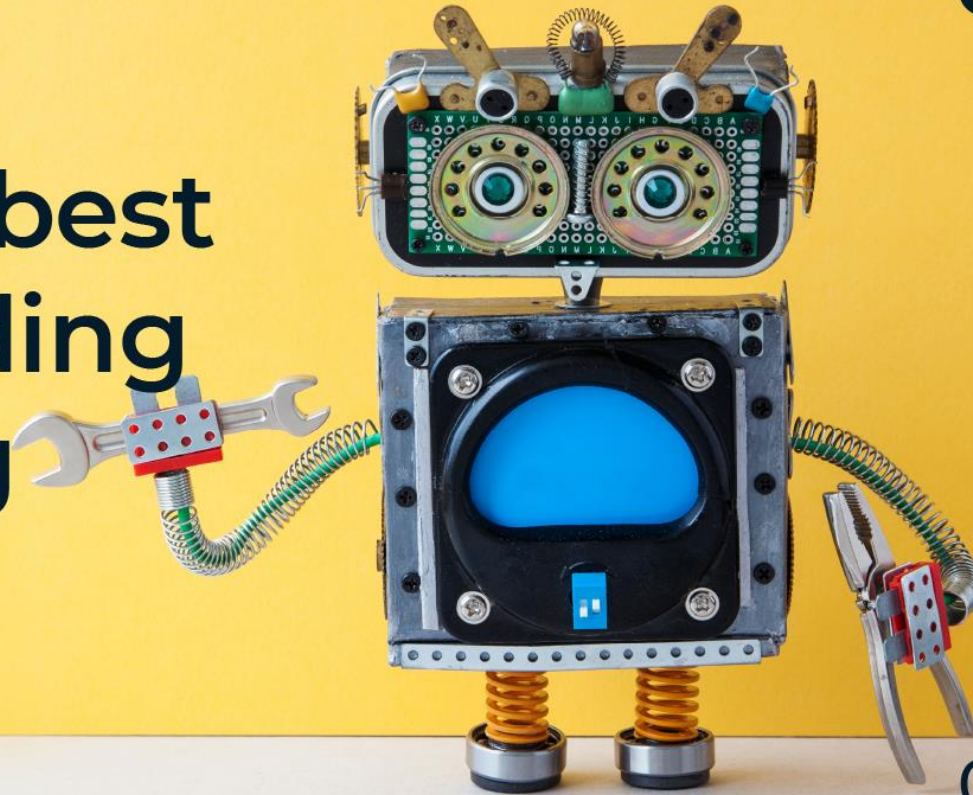


Virtual Coffee

MLOps. Software best practices for building Machine Learning solutions



**plain
concepts**



Rediscover
the meaning of technology



Rediscover
the meaning of technology

Founded in 2006 by 4 Microsoft MVPs, Plain Concepts was created to help companies adopt new technologies aimed at improving their productivity and processes.

Awarded in 2016 as **Microsoft Partner of the Year**, we currently have over **350 employees**, reaching a milestone in the technology sector by having **12 professionals** recognized as **Microsoft MVP** and over a dozen certifications at business level.

Present in **Spain, USA, UAE, United Kingdom, Germany and the Netherlands**, we have developed over 2,000 projects for companies across all industrial vertical.

Awards



Microsoft Partner | Mixed Reality



12 ★★★★★★★★★★
MICROSOFT
MOST VALUE
PROFESSIONAL
★★★★★★★★★★



MICROSOFT GOLDEN PROVIDER

- Application Development
Cloud Platform
- Cloud Productivity
Collaboration and
Content Datacenter
- DevOps



MICROSOFT SILVER PROVIDER

- Application Integration
Data Analytics
- Data Platform
Messaging
- Small and Midmarket
Cloud Solutions
- Windows and Devices

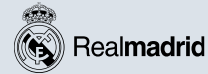
★★★★★★★★★★
AGILE ALLIANCE
CORPORATE MEMBER
★★★★★★★★★★

★★★★★★★★★★
XAMARIN PREMIUM
CONSULTING PARTNER
★★★★★★★★★★

★★★★★★★★★★
ALM PARTNER OF THE YEAR
FOR 7 CONSECUTIVE YEARS
★★★★★★★★★★

Our customers

Our customers are vast and cover a variety of business such as tourism and hospitality, construction, energy, telecommunications, banking, media, communications, entertainment and creative agencies.

