Speeding Up Machine Learning Development with MLflow

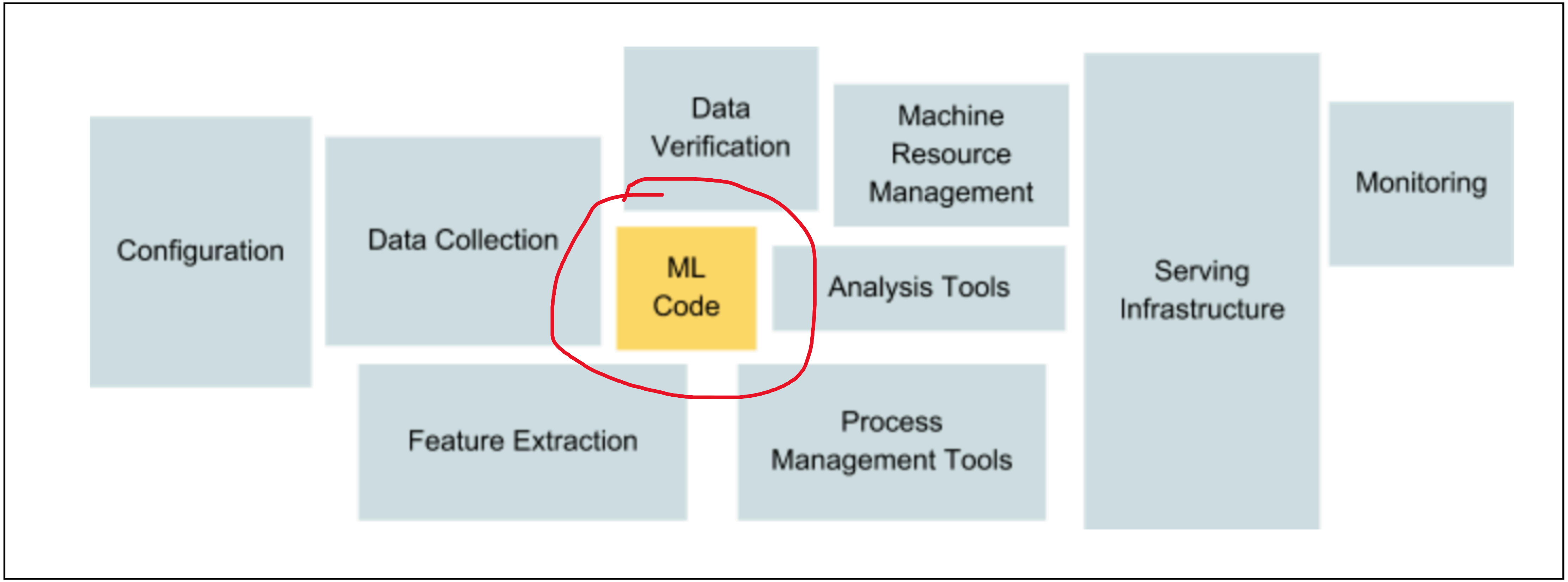
Hien Luu

Agenda

- Unique Challenges in ML Development Process
- Machine Learning Platform Tour
- Introduction to MLflow
  - Demo

ML Development Process Overview

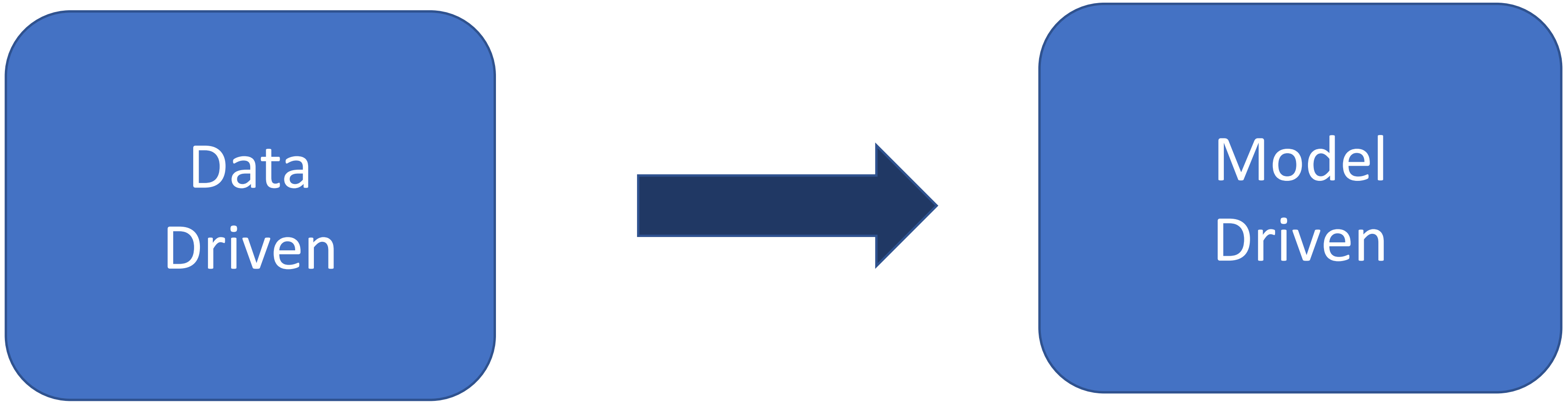
Hidden Technical Debt in Machine Learning Systems (paper from Google)



The required surrounding infrastructure is vast & complex

ML Development Process Overview

Software 2.0



New dawn of a new age of software

ML Development Process Overview

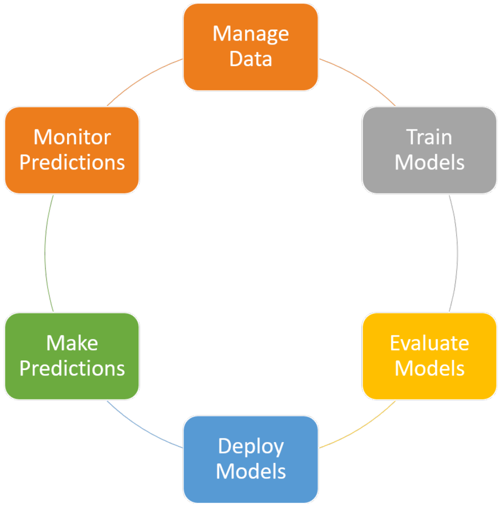
Software vs Model Development

	Software	Model
Goal	Meet function requirements	Optimize for business metric
Quality	Depend on code	Depend on data, code, algorithm, parameters
Tool	Standardized stack	Many tools and libraties
Outcome	Fairly deterministic	Changing with data

