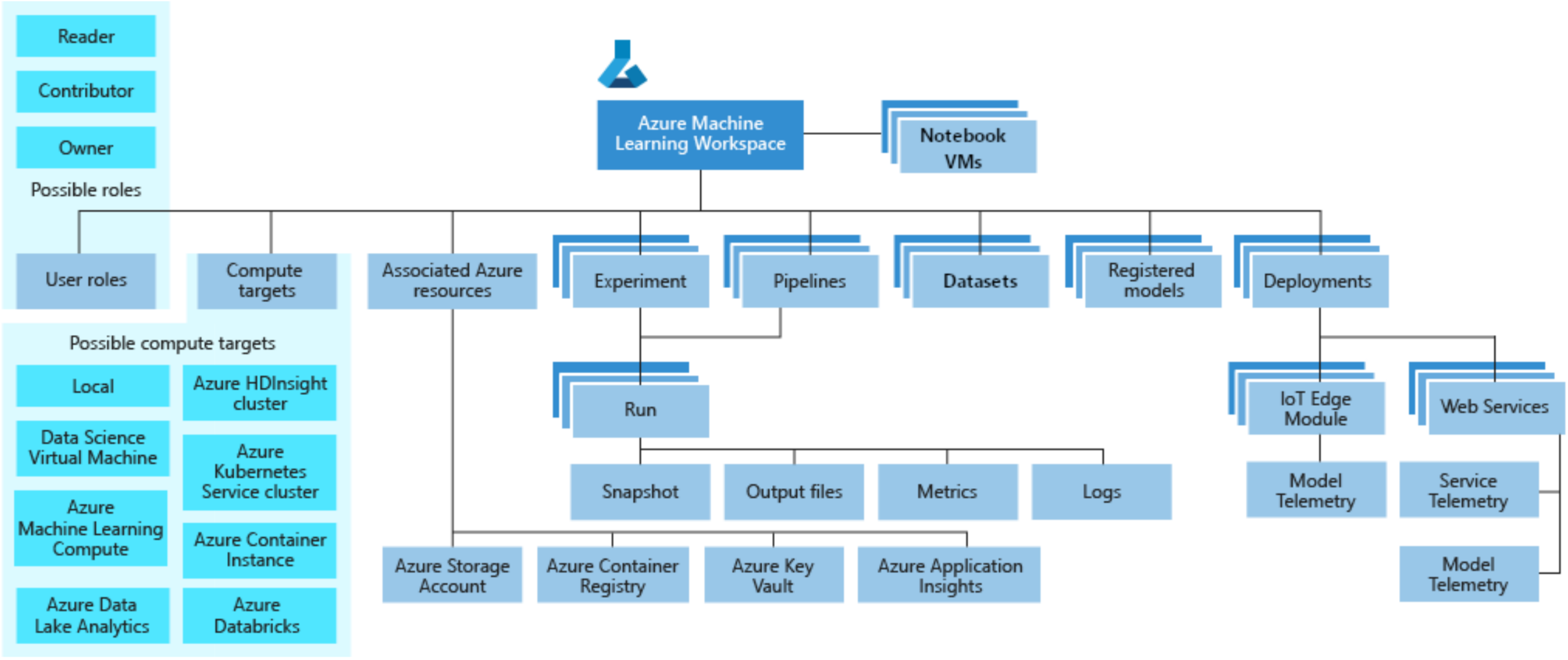
+

Python SDK/CLI/UX

That enables you to:

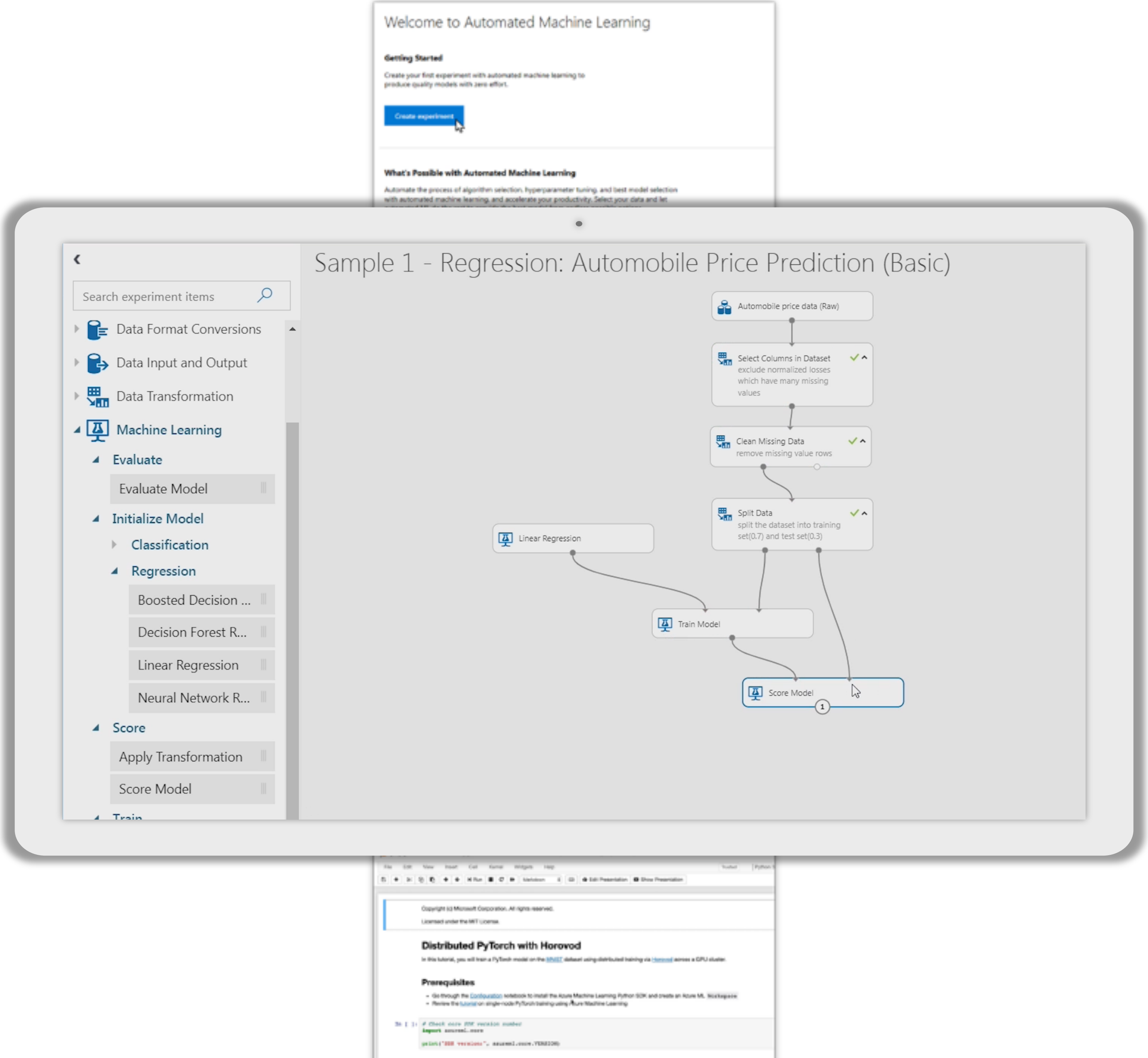
- ✓ Prepare Data
- ✓ Build Models
- ✓ Train Models
- ✓ Manage Models
- ✓ Track Experiments
- ✓ Deploy Models

Key Artifacts of Azure Machine Learning



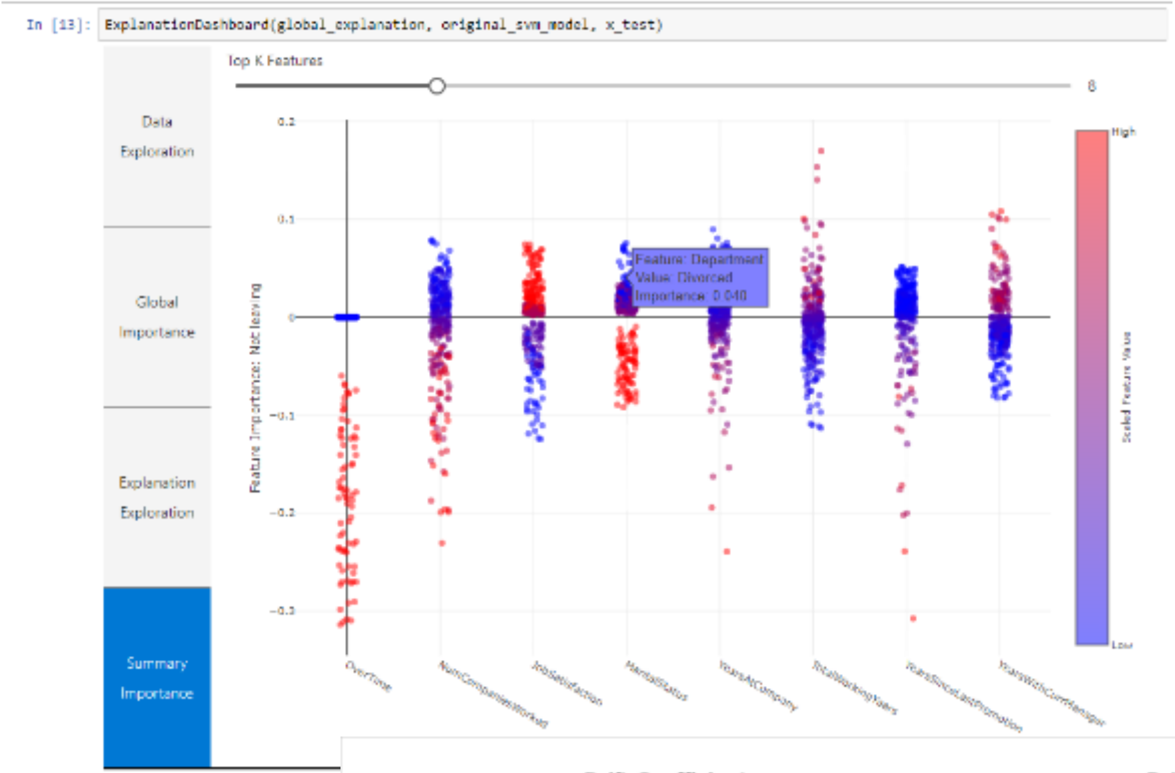
Azure Machine Learning

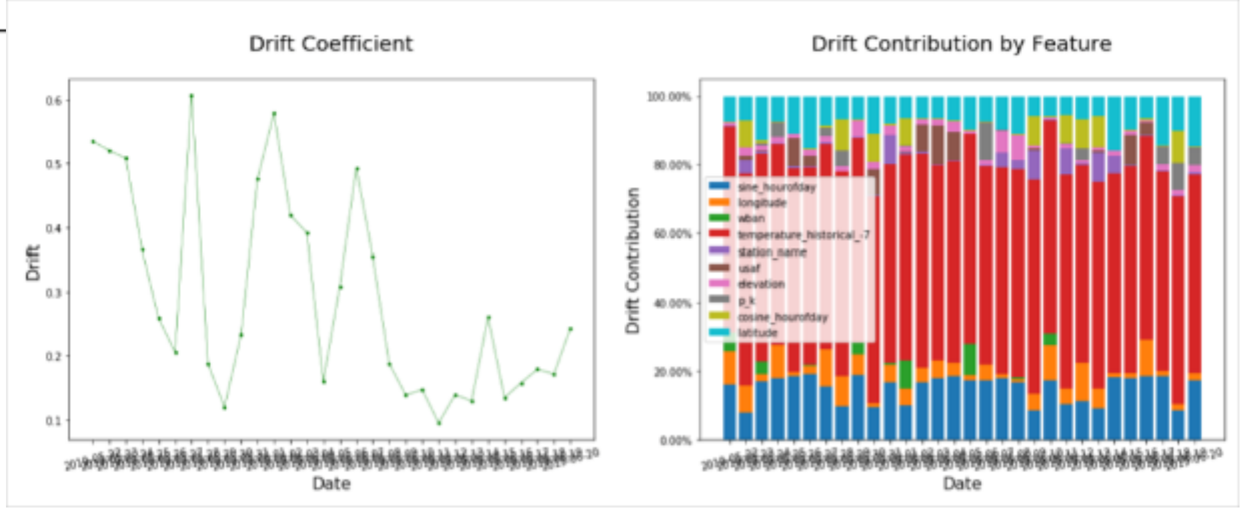
For all skill levels
Automated ML + drag & drop + code
Industry leading MLOps
Open & Interoperable
Trusted



Azure Machine Learning

For all skill levels
Industry leading MLOps
Open & Interoperable
Trusted
Model explainability and fairness





Azure ML Demos

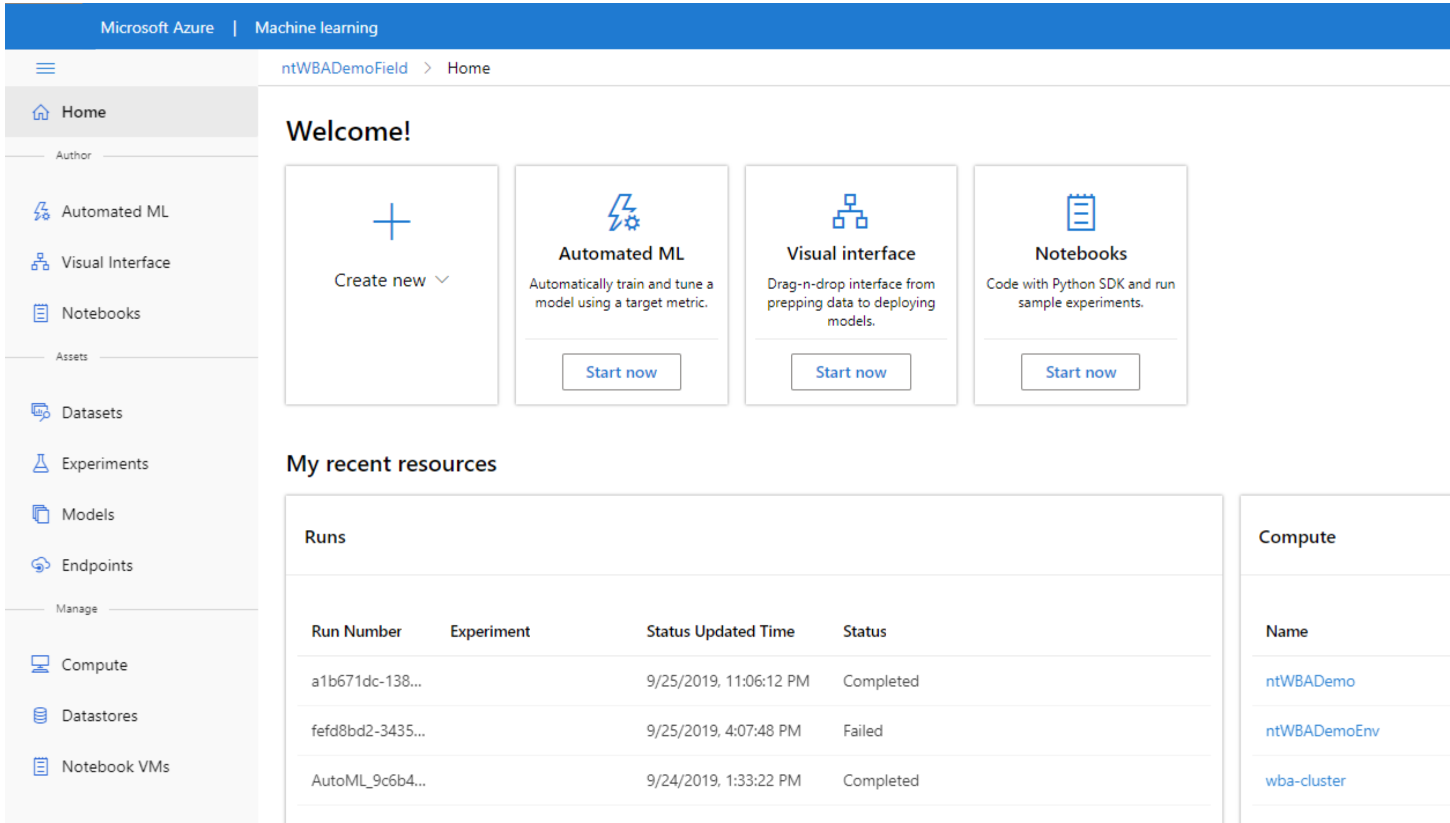
Overview and key features

- Compute Instance
- Experiment tracking
- Metrics Logging
- Datasets
- Model Register