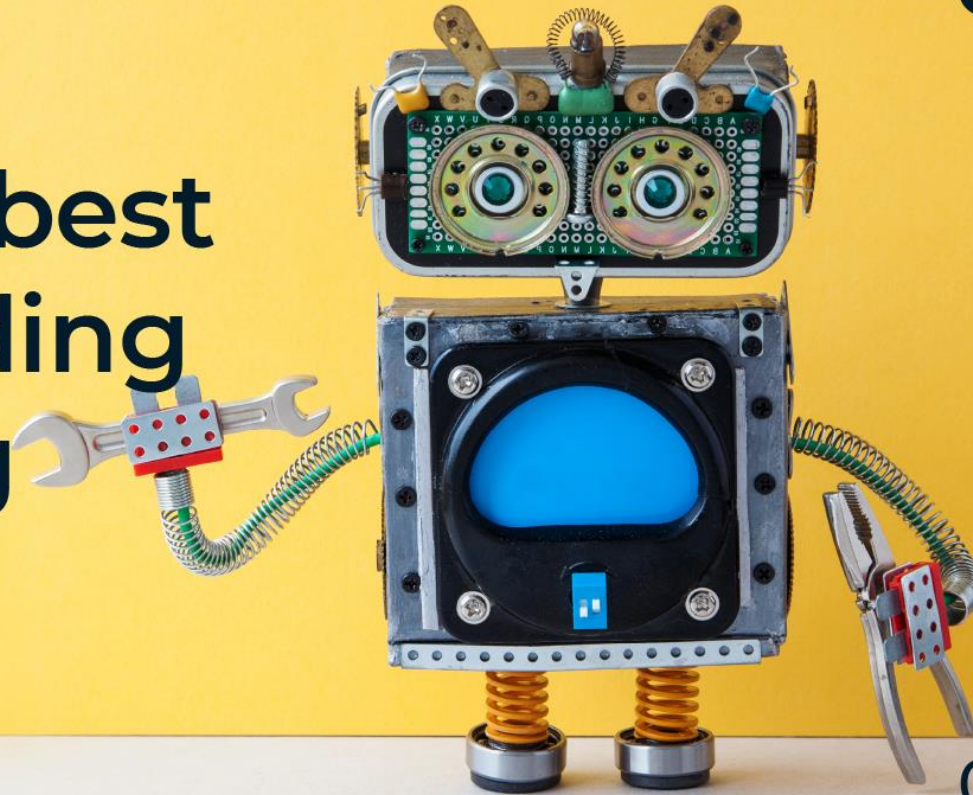


Our services



Virtual Coffee

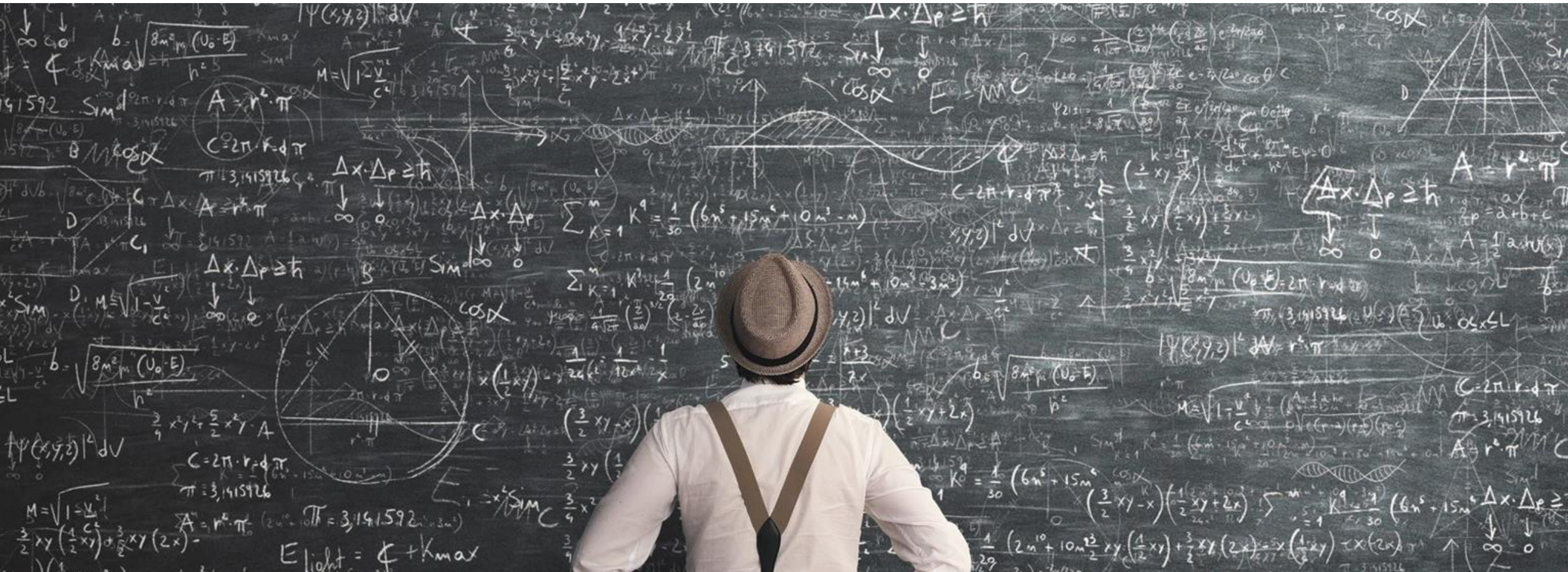
MLOps. Software best practices for building Machine Learning solutions