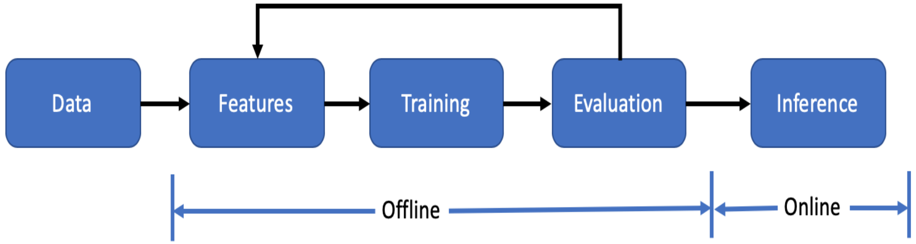
ML Development Process Overview

Machine Learning Development Dimensions

Experimentation



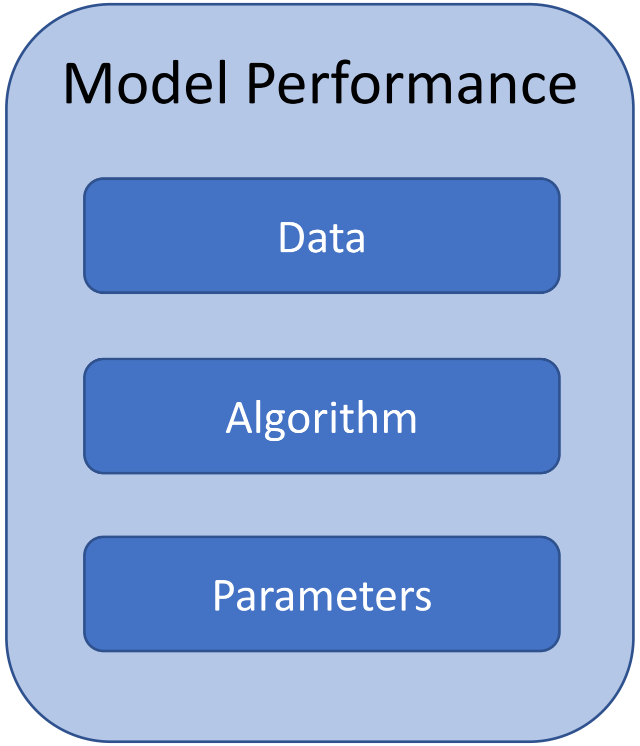
Environments

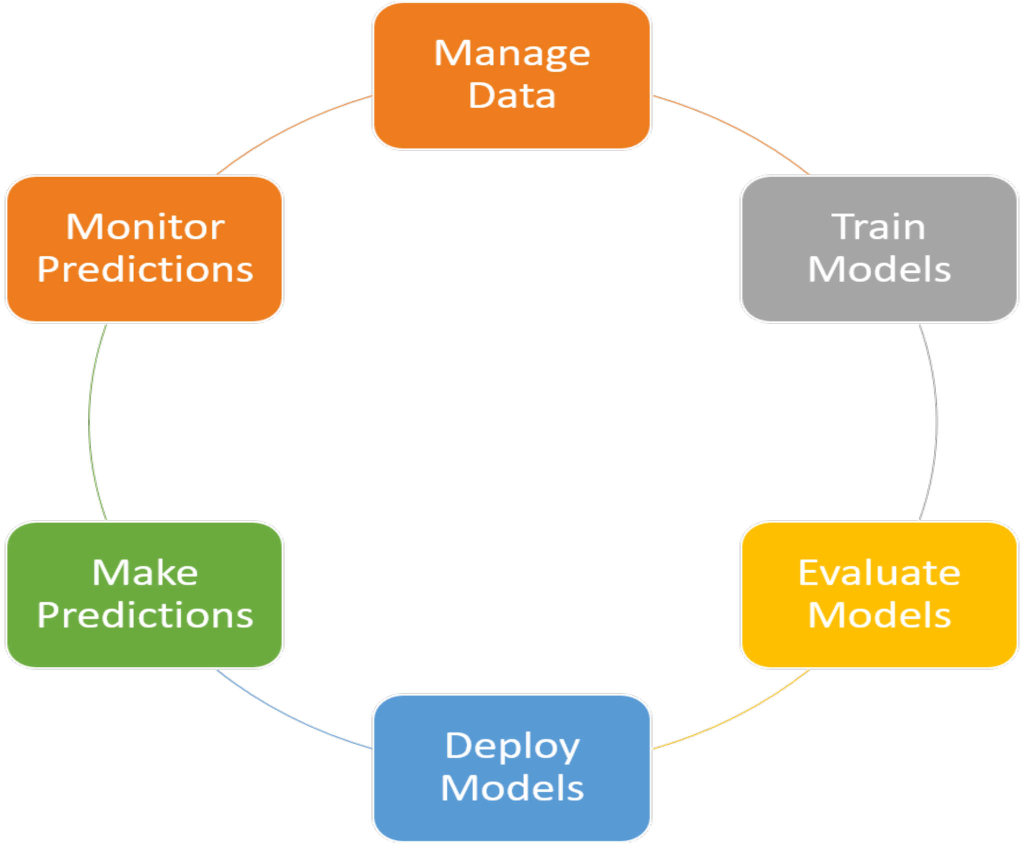


Big Data

ML Development Process Overview

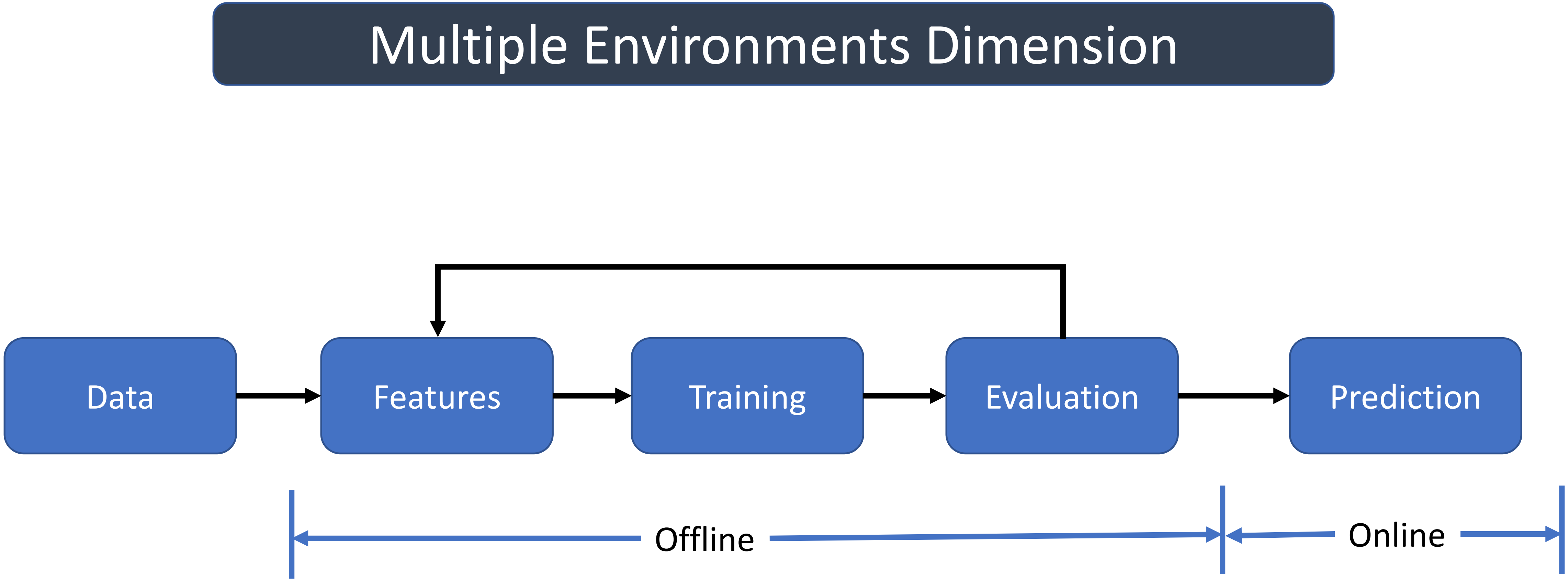
Experimentation Dimension





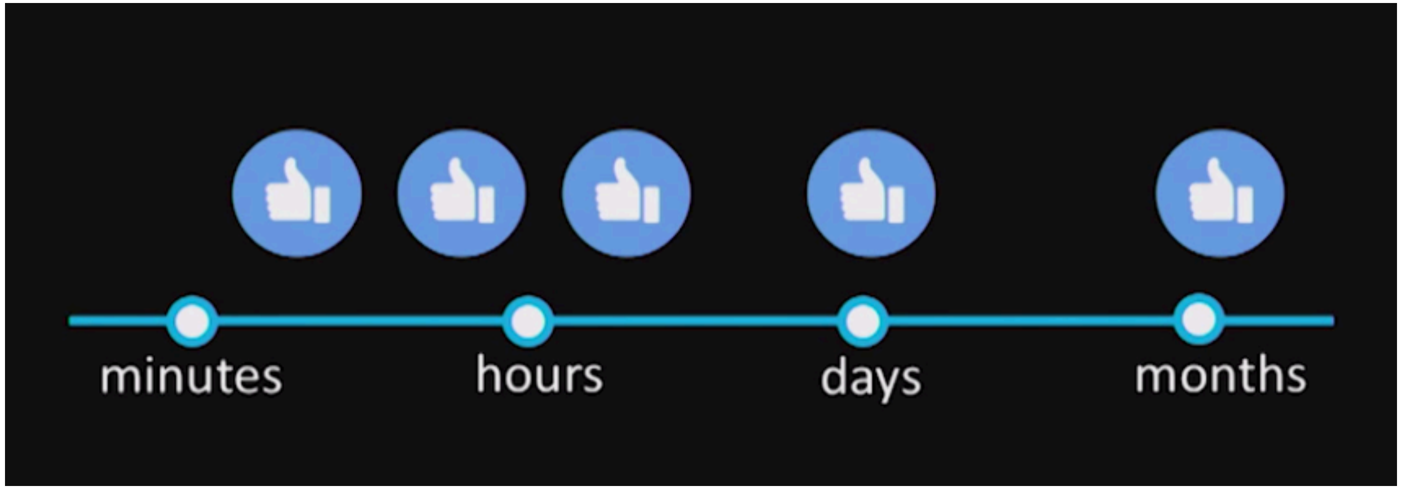
Machine Learning develoolment is a scientific endeavor

ML Development Process Overview



ML Development Process Overview

Data Volume Dimension

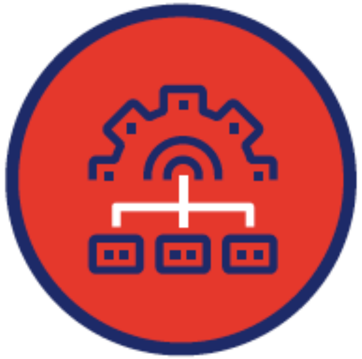


Machine Learning Pipeline & ML Training Infrastructure

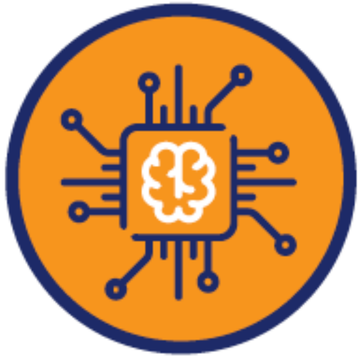
Machine Learning Platform Tour

Other Challenges Moving To Model Driven

People



Process



Technology