- Model Deployment

- Designer
- AutoML

Azure ML



Microsoft Azure | Machine learning
ntWBADemoField > Home
Home
Author
Automated ML
Visual Interface
Notebooks
Assets
Datasets
Experiments
Models
Endpoints
Manage
Compute
Datastores
Notebook VMs
Welcome!
Create new
Automated ML
Automatically train and tune a model using a target metric.
Start now
Visual interface
Drag-n-drop interface from prepping data to deploying models.
Start now
Notebooks
Code with Python SDK and run sample experiments.
Start now
My recent resources
Runs
Run Number
Experiment
Status Updated Time
Status
a1b671dc-138...
9/25/2019, 11:06:12 PM
Completed
fefd8bd2-3435...
9/25/2019, 4:07:48 PM
Failed
AutoML_9c6b4...
9/24/2019, 1:33:22 PM
Completed
Compute
Name
ntWBADemo
ntWBADemoEnv
wba-cluster

The Machine Learning Workflow

Prepare

Clean, transform and merge data

Experiment

Train &
Test Model

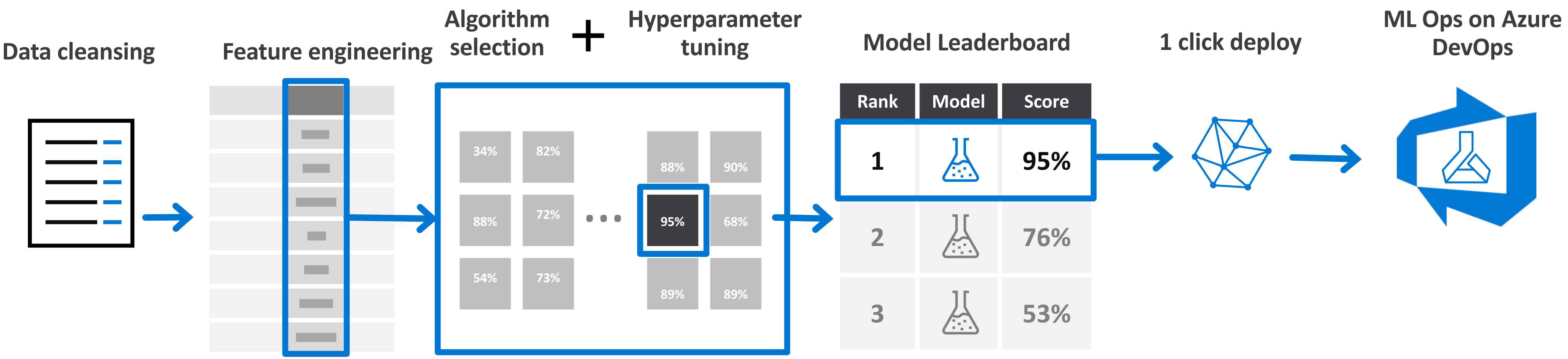
Deploy

Deploy to cloud or edge

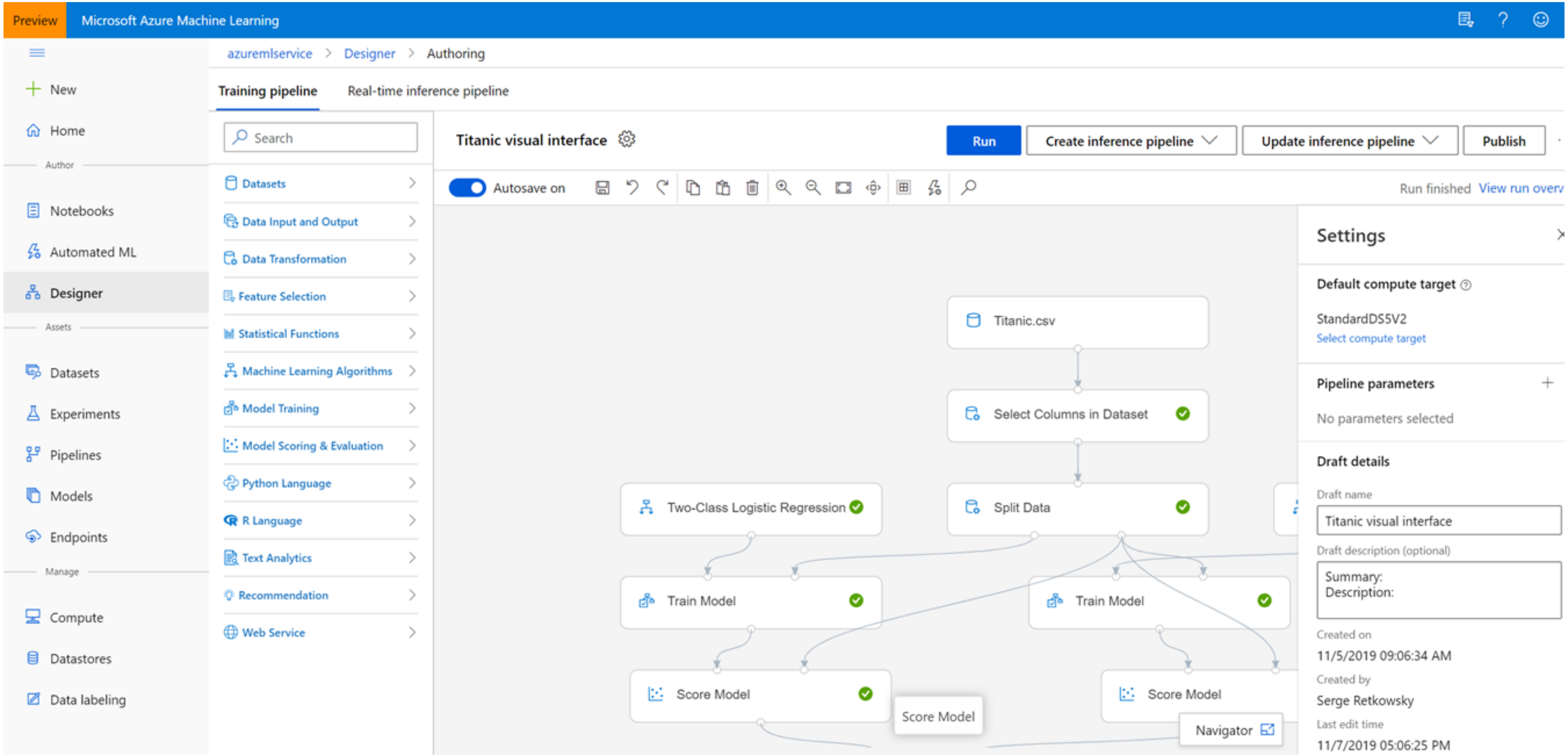
Operationalize

Monitor, retrain, CI/CD

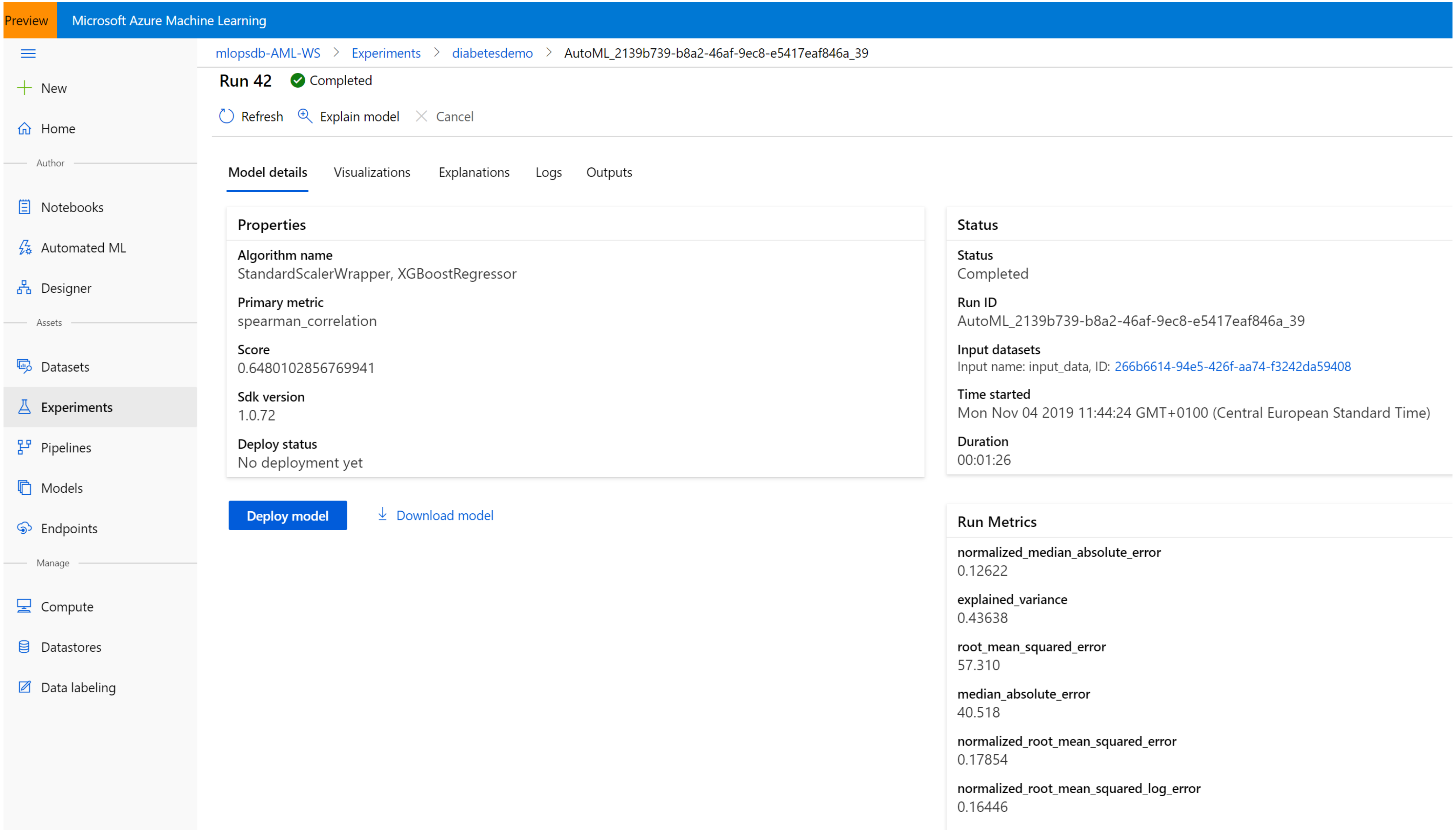
Azure Machine Learning Service Automation



Azure ML – Designer



Azure ML – AutoML



Distributed Training

You submit a model training 'job' – the infrastructure is managed for you.

- Jobs run on a native VM or Docker container.
- Supports Low priority (Cheaper) or Dedicated (Reliable) VMS.

Auto-scales: Just specify min and max number of nodes.

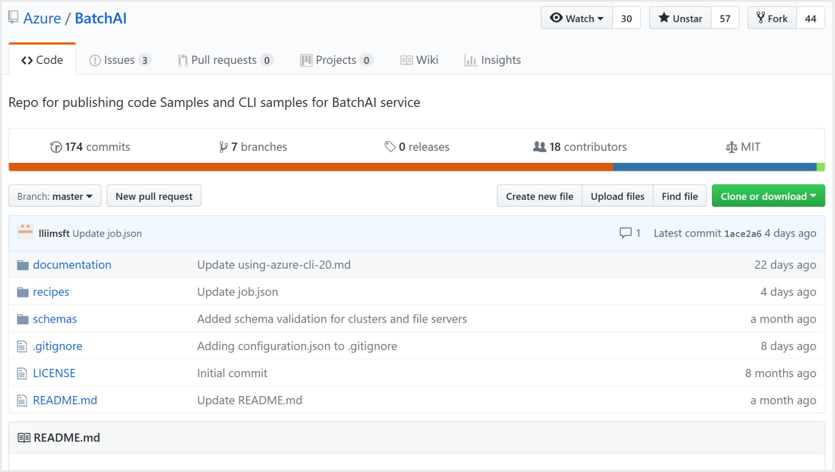
- If min is set to zero, *cluster is deleted when no jobs are running; so pay only for job duration*.

Works with most popular frameworks and multiple languages.

Supports distributed training with Horovod.

Cluster can be shared; multiple experiments can be run in parallel.

Supports most VM Families, including latest NVidia GPUs for DL model training.