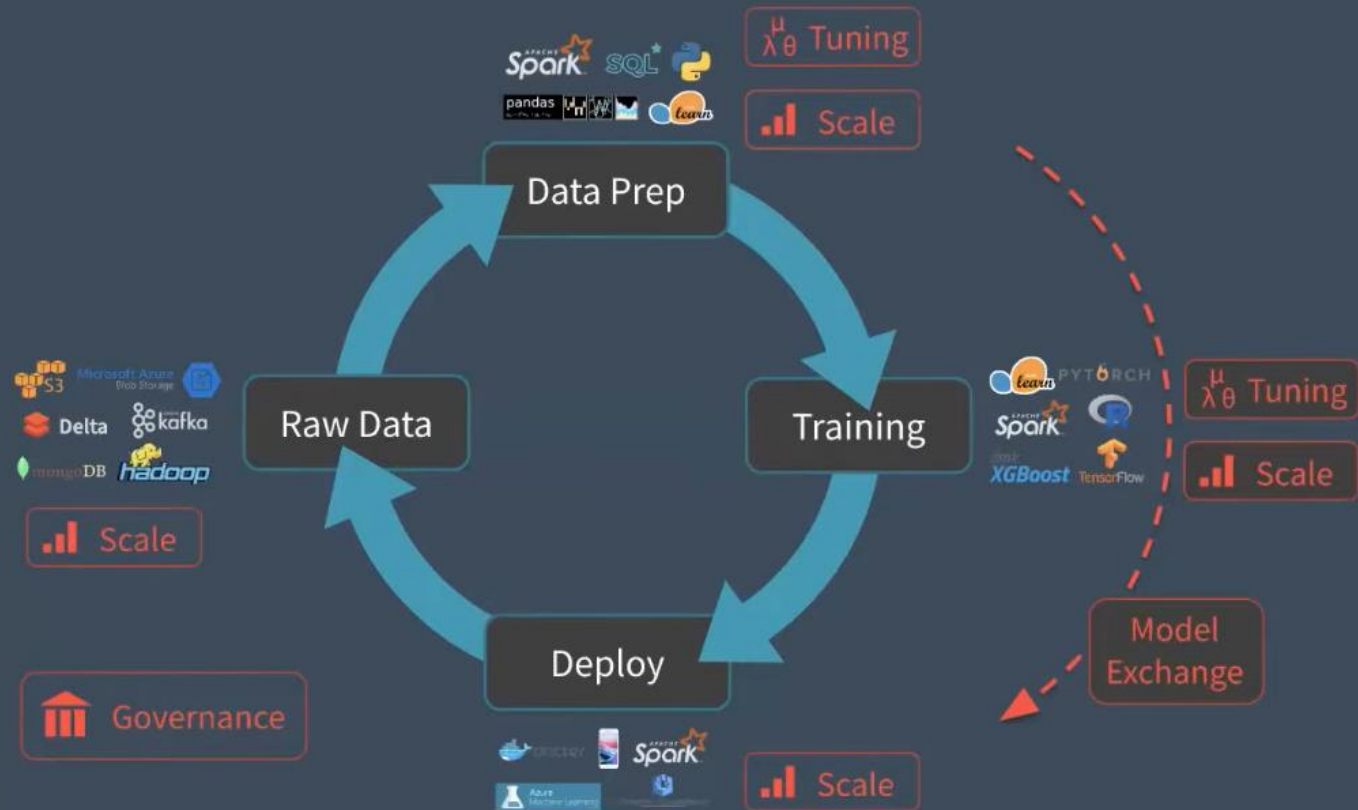
Agenda

1. Machine Learning is hard
2. MLOps **Components**
3. MLOps **Teams and Processes**
4. MLOps **Flows**

Machine Learning is hard



Machine Learning **lifecycle**



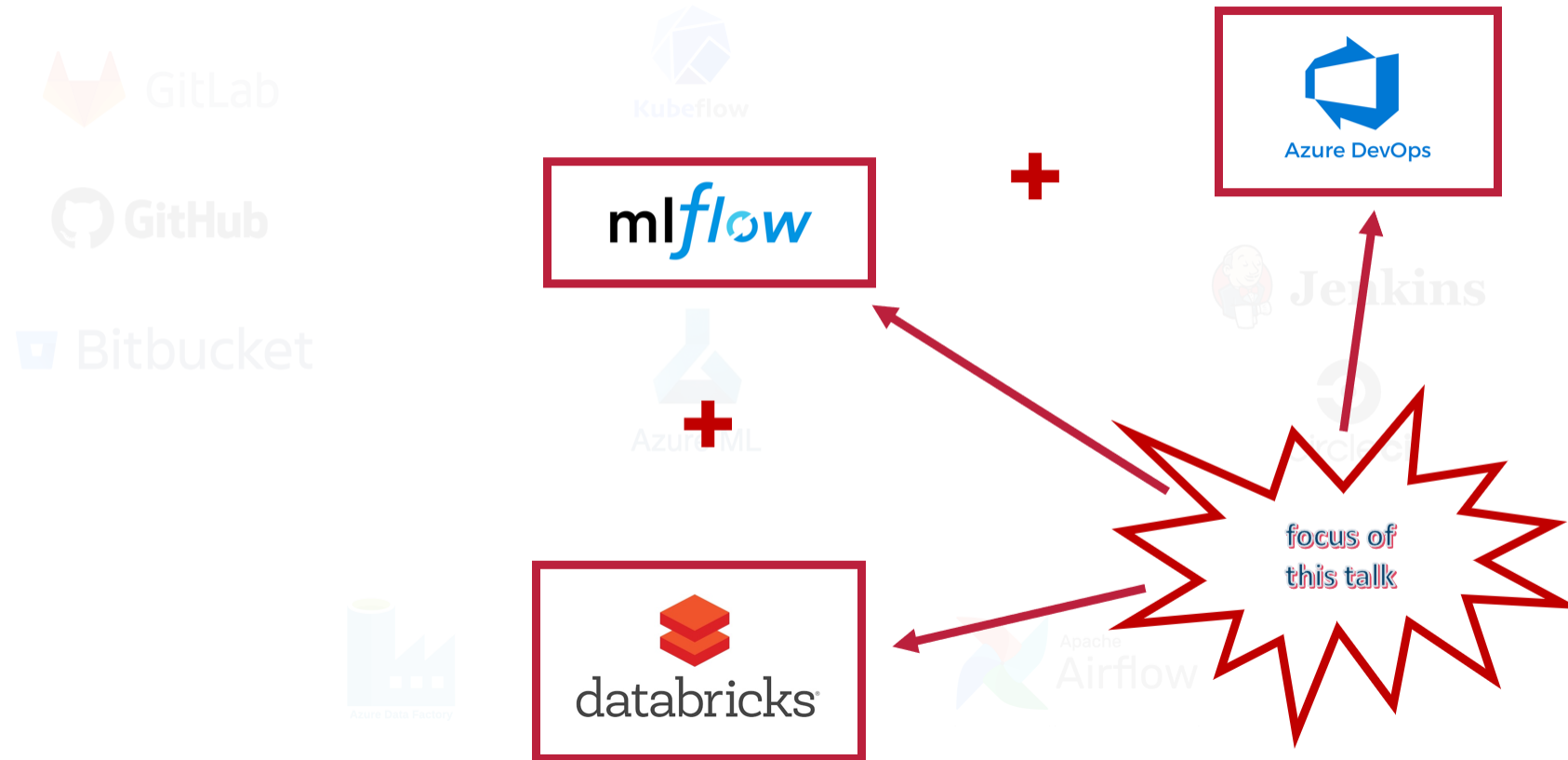
Machine Learning **best practices**

“

MLOps is a new practice for collaboration and communication between data scientists and IT professionals while **automating and productizing** machine learning algorithms