Scientific exploration and engineering rigorous and automation

ML Development Process Overview

Separation of Concerns

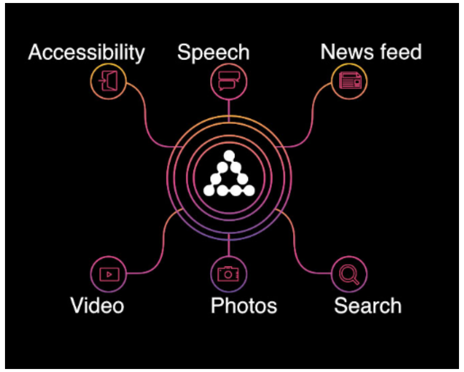
https://medium.com/netflix-techblog/open-sourcing-metaflow-a-human-centric-framework-for-data-science-fa72e04a5d9

ML Development Process Overview

ML development vs traditional software development

Machine Learning Platform Tour

In-house Machine Learning Platforms



FBLearner

AI Backbone

UBER

Michelangelo

ML as-a-service

Google TFX

Deploying production ML pipelines

ML as-a-service

Automation, productivity and fast iteration

Machine Learning Platform Tour

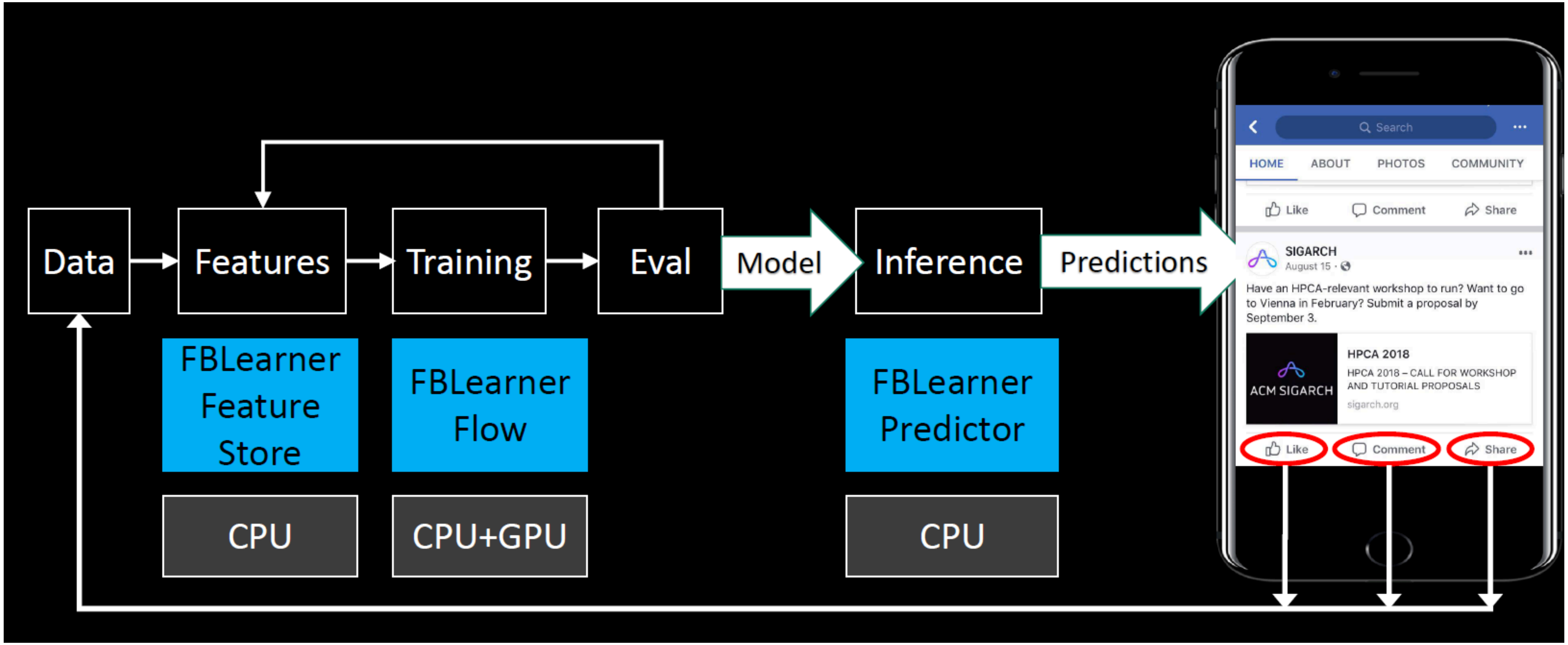
Anatomy of Machine Learning Platform

- Feature Infrastructure
- Model Training
- Model Management
- Model Serving
- Model Monitoring