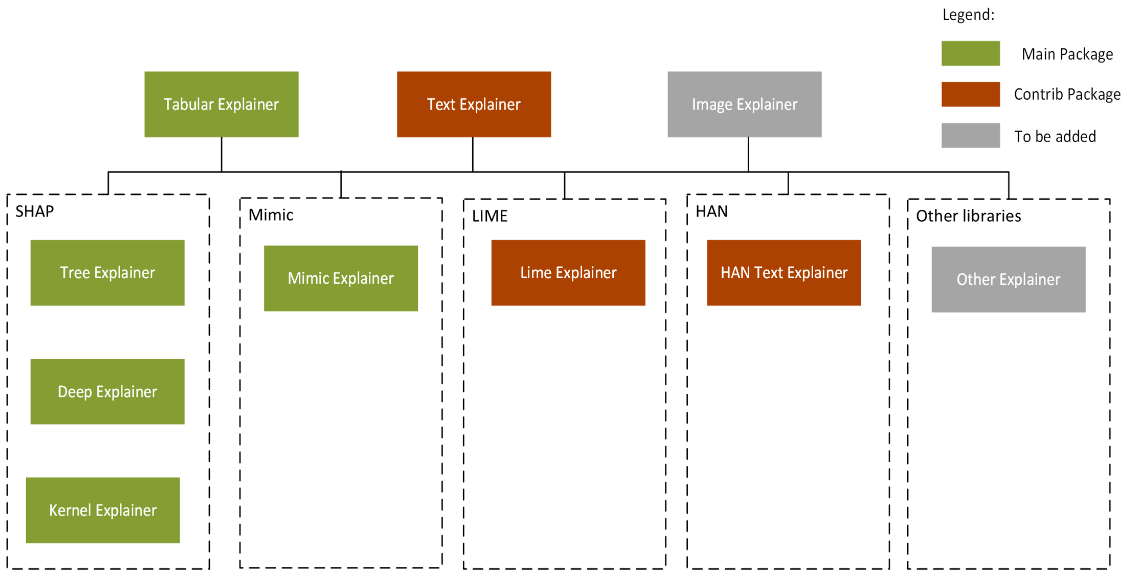


Azure ML – Model Interpretability

- Explain ML models **globally on all data, or locally on a specific data point** using the state-of-art technologies in an easy-to-use and scalable fashion.

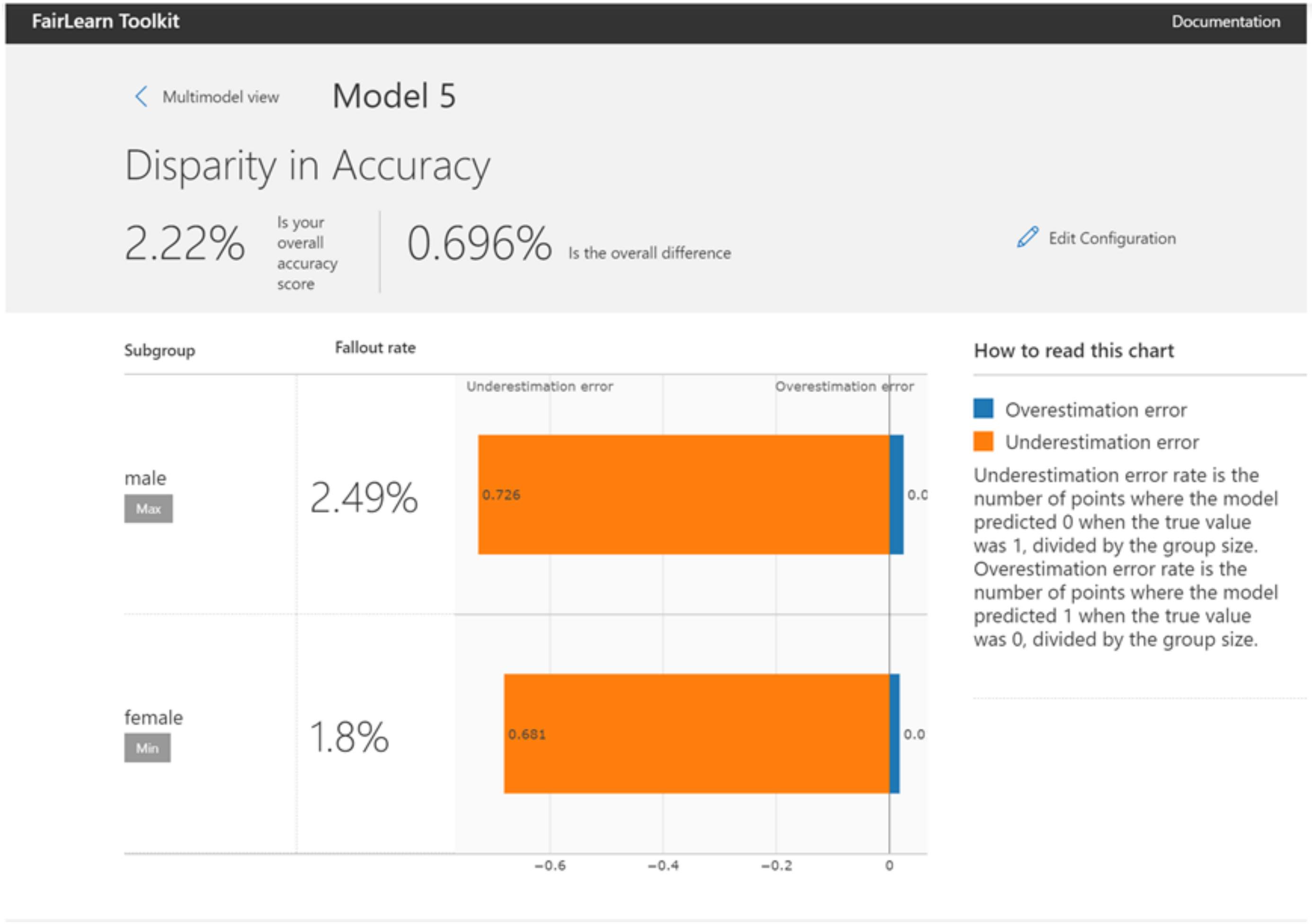
- Incorporate cutting-edge interpretability solutions **developed by Microsoft** and leverages **proven third-party libraries**

- Create a **common API and data structure** across the integrated libraries and integrates AML services



Azure ML – Fairness

- **Fairlearn is a new open source unfairness assessment** and mitigation tool that assists business stakeholders, executives, developers, and data scientists to get insights about the unfairness in their model predictions.



Azure ML – Data Drift

What is Data Drift

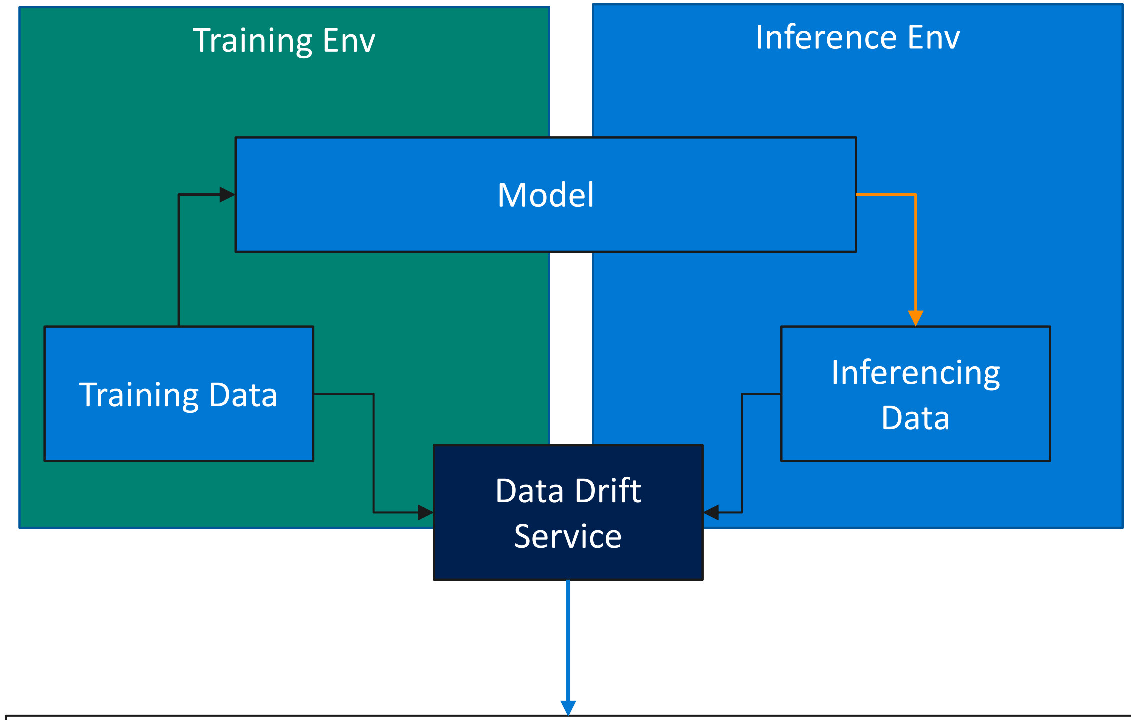
- When the inference data varies enough from the training data to cause model degradation, we consider it data drift

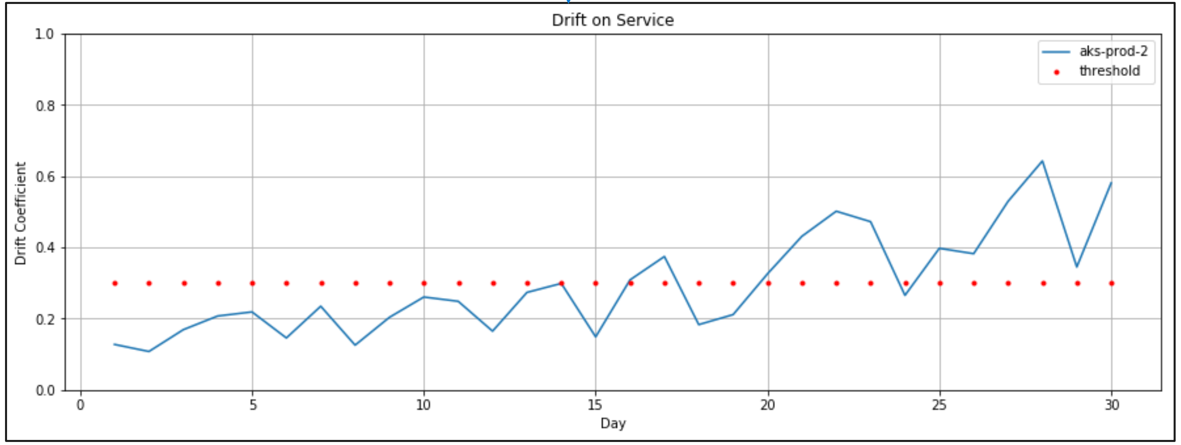
Why is Data Drift important

- in absence of labels, data drift can be used as a proxy for model performance metrics

What can I expect from data drift service

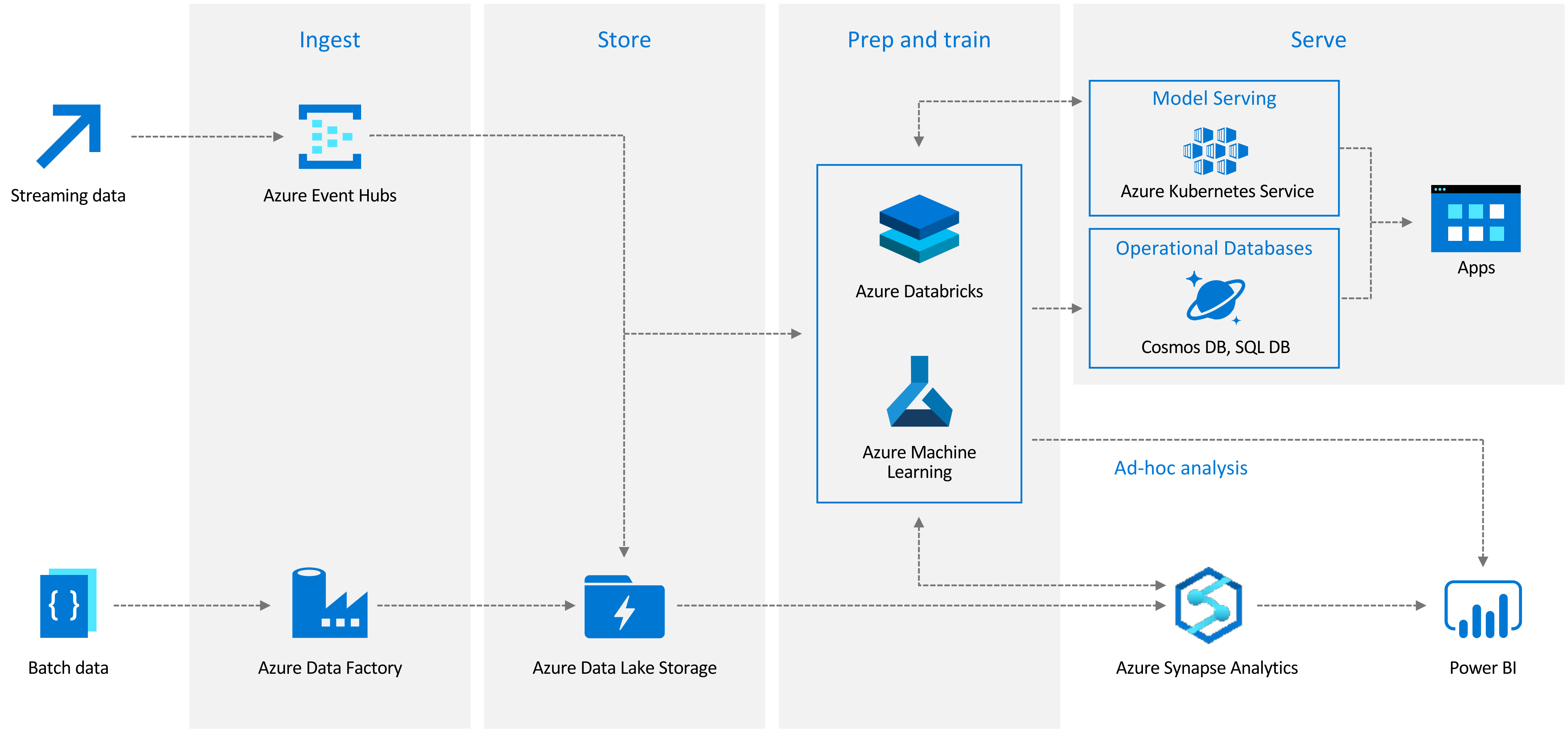
- **Alerting and Insights** in Previews
- **Drift Coefficient** - measures magnitude of drift
- **Drift Contribution by Feature** - measures feature that caused data drift
- **Distance Metrics by Feature** - distance metrics on features
- **Distribution Visualization** - capture and visualize change in distributions over time