”

MLOps Components



UNIFIED DATA ANALYTICS PLATFORM



DATA ENGINEERS DATA SCIENTISTS ML ENGINEERS DATA ANALYSTS



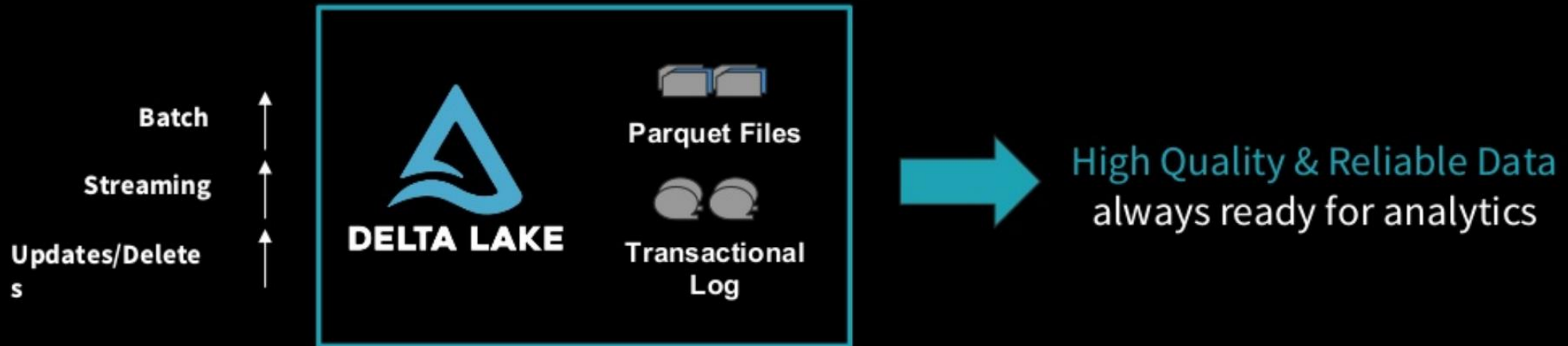
Accelerating data-driven
innovation across data
science, data engineering,
and business analytics



RAW DATA LAKE

<https://databricks.com/>

Delta Lake ensures data reliability



Key Features

- ACID Transactions
- Schema Enforcement
- Unified Batch & Streaming
- Time Travel/Data Snapshots



Azure Boards

Agile planning tools

Track work with configurable Kanban boards, interactive backlogs, and powerful planning tools. Unparalleled traceability and reporting make Boards the perfect home for all your ideas—big and small.

Azure Repos

Unlimited free private repos

Get flexible, powerful Git hosting with effective code reviews and unlimited free repositories for all your ideas—from a one-person project to the world's largest repository.

Azure Artifacts

Universal package repository

Share Maven, npm, NuGet, and Python packages from public and private sources with your entire team. Integrate package sharing into your CI/CD pipelines in a way that's simple and scalable.

Azure Pipelines

CI/CD for any platform

Build, test, and deploy in any language, to any cloud—or on-premises. Run in parallel on Linux, macOS, and Windows, and deploy containers to individual hosts or [Kubernetes](#).

Azure Test Plans

Manual and exploratory testing

Test often and release with confidence. Improve your overall code quality with manual and exploratory testing tools for your apps.

focus of
this talk

<https://azure.com/devops>

mlflow

Open source machine learning platform

- Works with any ML library, algorithm, language, etc
- *Open interface* design (use with any code you already have)

mlflow Tracking

Record and query
experiments: code,
data, confs, results

mlflow Projects

Packaging format
for reproducible
runs and workflows

mlflow Models

General format
that standardizes
deployment paths

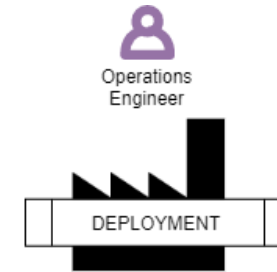
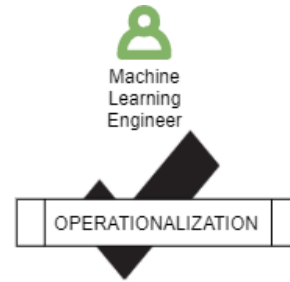
new