Minimize incidental complexity in Machine Learning to increase efficiency

Machine Learning Platform Tour

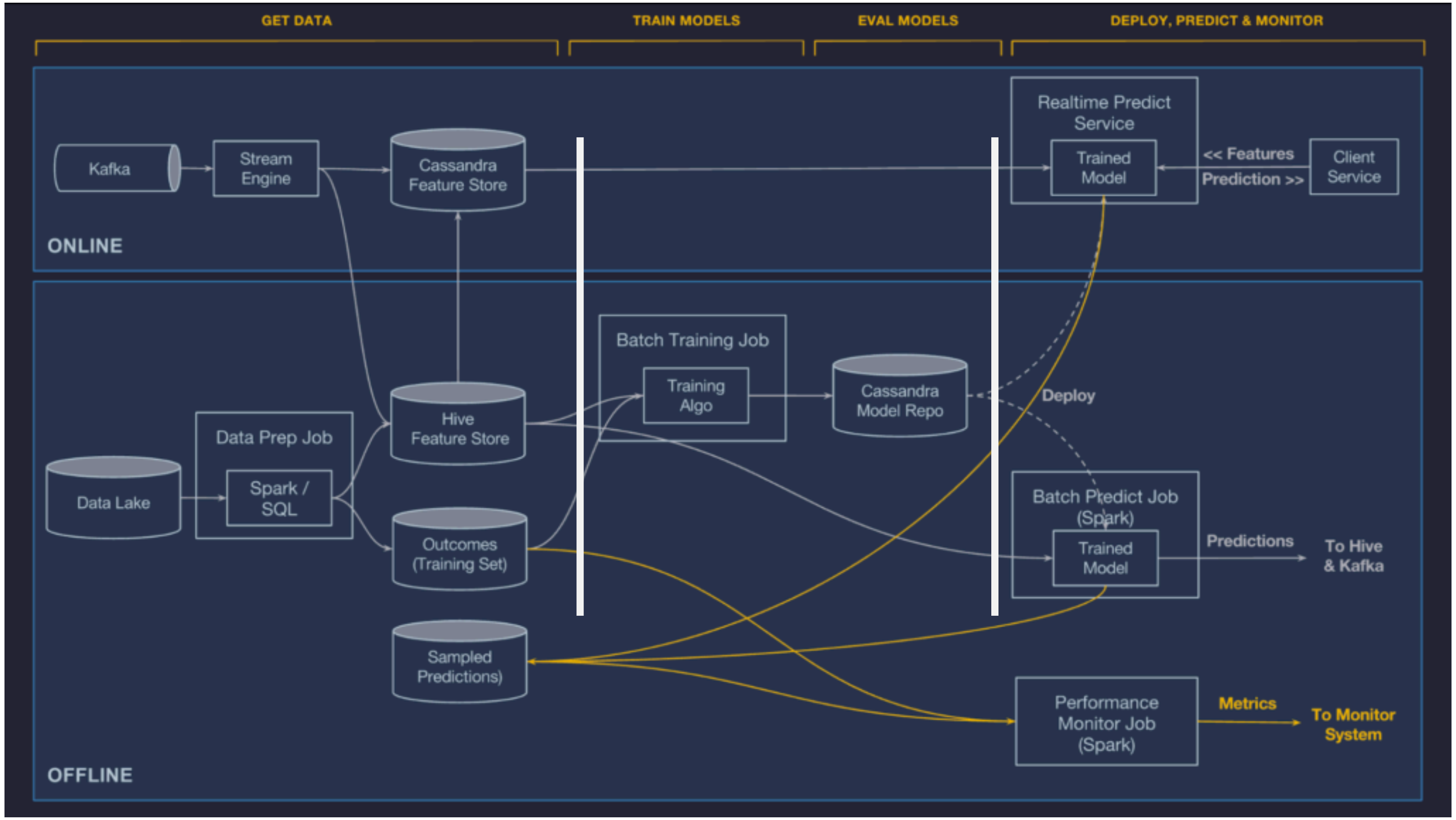
Facebook - FBLearner



Productivity, reusability, scalability, and ease-of-use

Machine Learning Platform Tour

Uber - Michelangelo



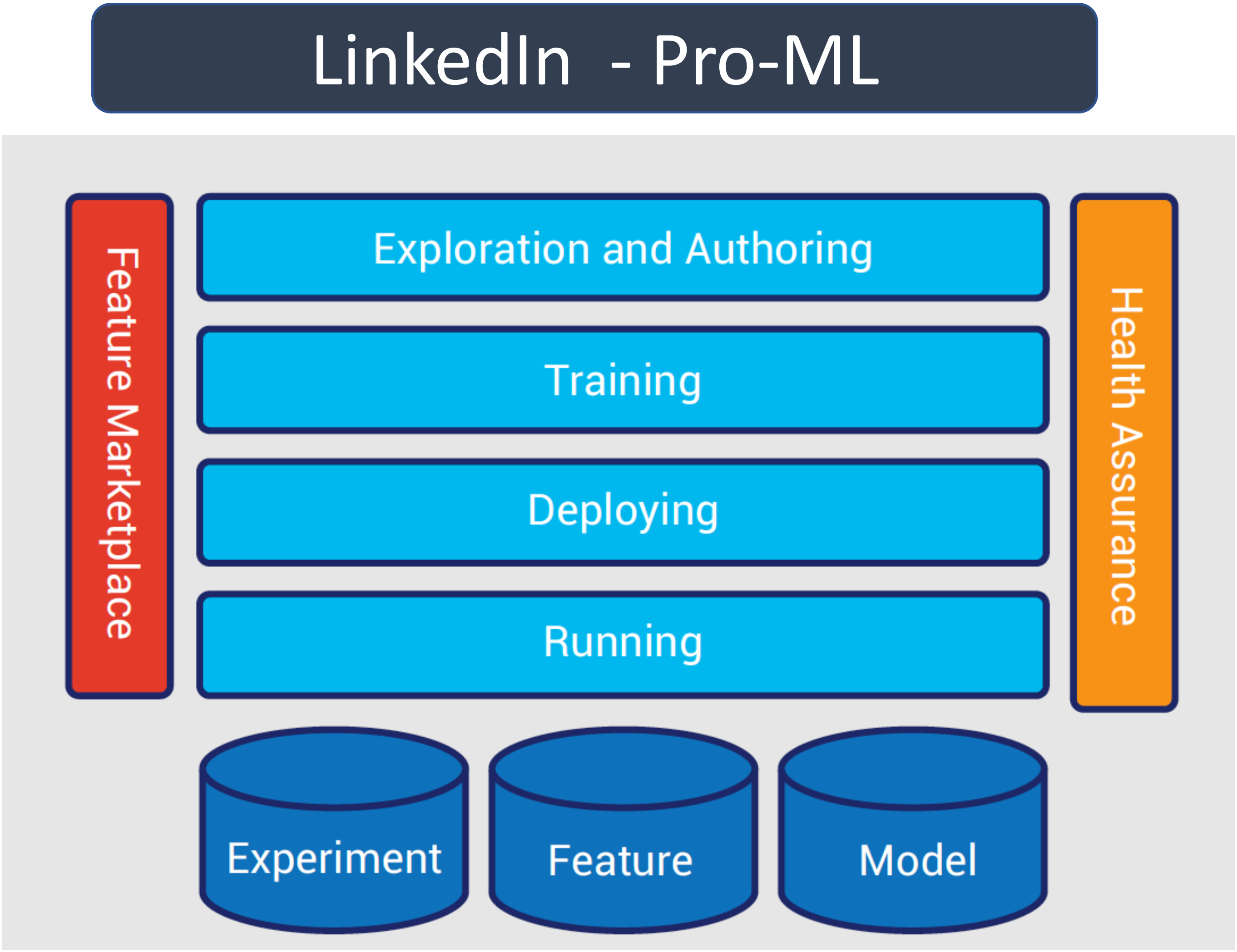
Iterative, tested, and methodical

Democratize & scale AI to make it as easy as requesting a ride

ML as software engineering

Machine Learning Platform Tour

To double effectiveness of ML engineers



Use Cases

- Feeds
- Search
- Recommendation
- Advertisement
- Fraud

https://towardsdatascience.com/introducing-pro-ml-68f37574e1f4

Machine Learning Platform Tour

Cloud Based Machine Learning Platforms

Machine Learning as a Service - MLaaS

AWS SageMaker

Azure ML

Google Cloud
Machine Learning Engine

Major pain points associated with a machine learning project dramatically change as the scale of the project increases.