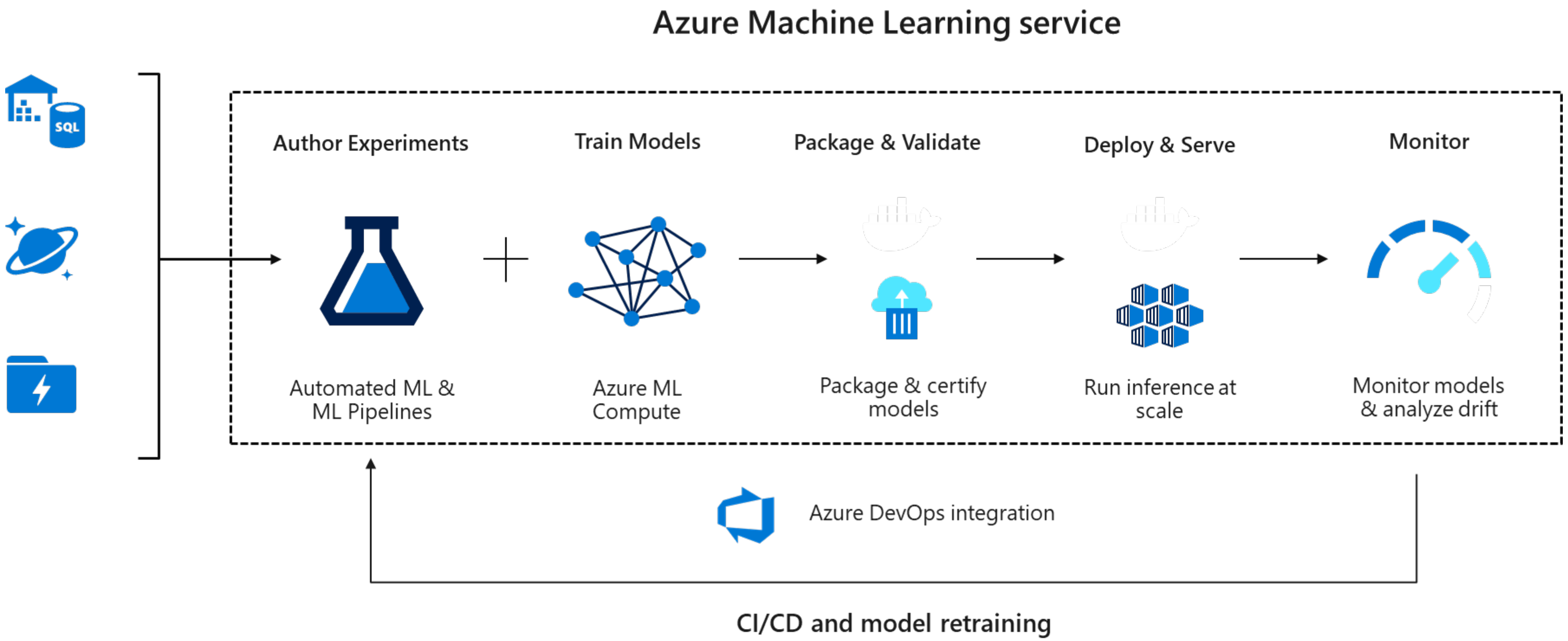
Reference Architecture

Machine learning



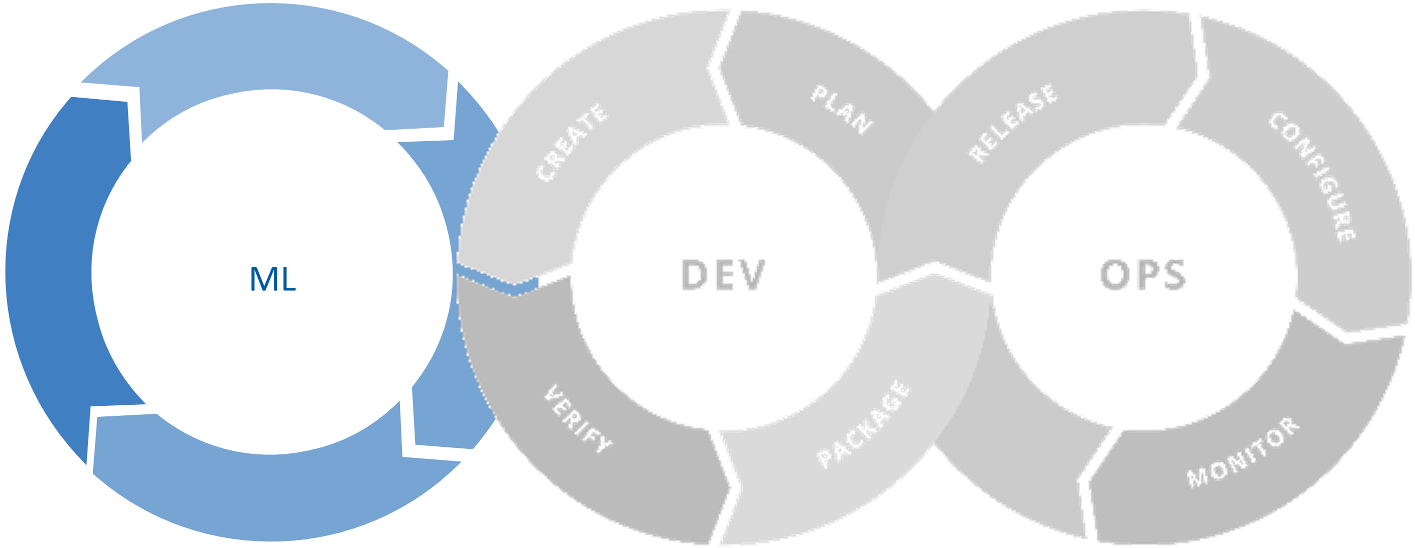
MLOps

MLOps


Azure Machine Learning service
Author Experiments
Train Models
Package & Validate
Deploy & Serve
Monitor
SQL
Automated ML &
ML Pipelines
Azure ML
Compute
Package & certify
models
Run inference at
scale
Monitor models
& analyze drift
Azure DevOps integration
CI/CD and model retraining

MLOps = ML + DEV + OPS

Help bring models to production



Experiment

Data Acquisition
Business Understanding
Initial Modeling

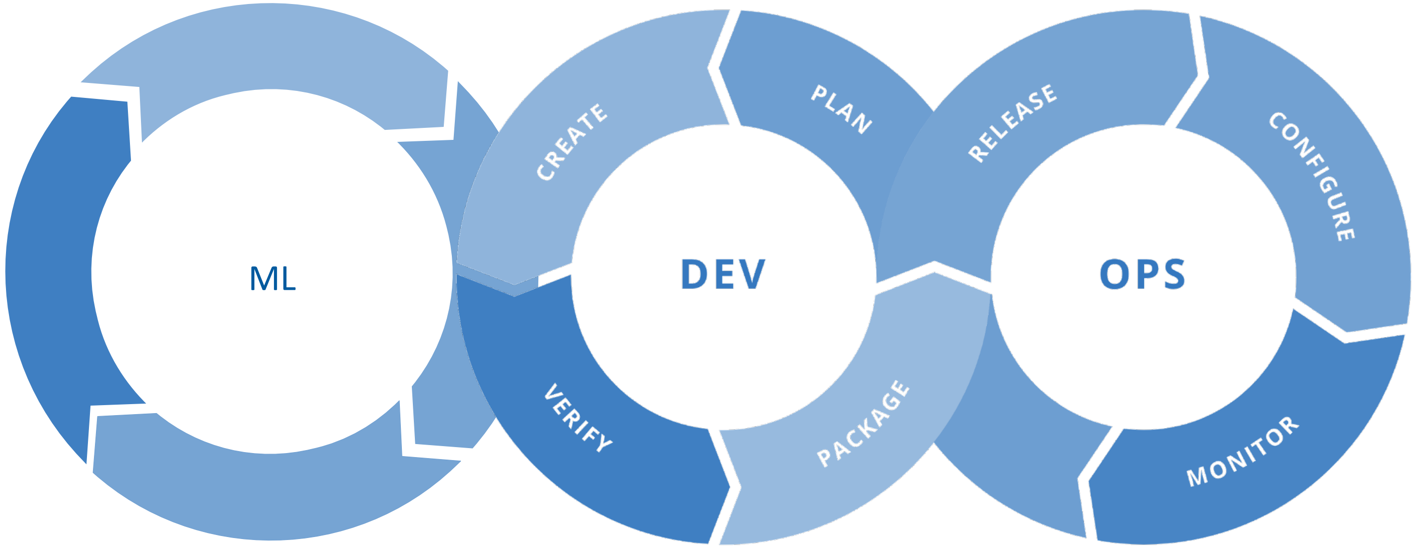
Develop

Modeling + Testing
Continuous Integration
Continuous Deployment

Operate

Continuous Delivery
Data Feedback Loop
System + Model Monitoring

MLOps = ML + DEV + OPS



Experiment
Data Acquisition
Business Understanding
Initial Modeling

Develop
Modeling + Testing
Continuous Integration
Continuous Deployment

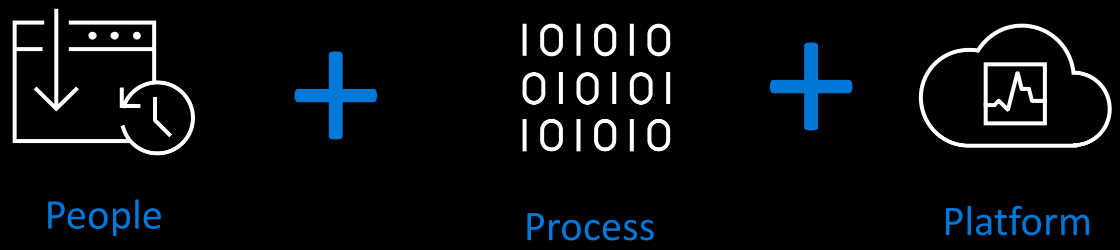
Operate
Continuous Delivery
Data Feedback Loop
System + Model Monitoring

MLOps – Focus Areas

Challenge	How to solve – Best Practices
Reproducibility & versioning	Track, snapshot & manage assets used to create the model Enable collaboration and sharing of ML pipelines
Auditability & explainability	Maintain asset integrity & persist access control logs Certify model behavior meets regulatory & adversarial standards
Packaging & validation	Support model portability across a variety of platforms Certify model performance meets functional requirements
Deployment & monitoring	Release models with confidence Monitor & know when to retrain

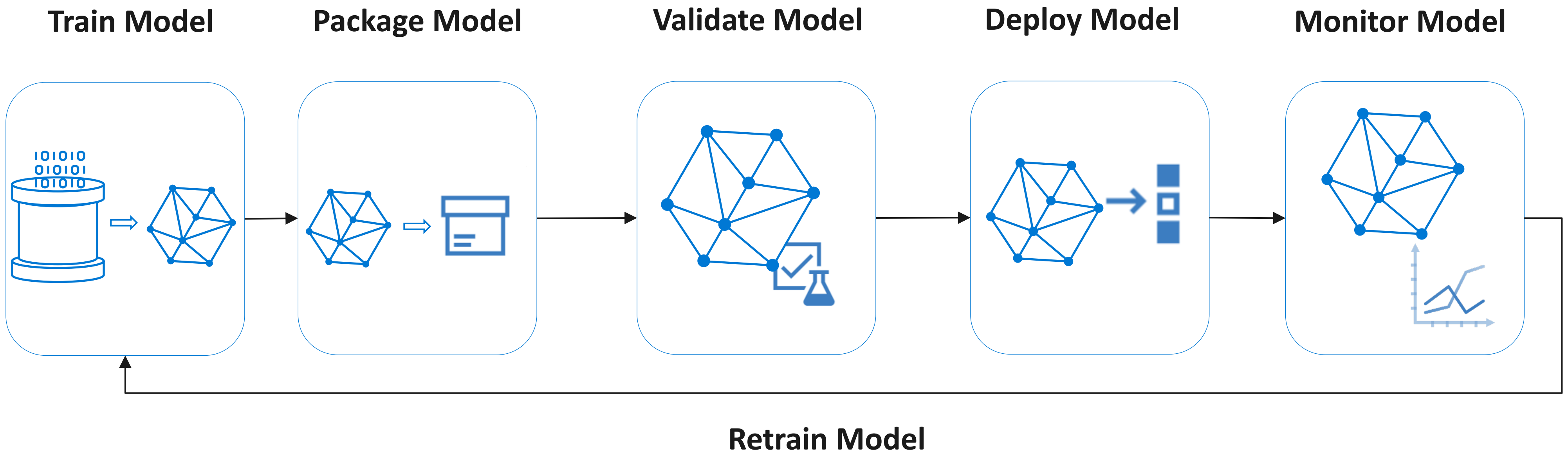
MLOps == How to bring ML to production

Bring together **people**, **process**, and **platform** to automate ML-infused software delivery & provide continuous value to our users.

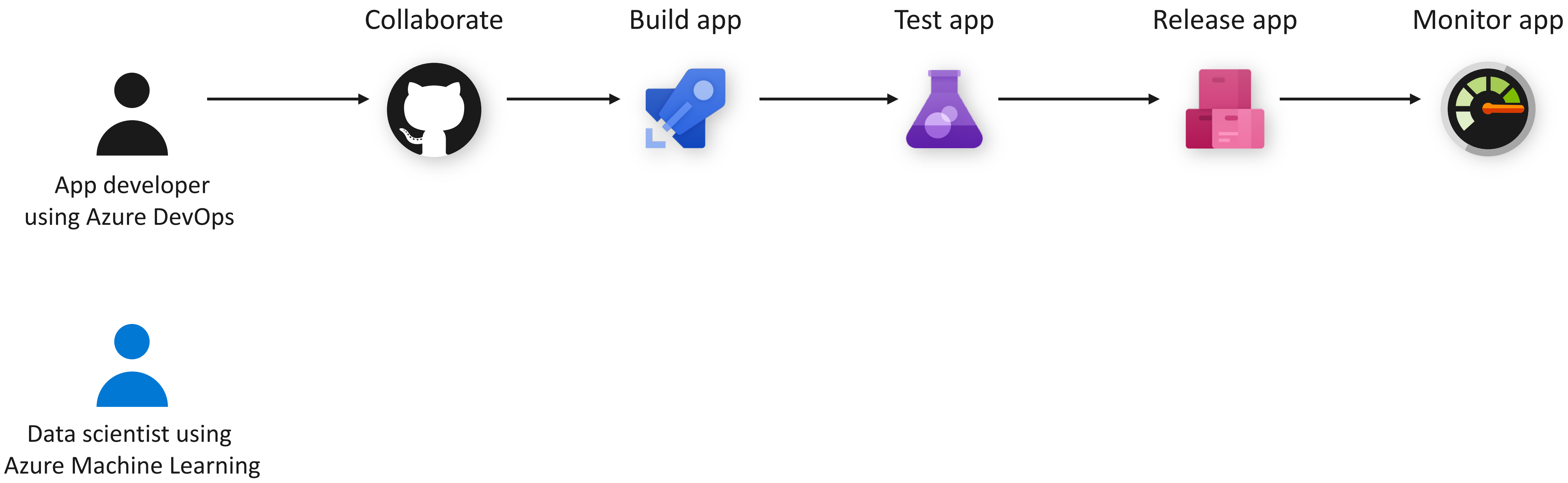


What is the E2E ML lifecycle?