mlflow Model Registry

Centralized model
management,
review & sharing

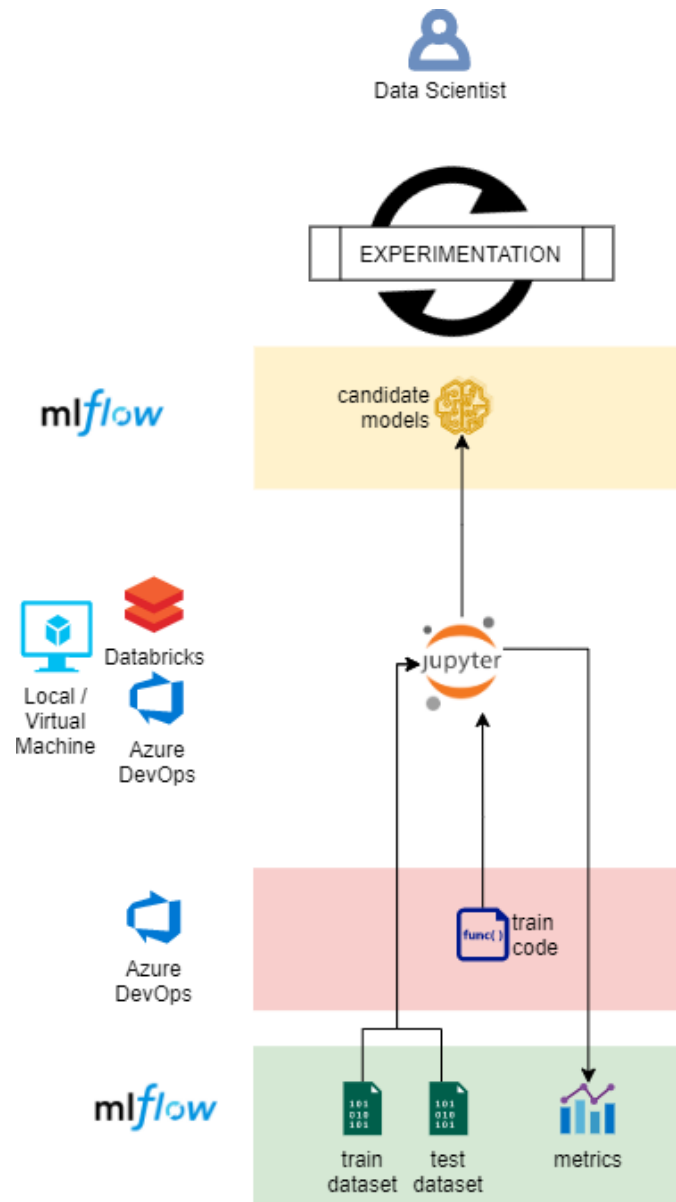
<https://mlflow.org/>

MLOps Teams and Processes



The three-legged problem





EXPERIMENTATION

Model Selection (mlFlow Models)

EDA (Exploratory Data Analysis)

Preprocessing

- Data Collection
- Feature Engineering

Train Model

- Hyperparameter tuning

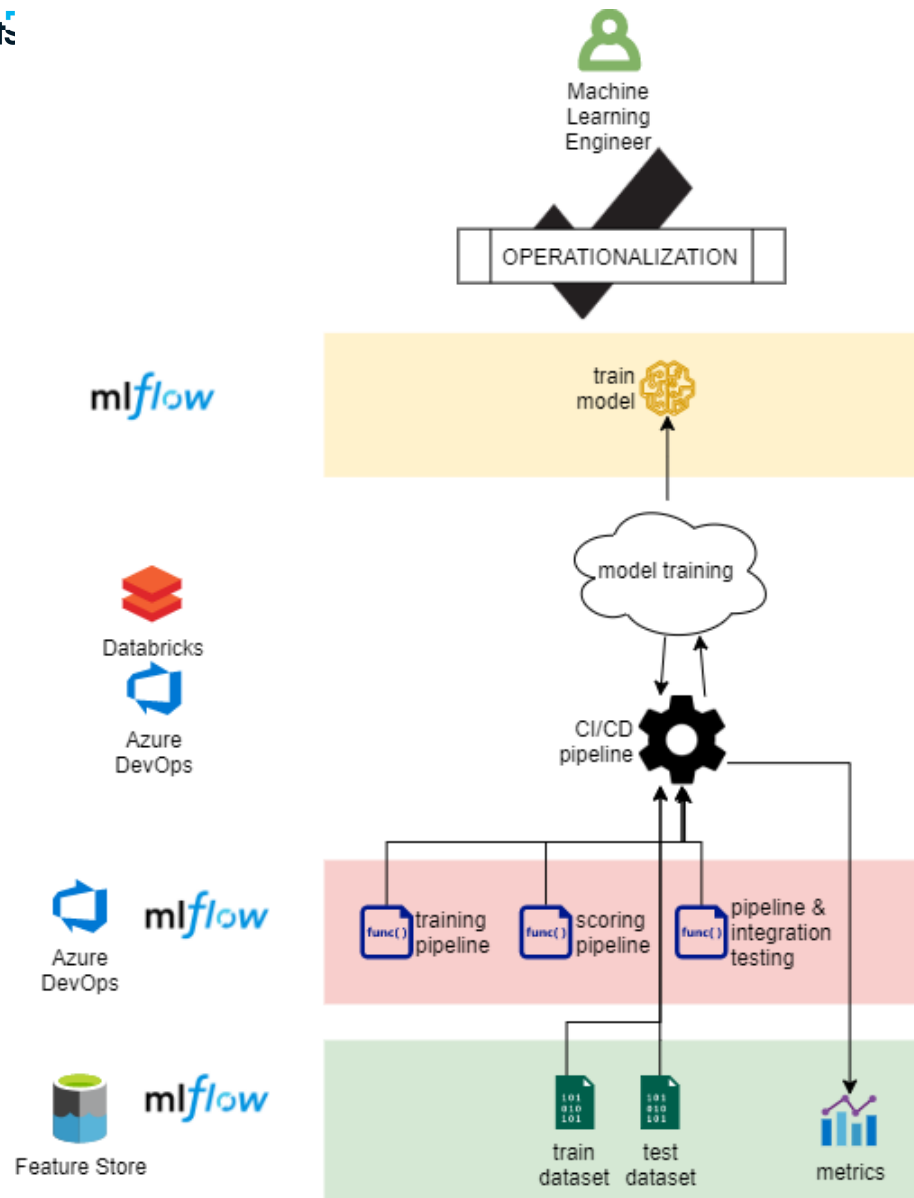
Evaluate Model

Feature branches

Notebooks (w/out Spaghetti code)

Dataset samples

Model Selection (mlFlow Tracking)



OPERATIONALIZATION

Model Store (mlFlow Models)

DevOps Pipelines:

- Build
- Release

Simple Orchestration:

- code
- gated releases

Infrastructure deployment

Code versioning

- Git flows
- mlFlow project

ML Pipelines (train and scoring)