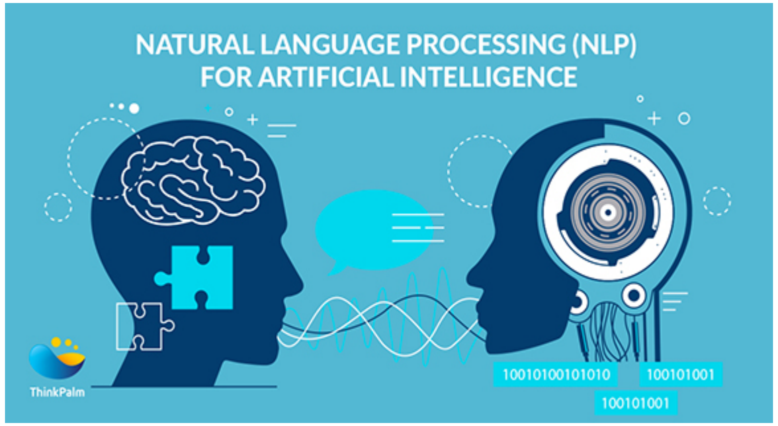
Machine Learning Platform Tour

Pre-trained Machine Learning Models

Vision

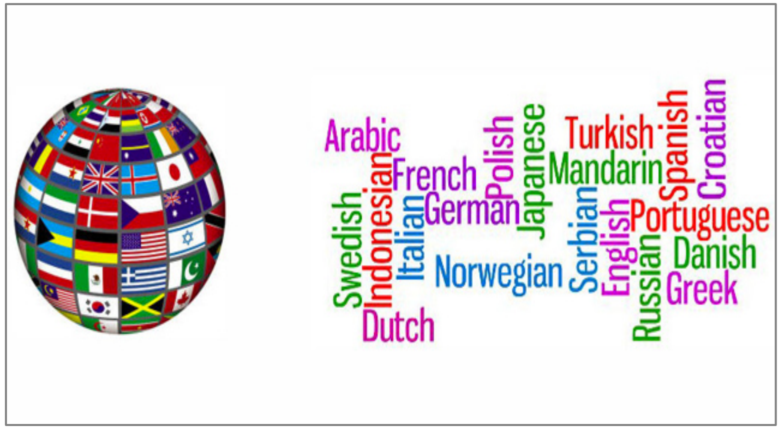
Text within image, facial expressions

Speech



Chat bots, disease predictions, fake news

Language



Translation, language detection

Machine Learning Platform Tour

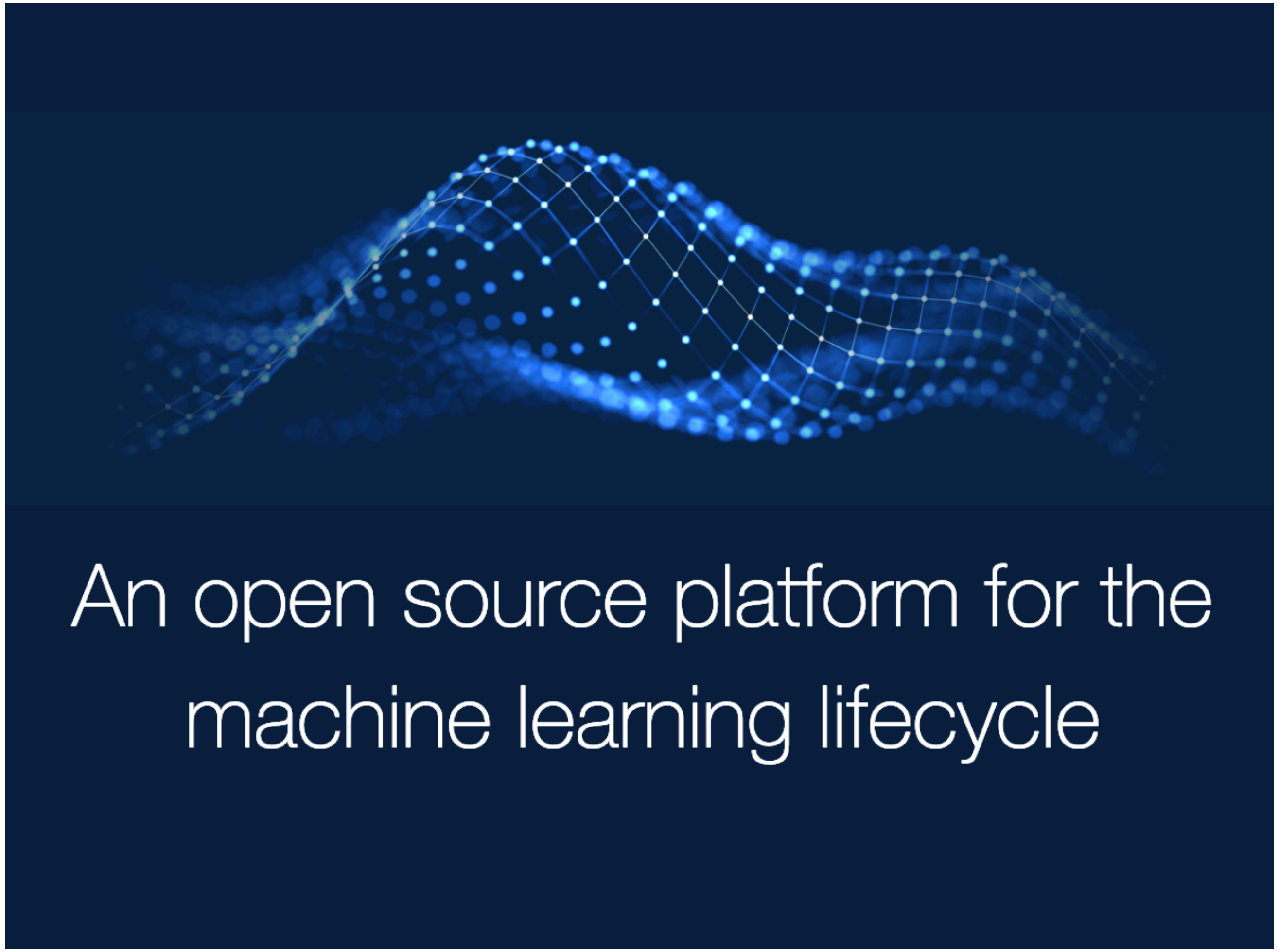
Cloud MLaaS

AI Services
(Computer vision, object recognition, NLP)

ML Services
(ML IDE, experimentation, model training, management & monitoring)

ML Frameworks & Compute Infrastructure
(Tensorflow, Pytorch, Caffe, GPUs, Kubernetes, prediction infrastructure)

MLflow – Open ML Platform



Principles

- Open
- Ease of use
- Extensible
- Scalable

Manage the ML lifecycle, including experimentation, reproducibility and deployment

MLflow – Open ML Platform

mlflow

Tracking

Record and query experiments (code, configs, data,result)

Project

Packaging format for reproducible runs on various platforms

Model

Model format that standardize deployment

Model Registry

Model lifecycle management

MLflow – Tracking

mlflow

Tracking

Record and query experiments: code, configs, results, …etc

Track and analyze experiments

MLflow – Tracking

Tracking Experiments

11	**Dataset**	**Split (trian/dev/test)**	0.7/0.2/0.1	0.7/0.2/0.1	0.7/0.2/0	[illegible]	[illegible]
12		**Class ratio (train/dev/test)**	0.42/0.42/0.42	0.42/0.42/0.42	0.42/0.4	[illegible]	[illegible]
13		**train/dev/test size**	4871/1392/696	4871/1392/696	5315/15	[illegible]	[illegible]