- **Develop & train model** with reusable ML pipelines
- **Package model** using containers to capture runtime dependencies for inference
- **Validate model behavior—**functionally, in terms of responsiveness, in terms of regulatory compliance
- **Deploy model—**to cloud & edge, for use in real-time/streaming/batch processing
- **Monitor model** behavior & business value, know **when to replace/deprecate a stale model**

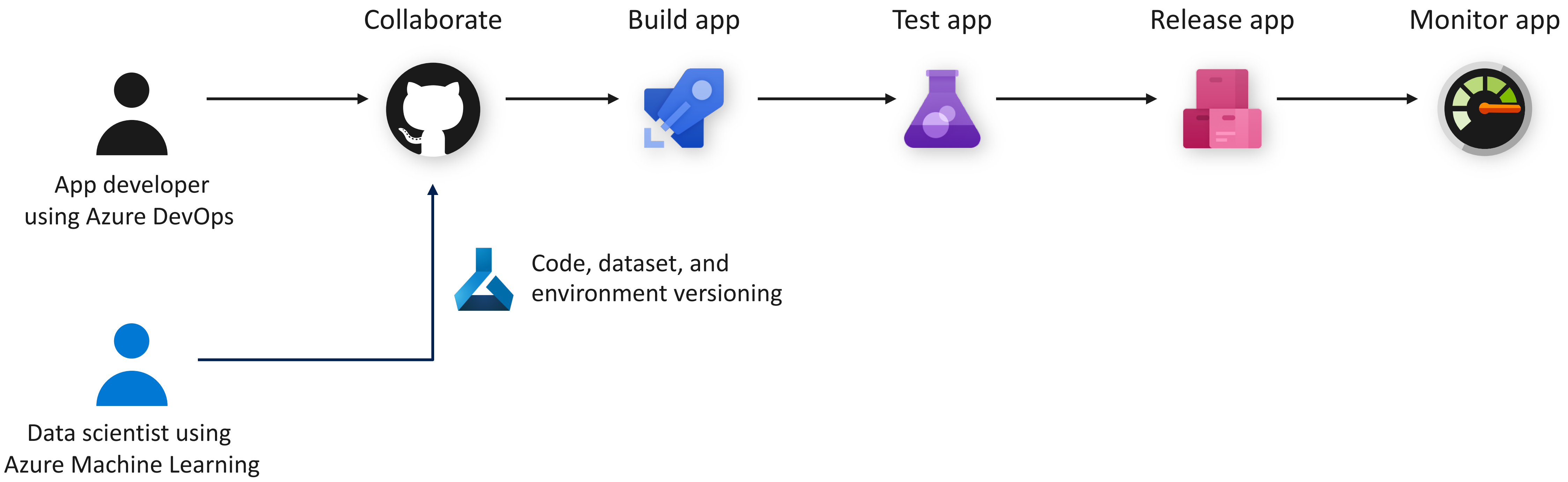


MLOps Workflow



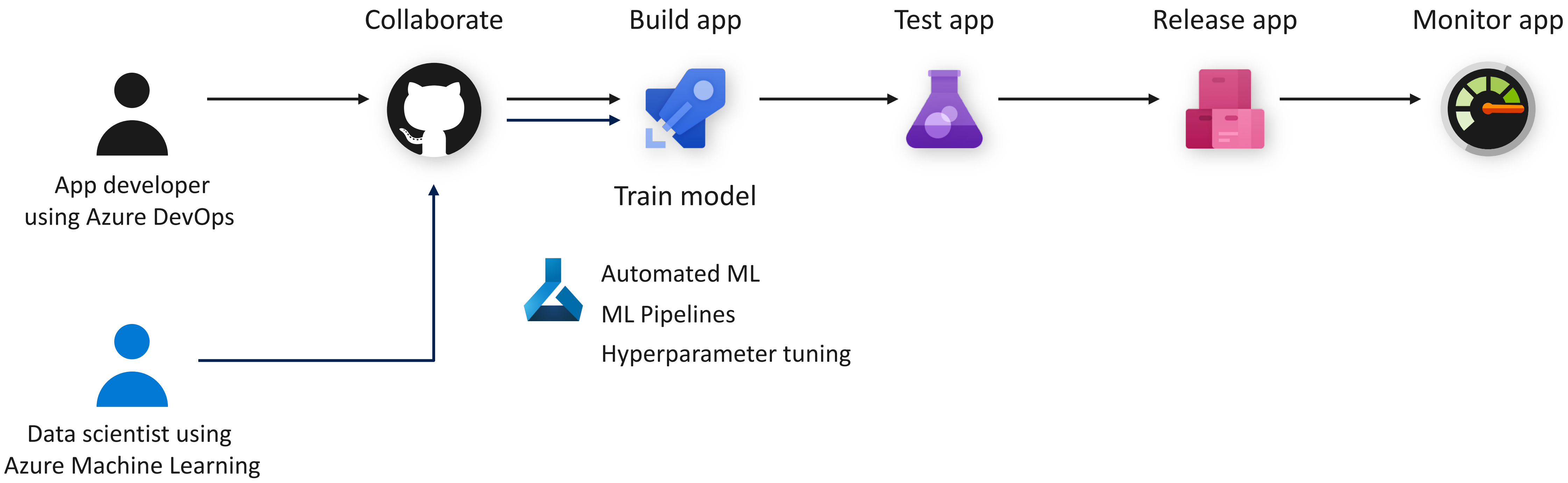
- ✕ Model reproducibility
- ✕ Model validation
- ✕ Model deployment
- ✕ Model retraining

MLOps Workflow



- ✓ Model reproducibility
- ✗ Model validation
- ✗ Model deployment
- ✗ Model retraining

MLOps Workflow



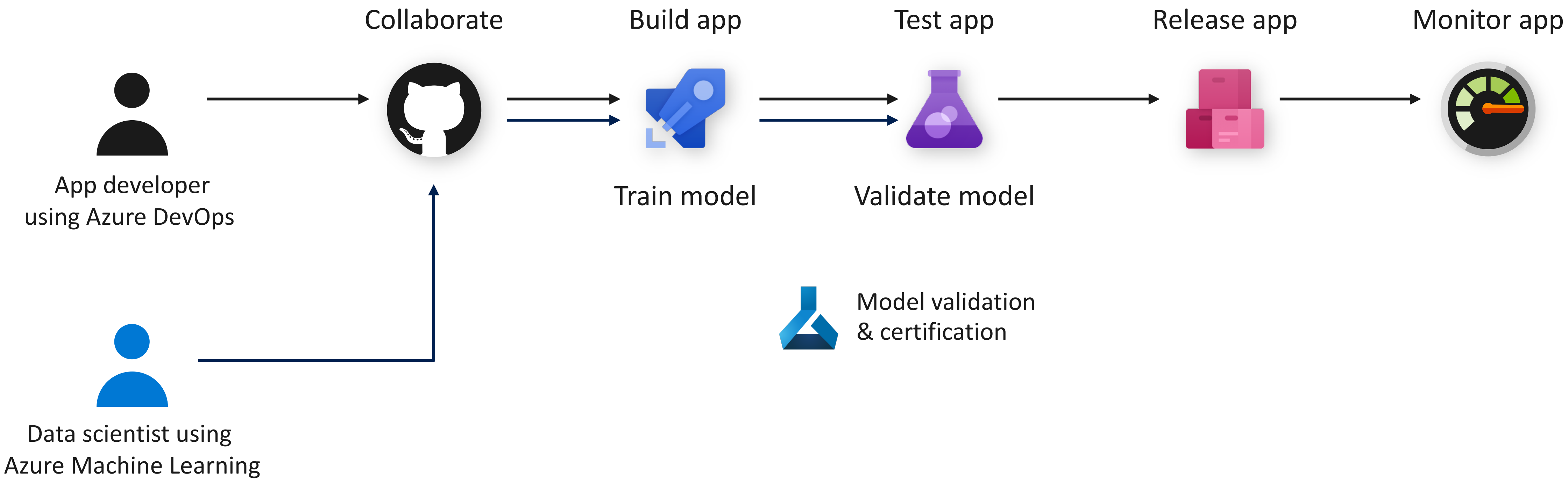
✓ Model reproducibility

✗ Model validation

✗ Model deployment

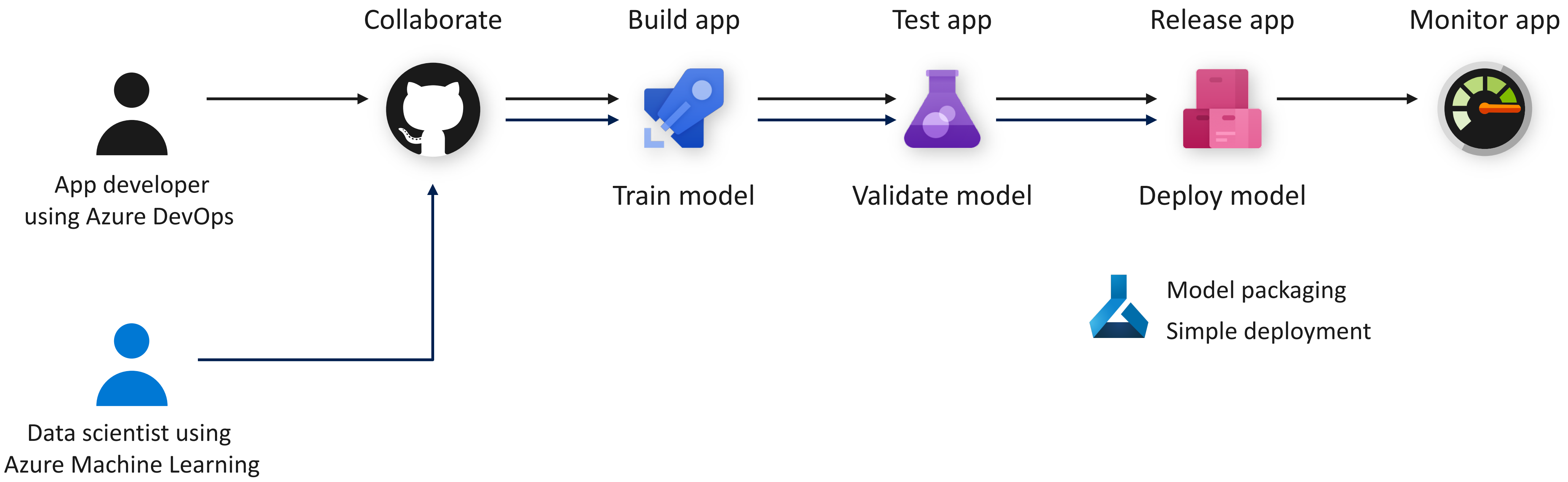
✗ Model retraining

MLOps Workflow



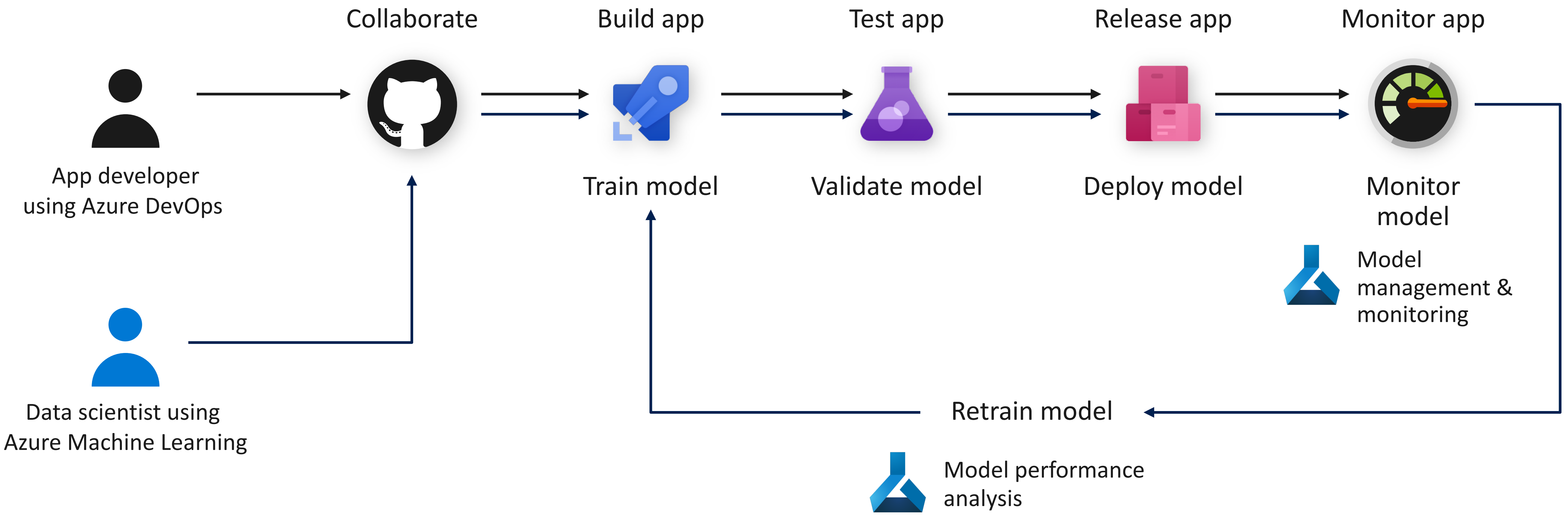
- ✓ Model reproducibility
- ✓ Model validation
- ✗ Model deployment
- ✗ Model retraining

MLOps Workflow



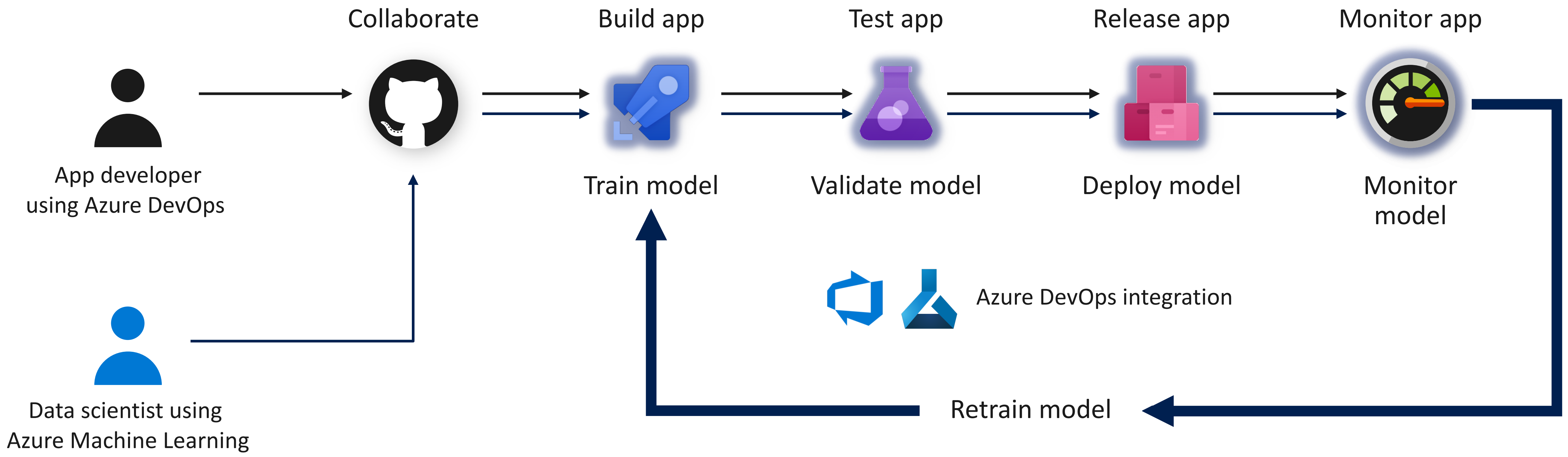
- ✓ Model reproducibility
- ✓ Model validation
- ✓ Model deployment
- ✗ Model retraining