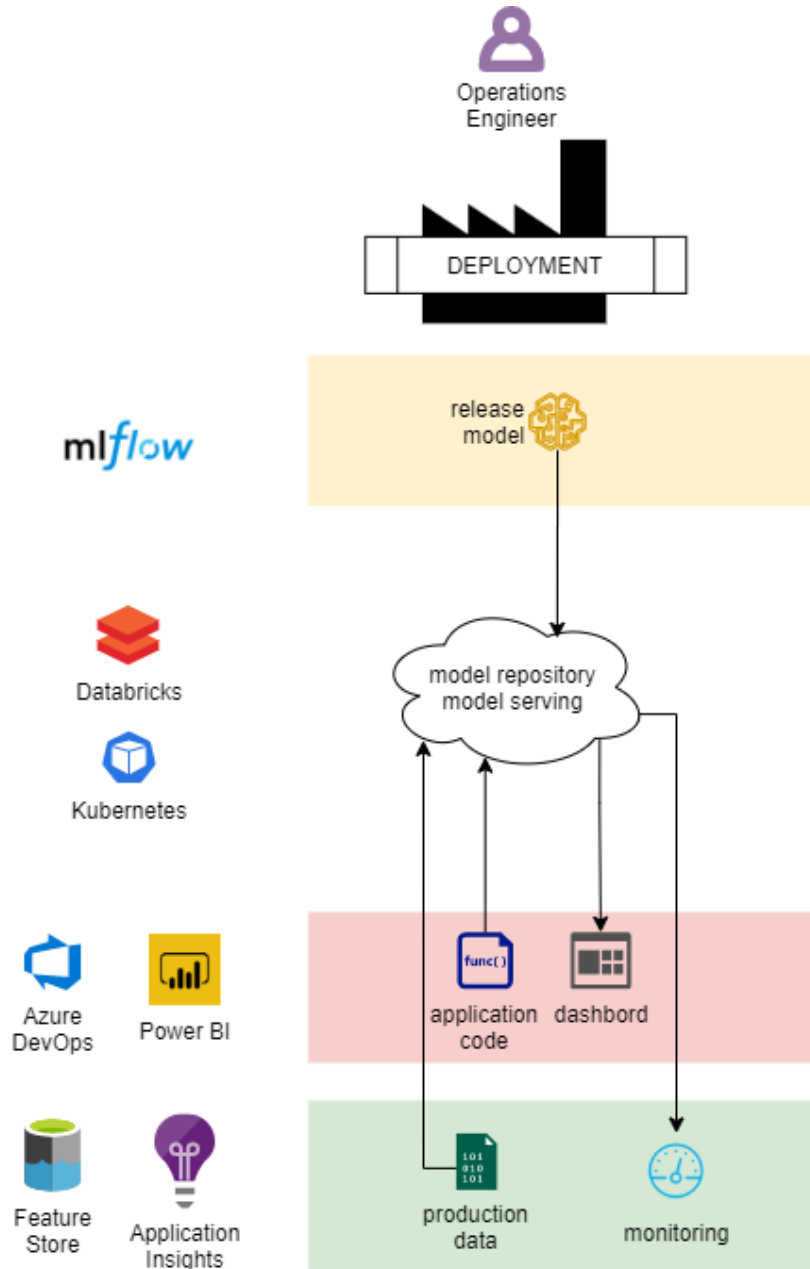
Unit and Integration testing

Train and Testing Dataset

Feature Store (data lineage)

Data Pipelines integration

Model evaluation (mlFlow tracking)



DEPLOYMENT

Model Registry (mlFlow Model Registry)

Model deployment:

- Single / Multiple model (partitioned models)
- Shadow model (canary deployment)
- Competing models (A/B testing)

Model serving:

- Embedded model (model + code)
- Model as a Service (REST API)
- Model as artifact (deployment by label)

Scalability:

- Kubernetes
- Databricks

Scoring:

- Batch (offline)
- Single Request (online)

Visualization

API's

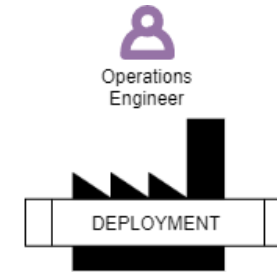
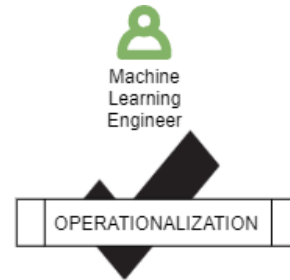
Analytics

Production Data (w/out past data)

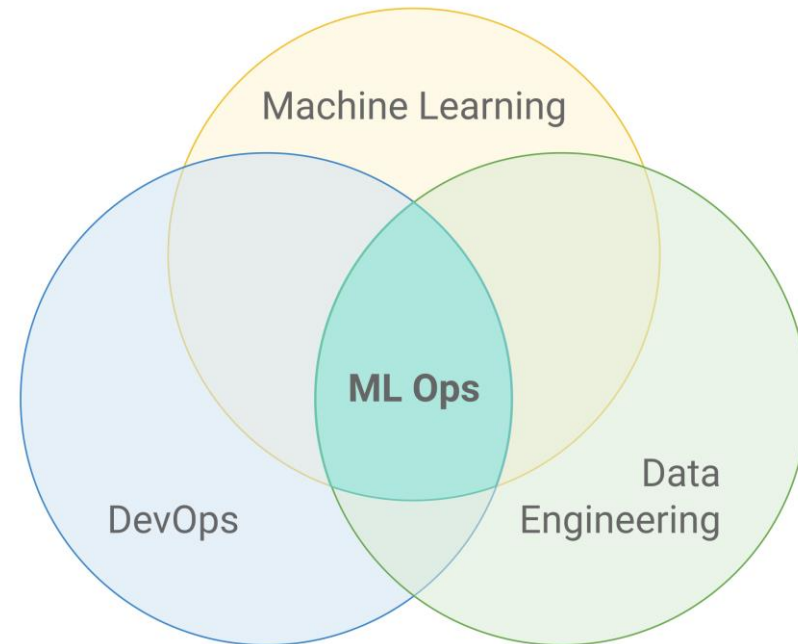
Monitoring:

- Concept drift
- Data drift

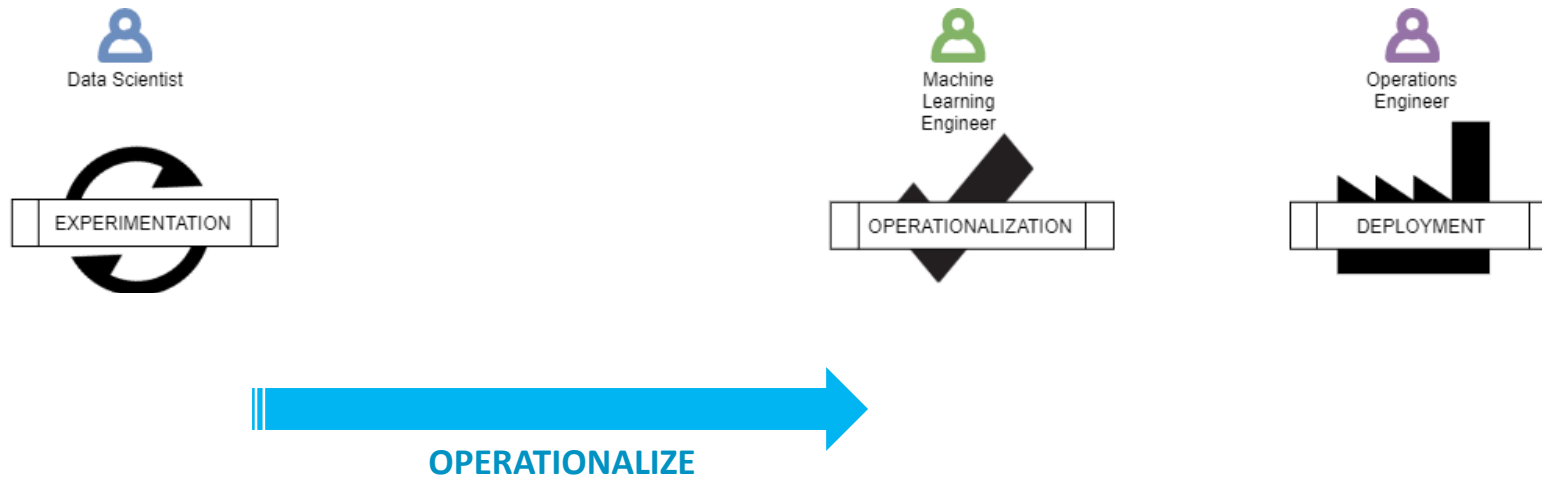
MLOps Teams and Processes



MLOps Teams and Processes

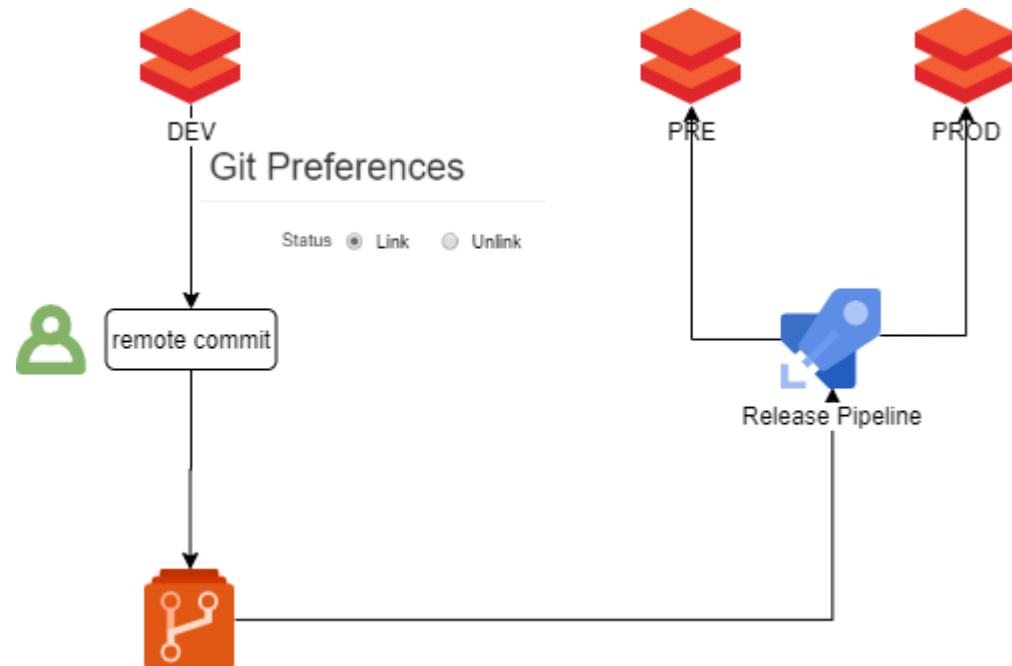


MLOps Flows



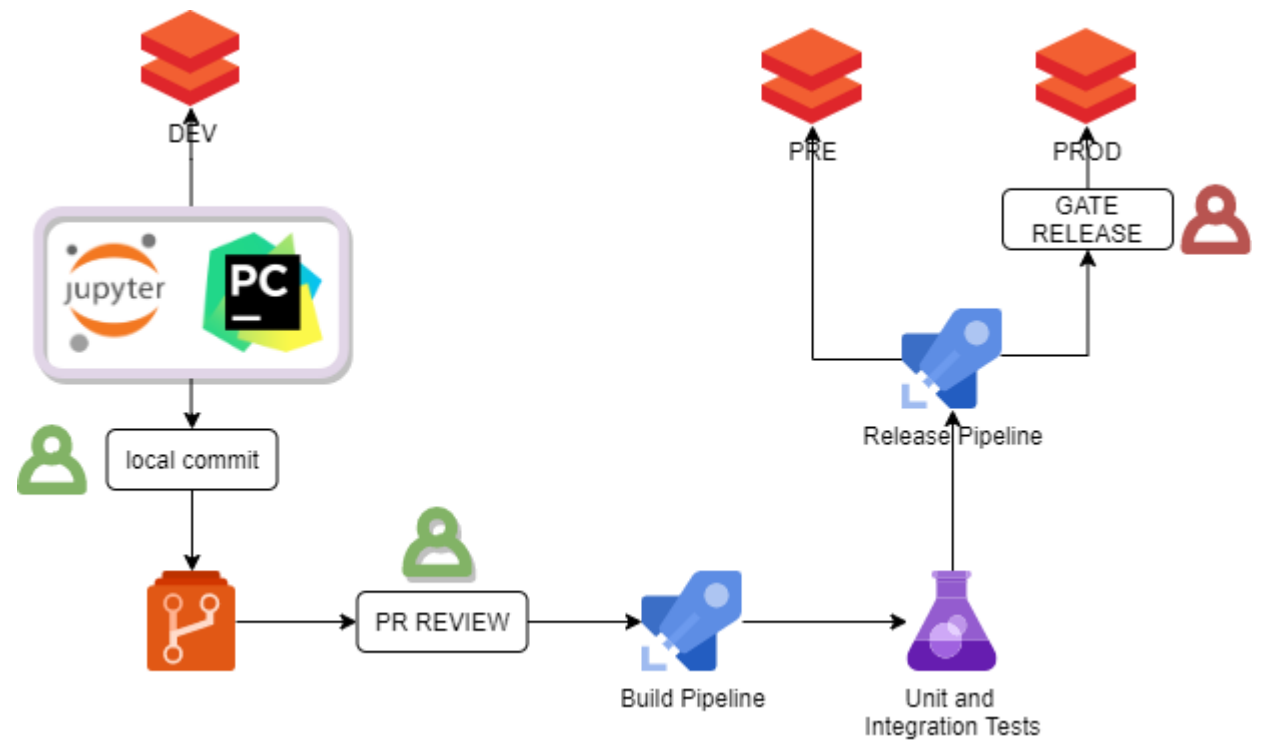