14	**Training hyperparameters**	**Learning rate**	1.00E-05	1.00E-05	[illegible]	[illegible]	[illegible]
15		**epoch**	3	2	[illegible]	[illegible]	[illegible]
16		**batch size**	32	32	[illegible]	[illegible]	[illegible]
17	**Results**	**accuracy**	0.88304595	0.8650862069	0.	[illegible]	[illegible]
18		**f1**	0.82495437	0.8108753316	0.8	[illegible]	[illegible]
19		**precision**	0.878865	0.7848381601	0.8	[illegible]	[illegible]
20		**recall**	0.7780239	0.8389705882	0.807	[illegible]	[illegible]
21		**tp**	1398	1402	14	1334	1130
22		**tn**	1692	1663	1707	1543	1504
23		**fp**	1113	1142	1161	1108	1148
24		**fn**	1189	1185	1190	1154	1357
25		**loss**	0.59637538	0.594134	0.594134	0.6037084	0.594134
26	**Test results**	**accuracy**	0.90747	0.90747	0.88026	0.88314	0.75847
27		**f1**	0.85636	0.85636	0.83108	0.83469	0.5915
28		**precision**	0.90934	0.90934	0.86689	0.87027	0.77626
29		**recall**	0.8099	0.8099	0.79846	0.80226	0.48604
30							

"ML experimentation is like the wild west. Ad-hoc tools and processes because of a lack of standardized tooling. Forget reproducibility, it is difficult to track experiments and results."

https://towardsdatascience.com/tracking-ml-experiments-using-mlflow-7910197091bb

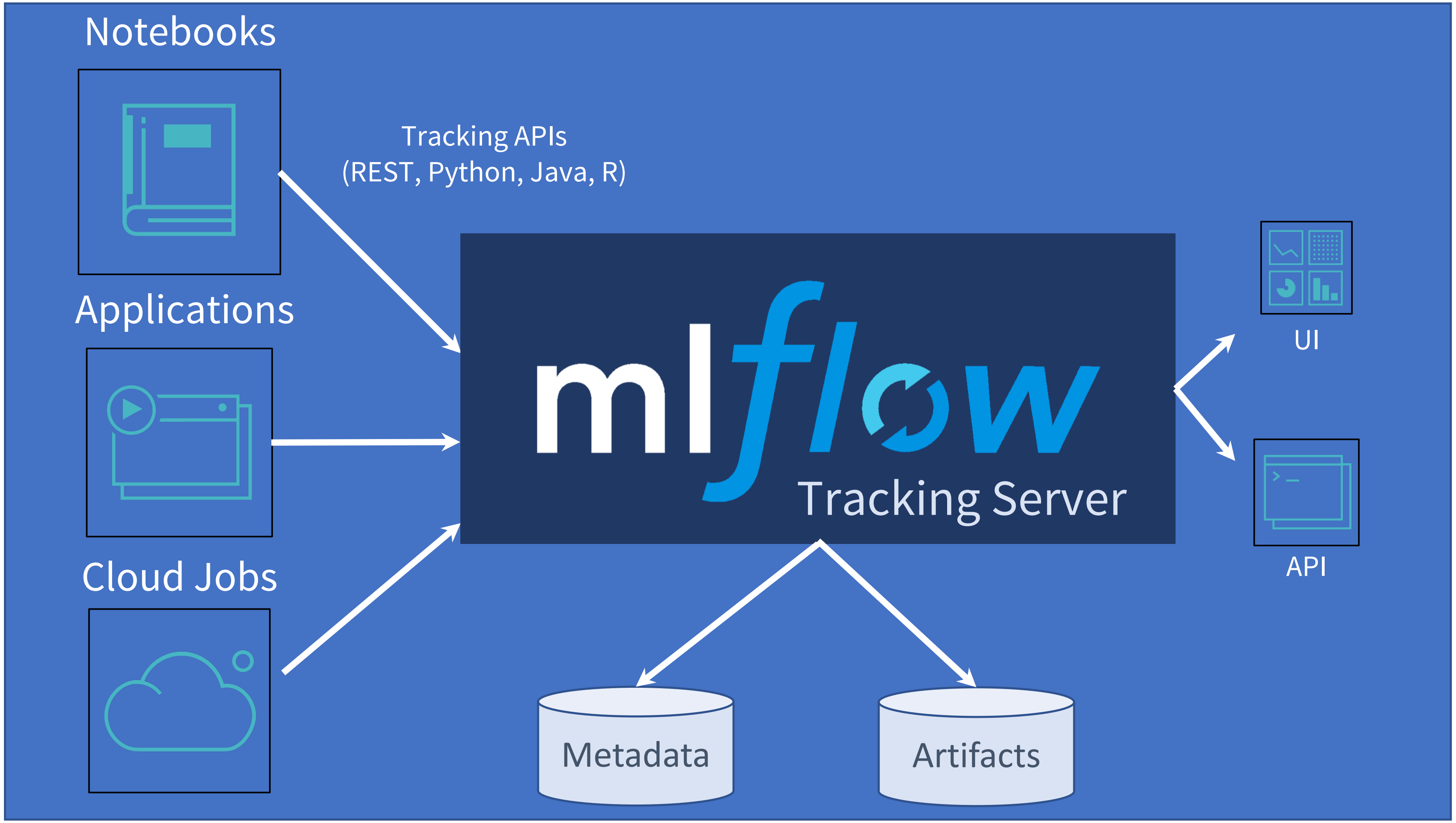
MLflow – Tracking

Tracking

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

Key Concepts

- K-V parameters
- Metrics
- Artifacts
- Source code
- Version (git)
- Tags & notes