MLOps with Azure Machine Learning



- ✓ Model reproducibility
- ✓ Model validation
- ✓ Model deployment
- ✓ Model retraining

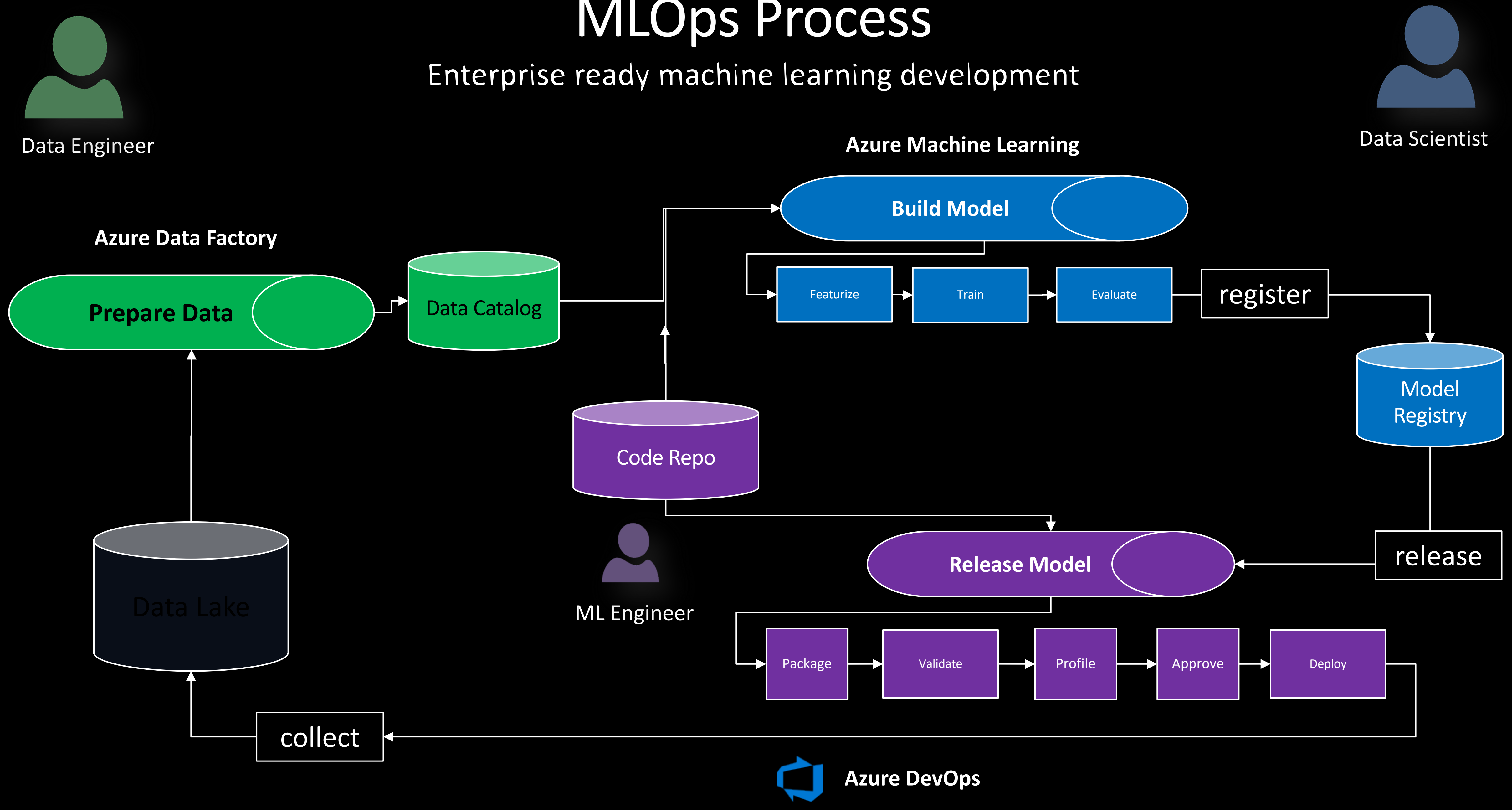
MLOps with Azure Machine Learning



- ✓ Model reproducibility
- ✓ Model validation
- ✓ Model deployment
- ✓ Model retraining

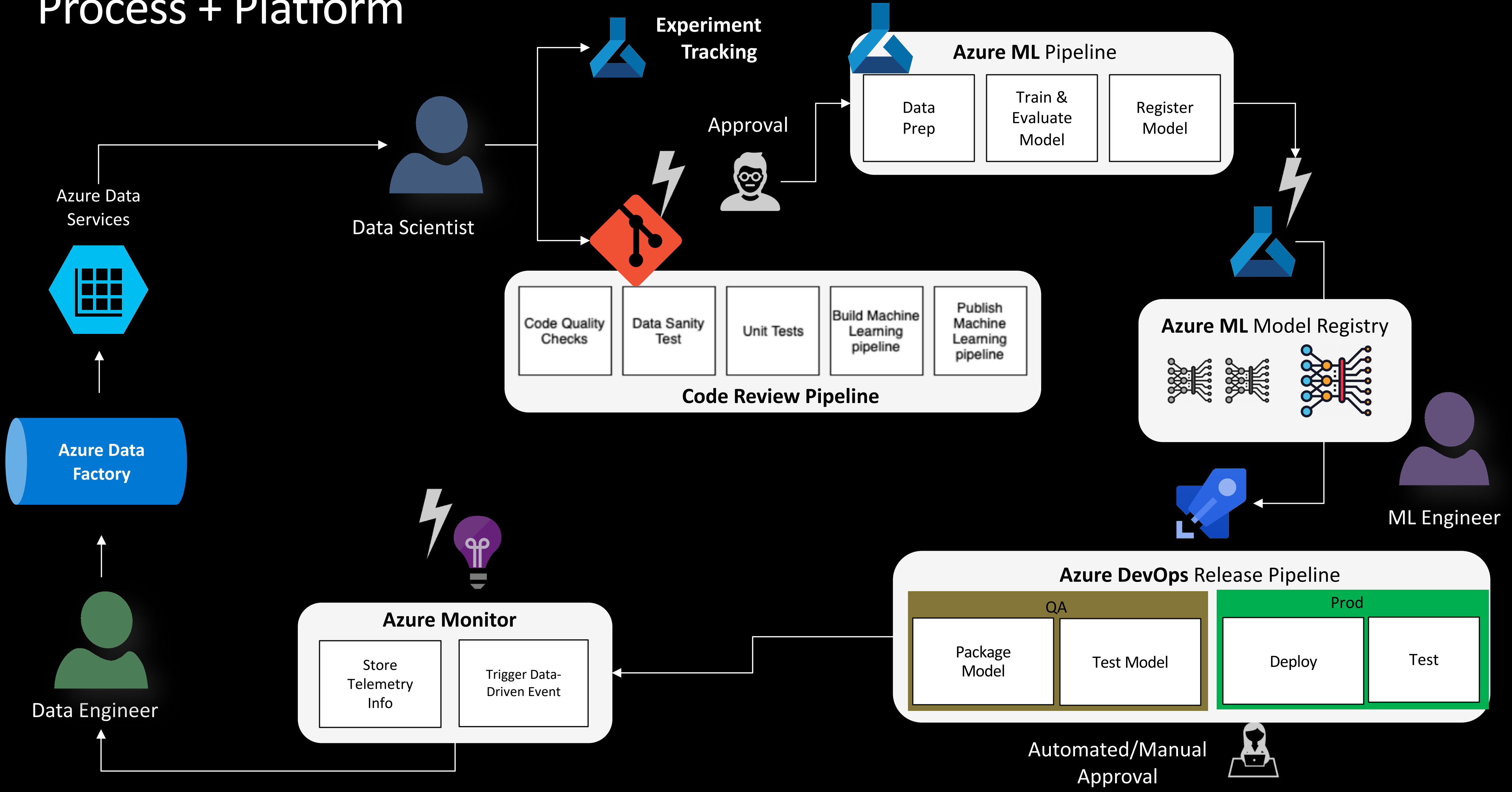
MLOps Process

Enterprise ready machine learning development

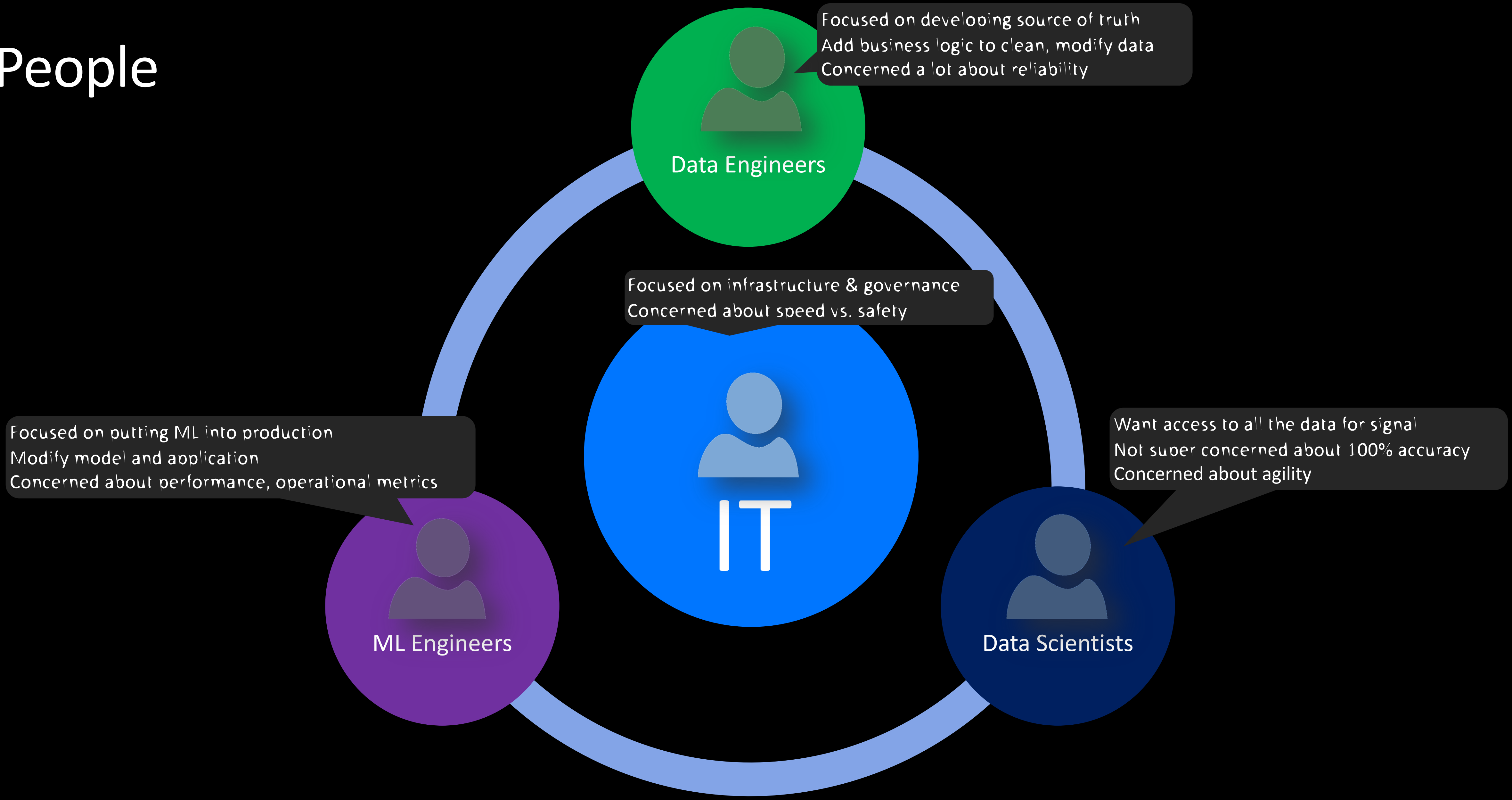


MLOps Demo

Process + Platform



People


Focused on developing source of truth
Add business logic to clean, modify data
Concerned a lot about reliability
Data Engineers
Focused on infrastructure & governance
Concerned about speed vs. safety
IT
Focused on putting ML into production
Modify model and application
Concerned about performance, operational metrics
ML Engineers
Want access to all the data for signal
Not super concerned about 100% accuracy
Concerned about agility
Data Scientists

Summary

People

- Machine Learning success will need culture change
- Different Personas have different needs

Process

- Need to have policies and governance with a self-service platform
- Integrate it with the rest of the IT/App development lifecycle

Platform

- Automatable and Reproducible
- Self-Service

MLOps Benefits

Reproducibility / Auditability

- Code drives **generation** and **deployments**
- Pipelines are **reproducible** and **verifiable**
- All artifacts can be **tagged** and **audited**

Validation

- SWE best practices for quality control
- Offline comparisons of model **quality**
- Minimize **bias** and enable **explainability**

Automation / Observability

- Controlled rollout capabilities
- Live comparison of predicted vs. expected performance
- Results fed back to watch for drift and improve model