Databricks Development

Databricks based development

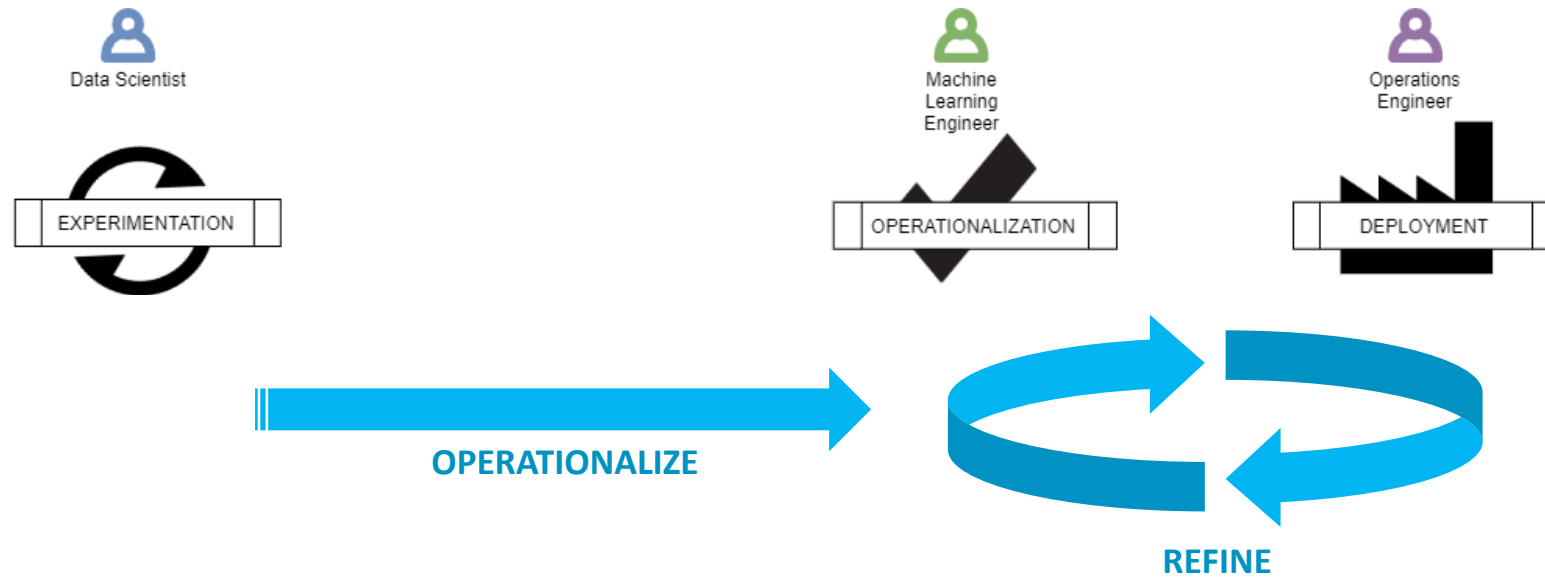


Databricks Development

Mixed based development



MLOps Flows





Thank you

fplopez@plainconcepts.com

kalbes@plainconcepts.com

www.plainconcepts.com