MLflow – Tracking

Python, R, Java, REST

Tracking

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

```
import mlflow

 # log model's tuning parameters

with mlflow.start_run():
  mlflow.log_param("layers", layers)
  mlflow.log_param("alpha", alpha)

  # log model's metrics
  mlflow.log_metric("mse", model.mse())
  mlflow.log_artifact("plot", model.plot(test_df))
```

Recently added: `mlflow.keras.autolog()`

MLflow – Open ML Platform

- UI Demo
  - Show MLFlow parameter and metric logging
  - Visualize the experiments
  - `mlflow ui`
    - http://127.0.0.1:7000/#/

MLflow – Projects

mlflow

Projects

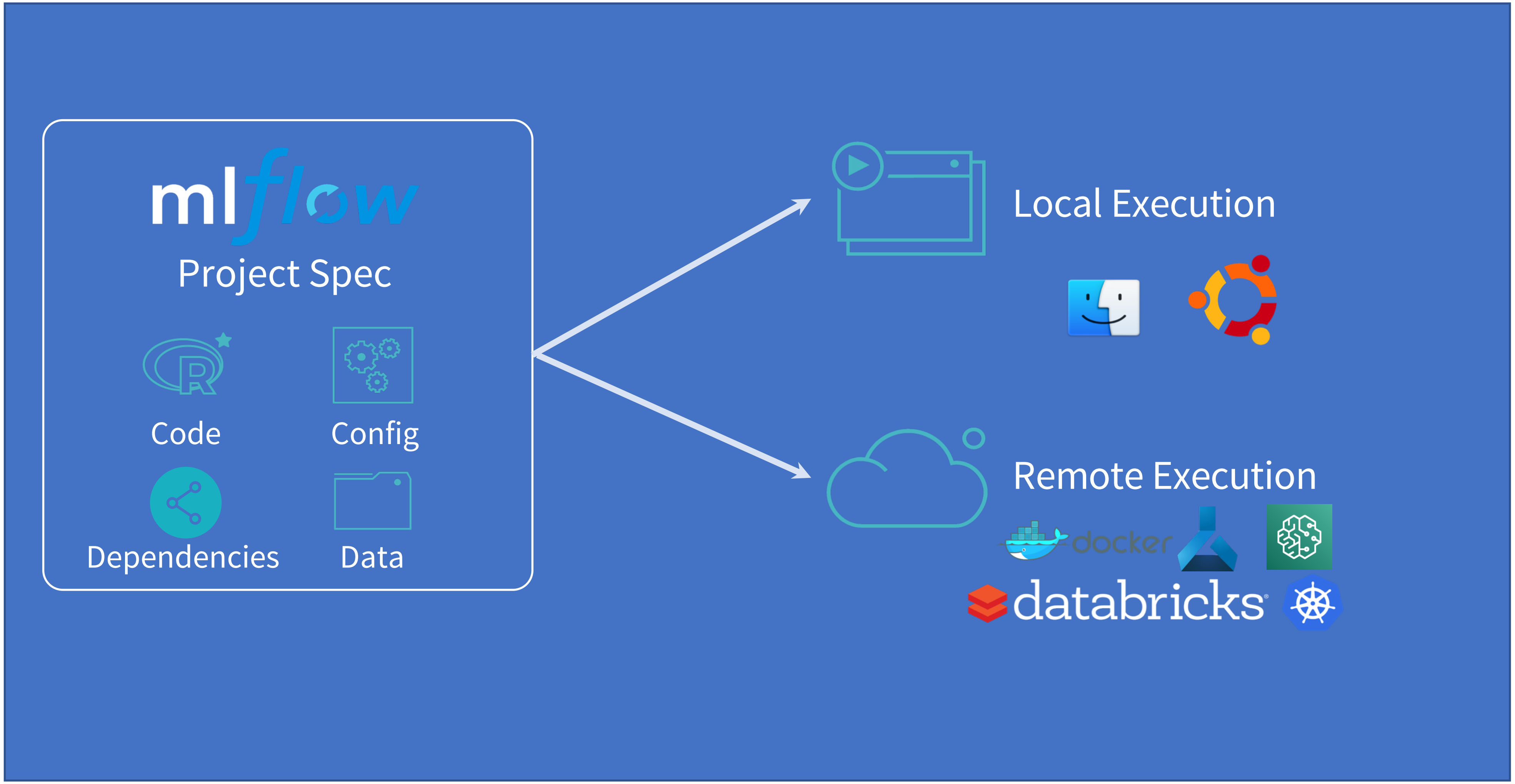
Packaging format
for reproducible runs on any
platform

Reproducibility via self-contained ML project specification

MLflow – Open ML Platform

Projects

Packaging format for reproducible runs on any platform



Reproducibility, Sharing, Productionalization

MLflow – Projects

Projects

Packaging format for reproducible runs on any platform

```
my_project/
├── MLproject
│
│
│
│
│
├── conda.yaml
├── main.py
└── model.py
    ...
```

```yaml
conda_env: conda.yaml

entry_points:
  main:
    parameters:
      training_data: path
      lambda: {type: float, default: 0.1}
    command: python main.py {training_data}
             {lambda}
```

```
$ mlflow run <directory> or git://<my_project>
mlflow.run("<directory> or git://<my_project>")
```