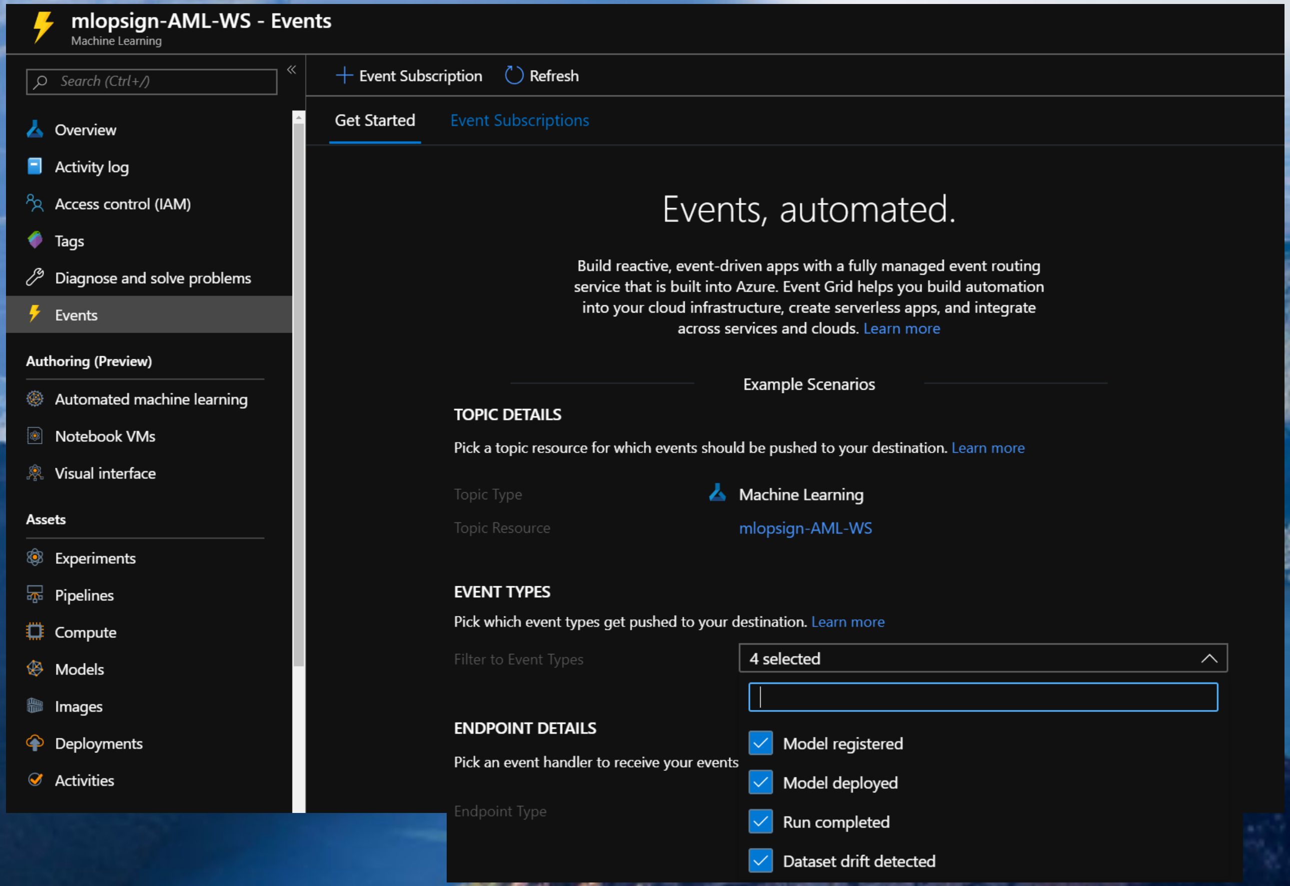
MLOps

Some additional capabilities…

Azure ML Event Grid integration

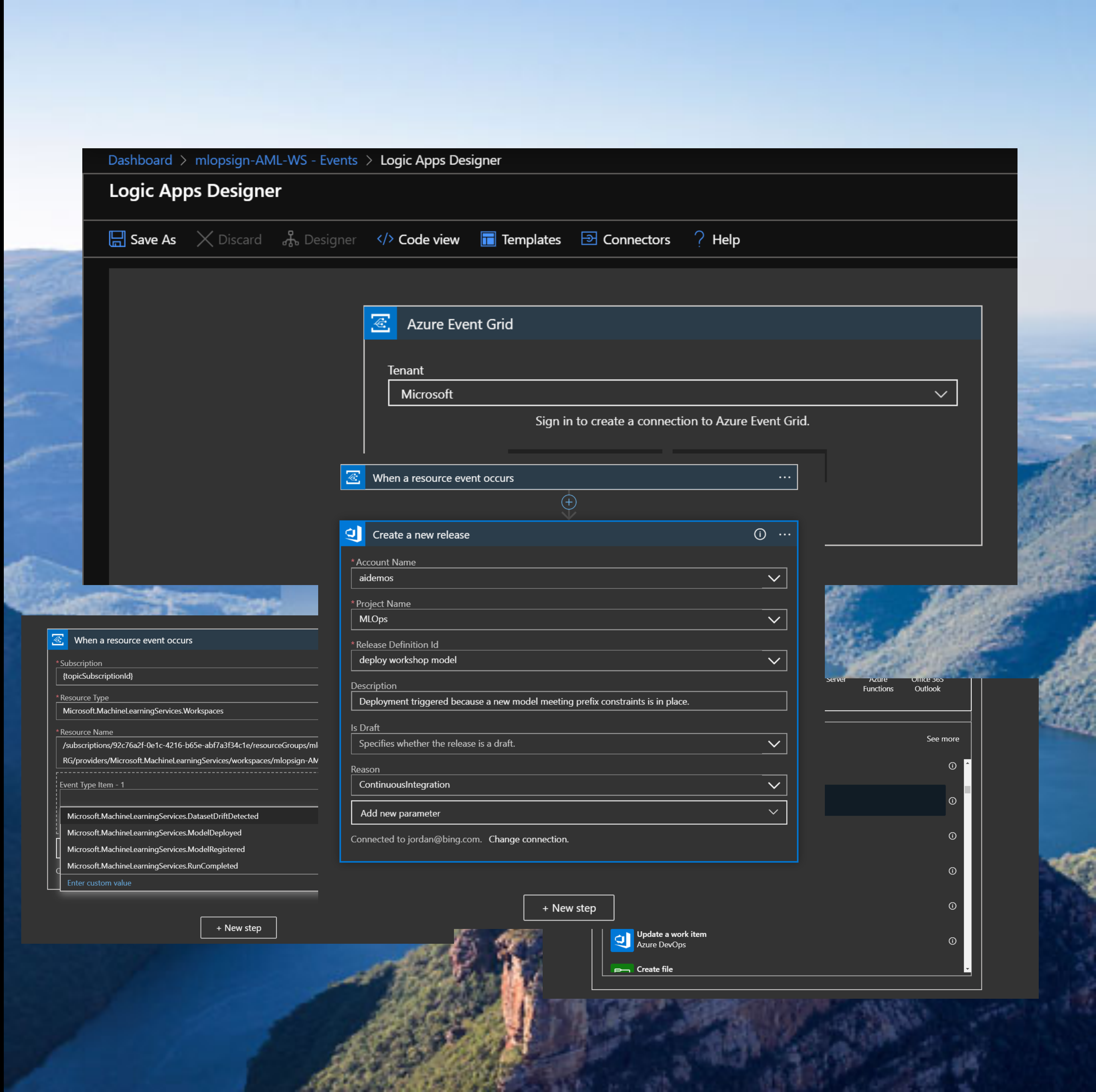
Fully managed event routing for all activities in the ML lifecycle

Let’s look at some examples…



Send emails / hit a webhook whenever a run completes

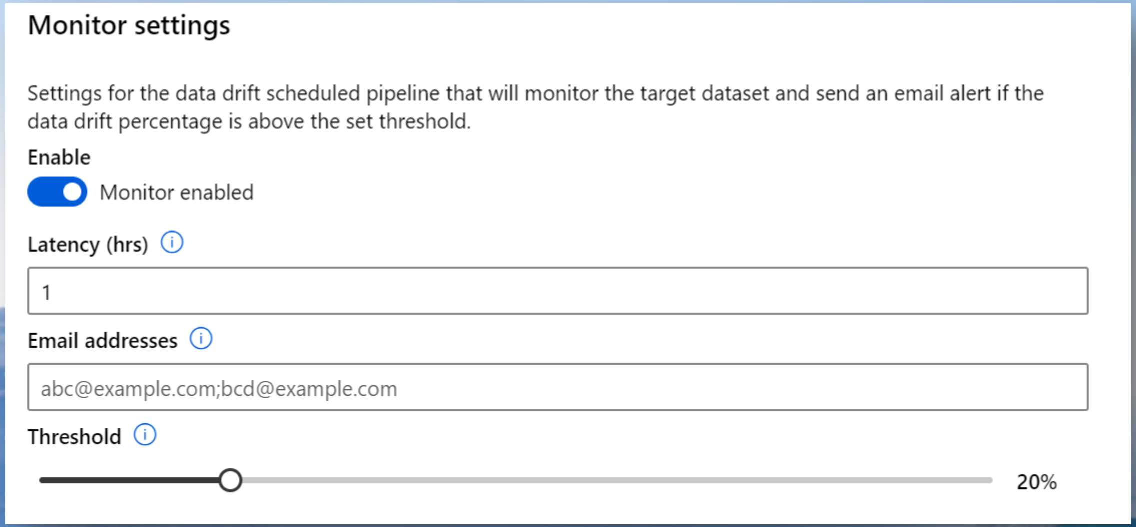
Streamline model reuse & consumption

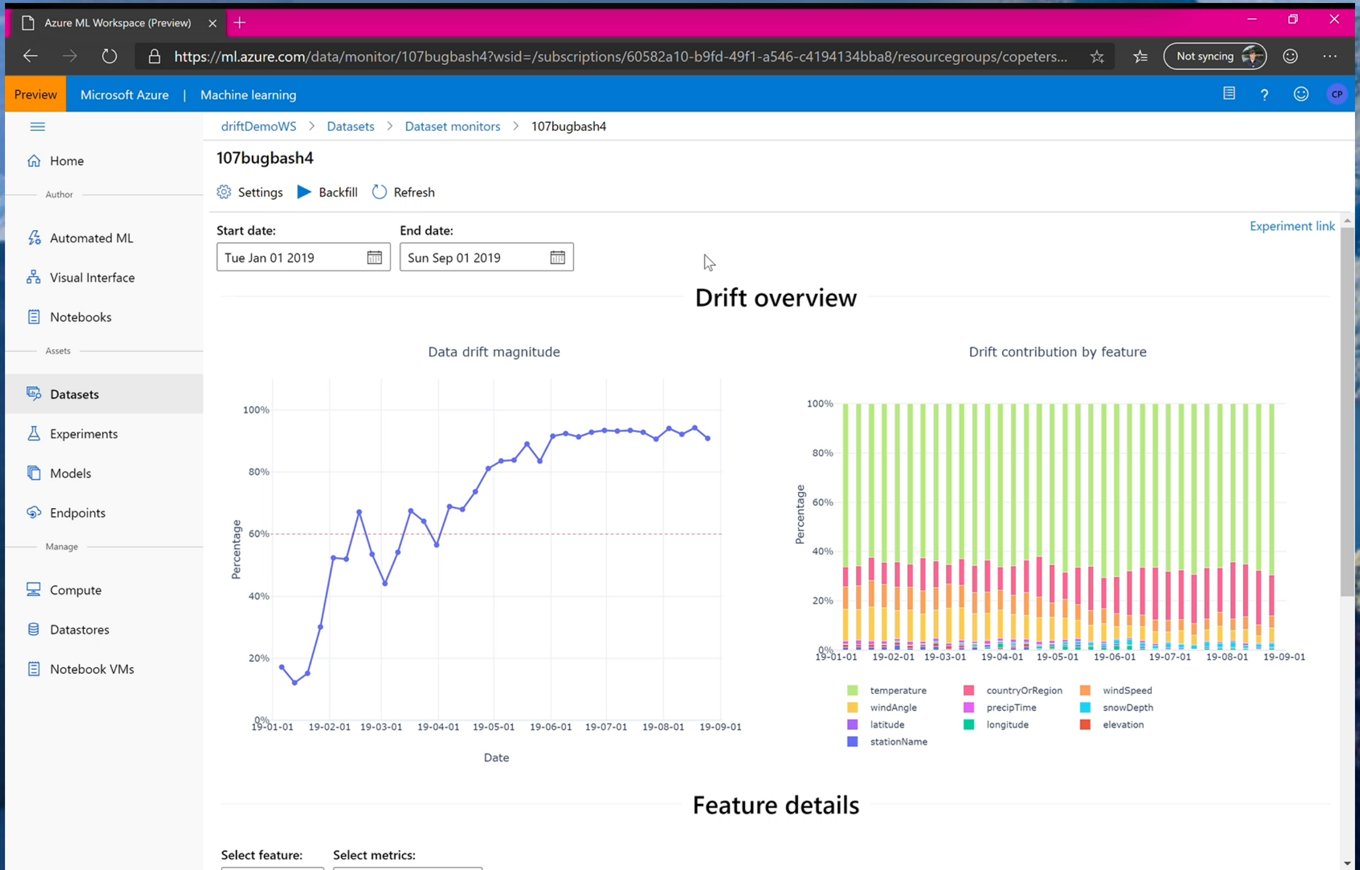


Set up a data drift monitor…

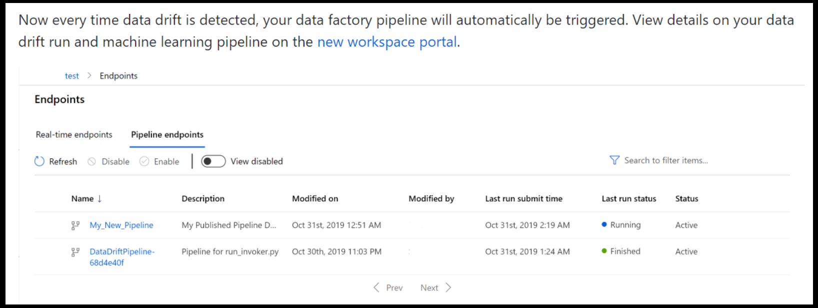
Compare datasets over time

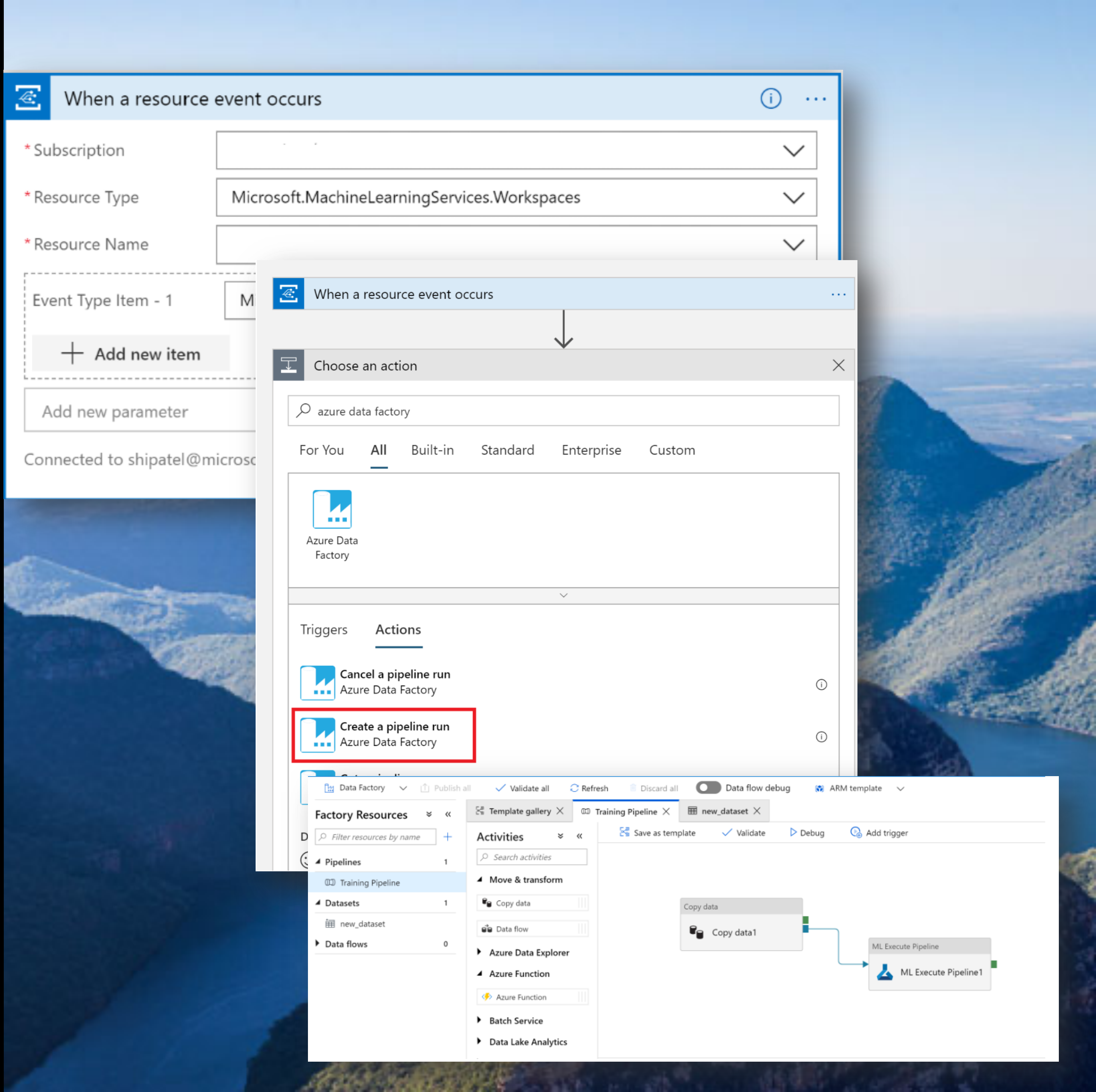
Determine when to take a closer look





Trigger Data Factory when data drift is detected





NEW MLOPS CAPABILITIES

Azure Machine Learning

Improved data management

ML pipelines YAML support

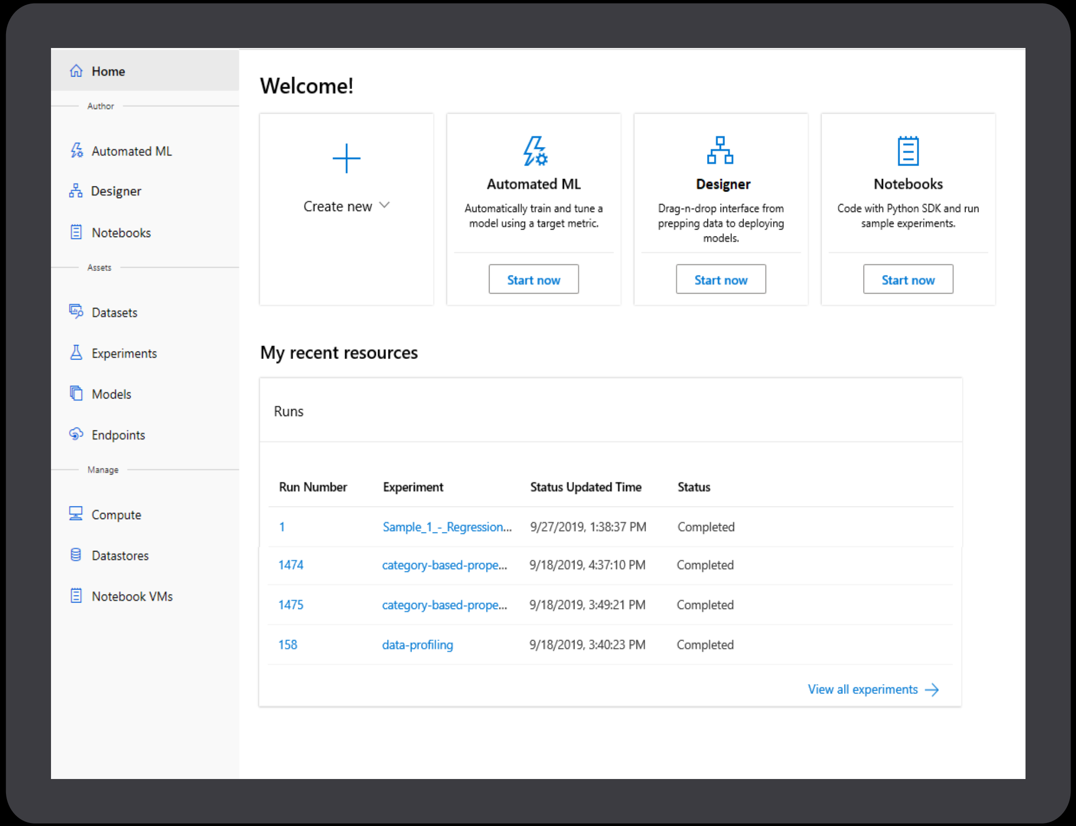
Event Grid integration

No-code model deployment

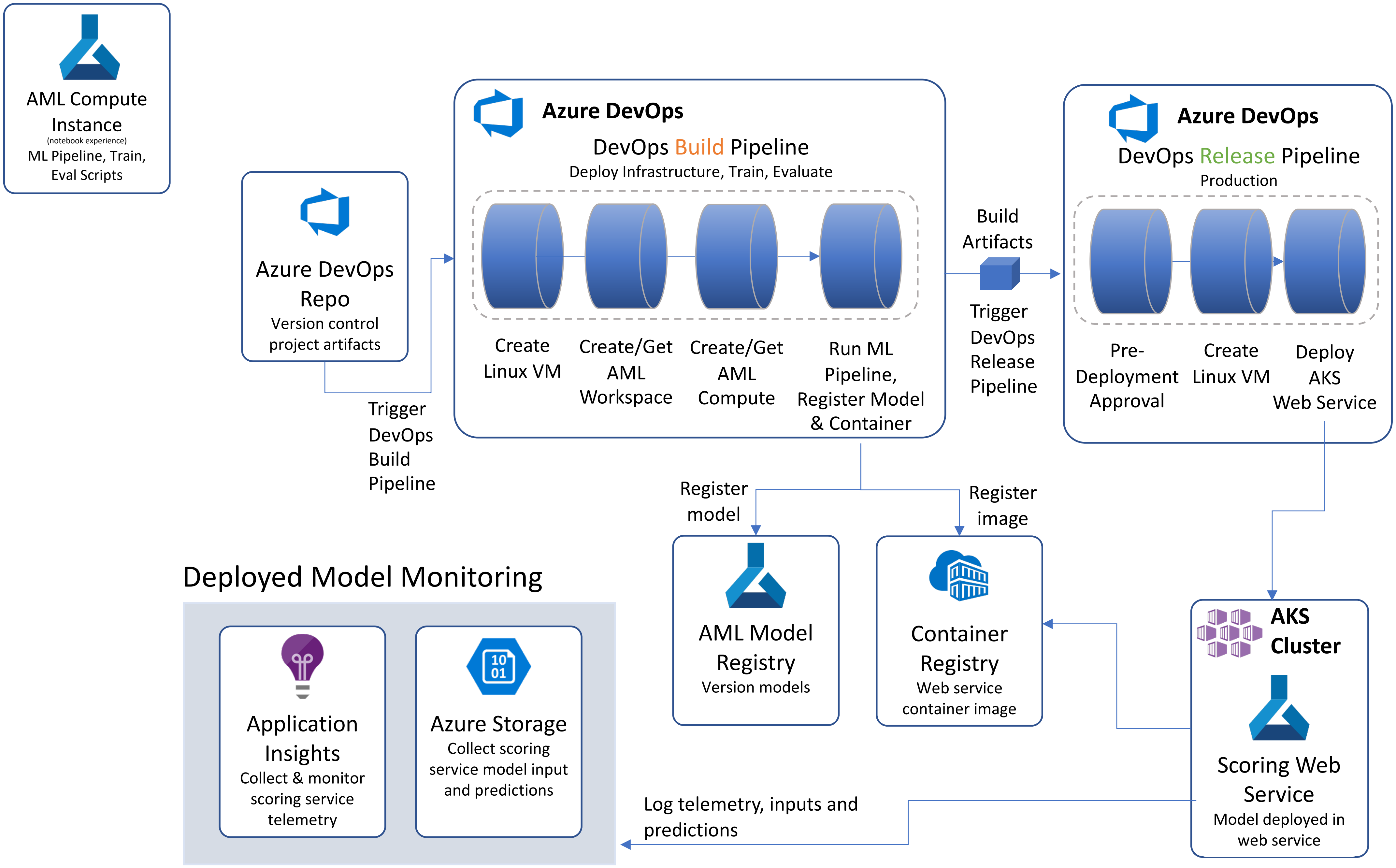
DevOps integration

Dataset drift analysis

Data Factory integration



Hands-on Lab


AML Compute Instance
(notebook experience)
ML Pipeline, Train, Eval Scripts
Azure DevOps Repo
Version control project artifacts
Trigger DevOps Build Pipeline
Azure DevOps
DevOps Build Pipeline
Deploy Infrastructure, Train, Evaluate
Create Linux VM
Create/Get AML Workspace
Create/Get AML Compute
Run ML Pipeline, Register Model & Container
Build Artifacts
Trigger DevOps Release Pipeline
Azure DevOps
DevOps Release Pipeline
Production
Pre-Deployment Approval
Create Linux VM
Deploy AKS Web Service
Register model
Register image
AML Model Registry
Version models
Container Registry
Web service container image
AKS Cluster
Scoring Web Service
Model deployed in web service
Deployed Model Monitoring
Application Insights
Collect & monitor scoring service telemetry
Azure Storage
Collect scoring service model input and predictions
Log telemetry, inputs and predictions

Hands-on Lab – setup guide overview

Full workshop content:

https://github.com/tiagoh/MCW-ML-Ops-1Day

How to get started ?

Follow-up the Before the hands-on lab setup guide

Potencial questions:

- AML Workspace ?
- AML Compute Instance ?
- Options to execute noteboks in Azure ?

Wrap-up + Q&A

Azure ML Components

Experience
SDK, Notebooks, Drag-n-drop, Wizard

MLOps
Reproducible, Automatable, GitHub, CLI, REST

Datasets
Profiling, Drift, Labeling

Training
Experiments, Runs

Model Registry
Models, Images

Inferencing
Batch, Realtime

Compute
Jobs, Clusters, Instances

IoT Edge
Security, Mgmt, Deployment

Cloud
CPU, GPU, FPGA

Edge
CPU, GPU, NPU

New Capabilities - Azure MLOps

Asset management & orchestration services to help manage the ML lifecycle.

ML Experimentation

Training Services

- Script Execution
- AutoML
- HyperDrive

Run History

Compute

- ML Compute
- ML VMs
- Databricks

Storage

- Blob Store
- Files
- Data Lake

ML Ops

Asset Management

- **Environments** **
- **Code** **
- **Datasets** **
- Models

Orchestration Services

- **ML Pipelines**
- Model Packaging
- **Model Validation** **
- Model Deployment **
- **ML Audit Trail** **
- **Azure DevOps ML Extension** **

ML Inference

Real-Time
Azure Kubernetes Service

Batch
Azure ML Compute

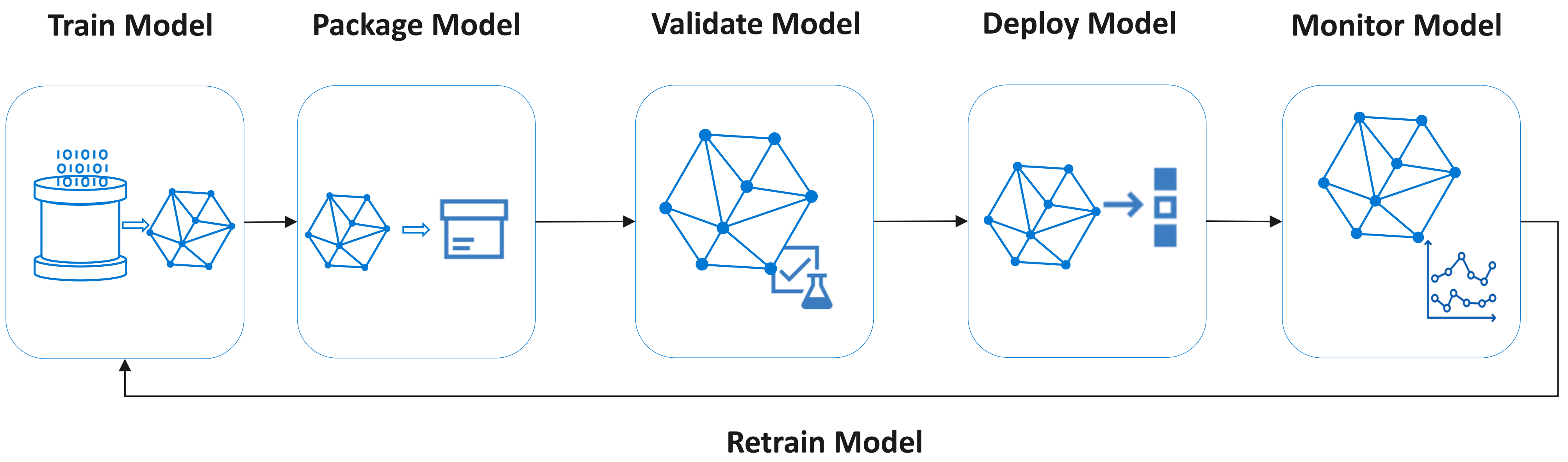
Edge
Azure IoT Hub

Monitoring

- Experimentation Monitoring
- Inference Monitoring
- **ML Data Drift** **

Recap—E2E ML Lifecycle

- **Develop & train model** with reusable ML pipelines
- **Package model** using containers to capture runtime dependencies for inference
- **Validate model behavior—**functionally, in terms of responsiveness, in terms of regulatory compliance
- **Deploy model—**to cloud & edge, for use in real-time/streaming/batch processing
- **Monitor model** behavior & business value, know **when to replace/deprecate a stale model**



MLOps accelerates going to Production

- We have pre-built solutions you can customize for your business
- We will continue to flesh these recipes out as we go
- Everything we build, we will build in the open!

Conclusions

MLOps makes it possible to do production-grade ML

Azure is investing heavily in **MLOps**

Try Azure ML for MLOps **today!**

https://aka.ms/mlops

Resources

Sources for this presentation:

- Ignite "Manage your end-to-end machine learning lifecycle with MLOps" | BRK3035
- Ignite: "Understanding enterprise readiness for machine learning solutions" | BRK2017

Related materials:

- R MLOps presentation

Thank You!

ευχαριστώ Salamat Po متشکرم شكراً Grazie

благодаря ありがとうございます Kiitos Teşekkürler 谢谢

ขอบคุณครับ Obrigado شكريه Terima Kasih Dziękuję

Hvala Köszönöm Tak Dank u Wel дякую Tack

Mulţumesc спасибо Danke Cám ơn Gracias

多謝晒 Ďakujem תודה நன்றி Děkuji 감사합니다