MLflow – Projects

- Demo
  - `mlflow run <local directory>`
  - `mlflow run <github>`

MLflow – Models

Models

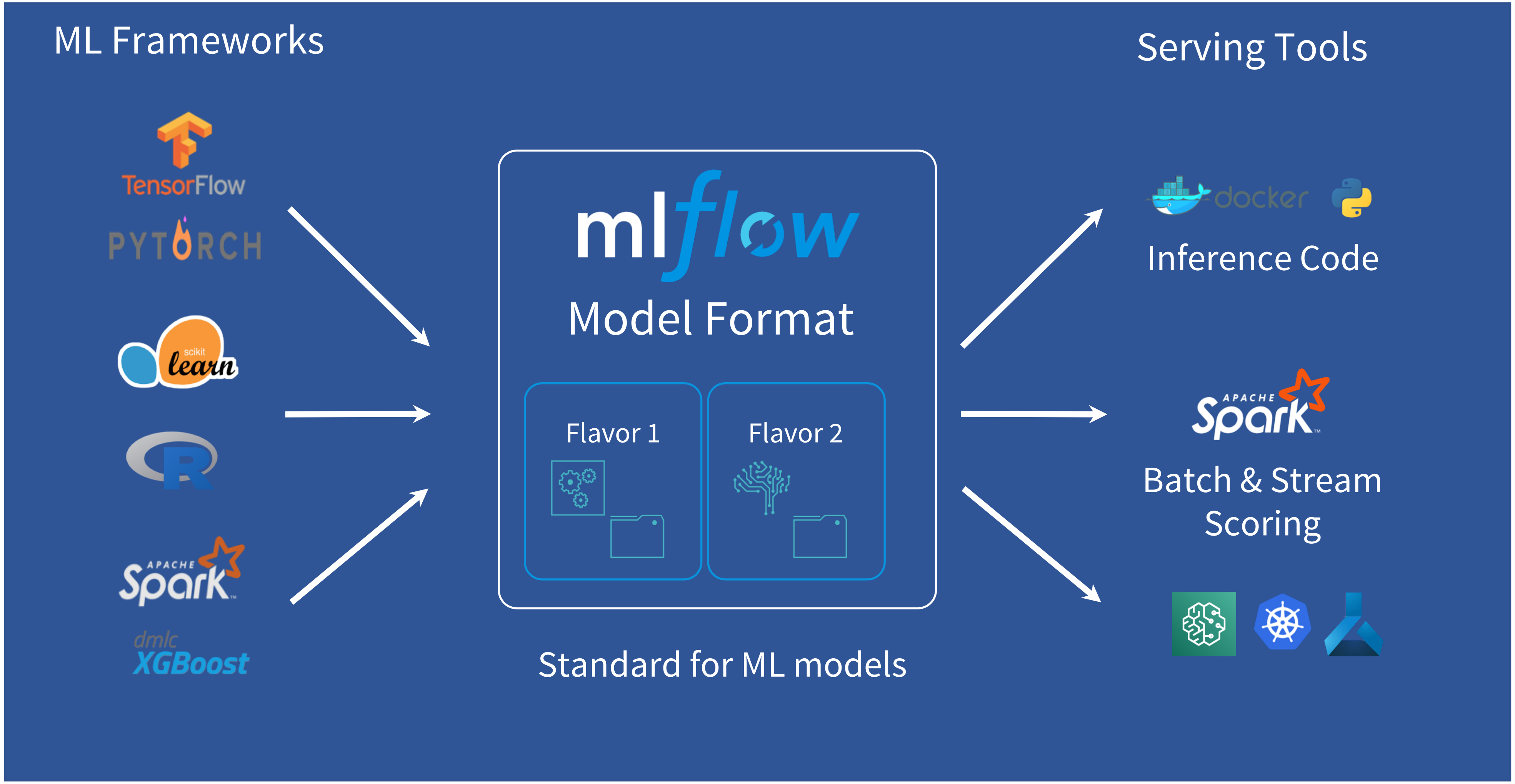
General model format that supports diverse deployment tools

Simplify model deployment

MLflow – Models

Models

General format for sending models to diverse deploy tools



MLflow – Models

Models

General format for sending models to diverse deploy tools

`mlflow.tensorflow.log_model(...)`

```
my_model/
├── MLmodel
│
│
│
│
│
└ estimator/
   ├─ saved_model.pb
   └─ variables/
      ...
```

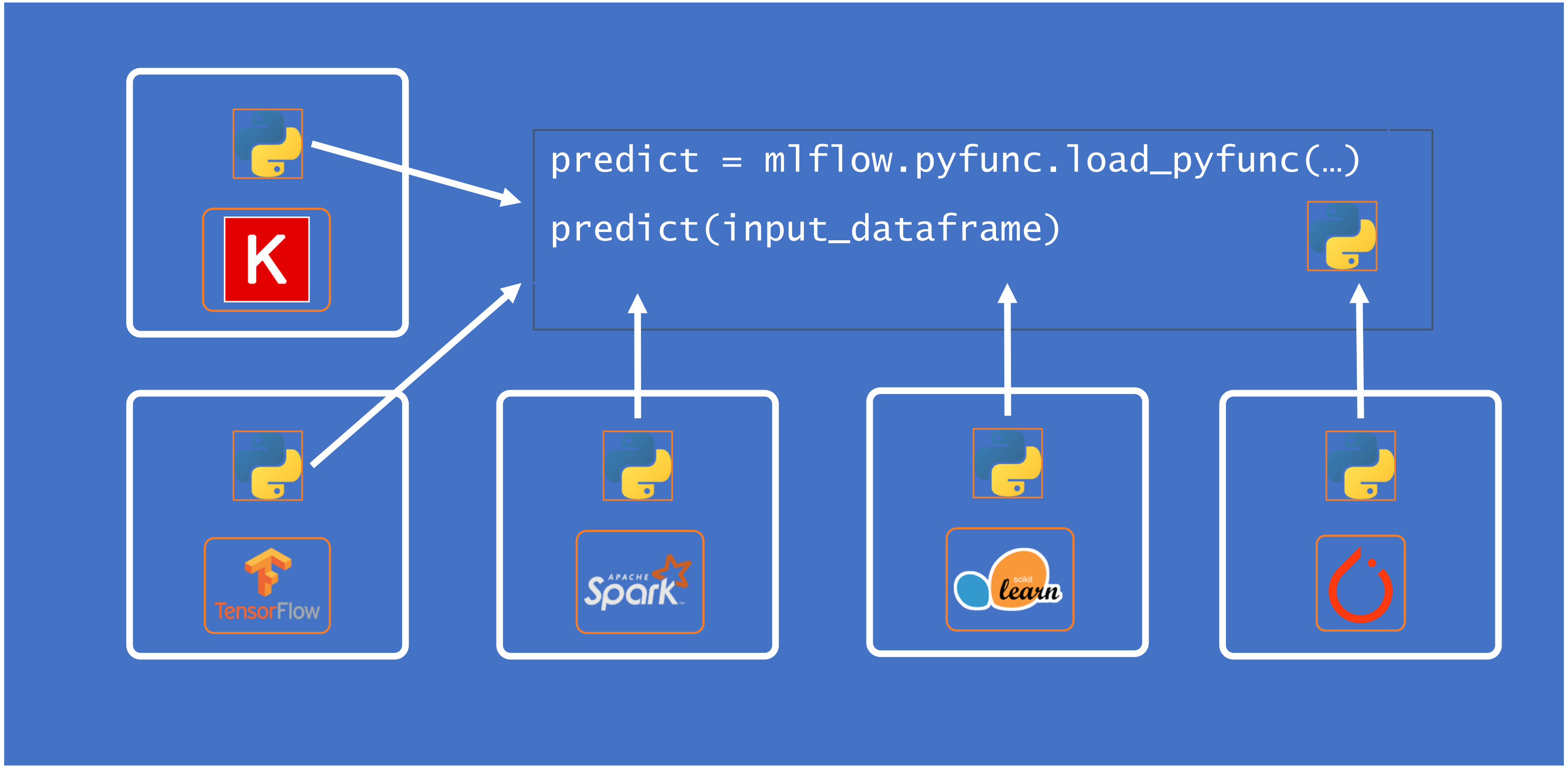
```
run_id: <uuid>
time_created: 2019-06-20T08:11
flavors:
  tensorflow:
    saved_model_dir: estimator
    signature_def_key: predict
  python_function:
    loader_module:
            mlflow.tensorflow
```

MLflow – Models

Models

General format for sending models to diverse deploy tools

Model Flavors



Built-In Flavors

- Python Function
- R,H_2O
- Keras
- Mleap
- PyTorch
- Skiki-learn
- Spark Mllib
- TensorFlow
- ONNX

MLflow – Model Registry

Model Registry

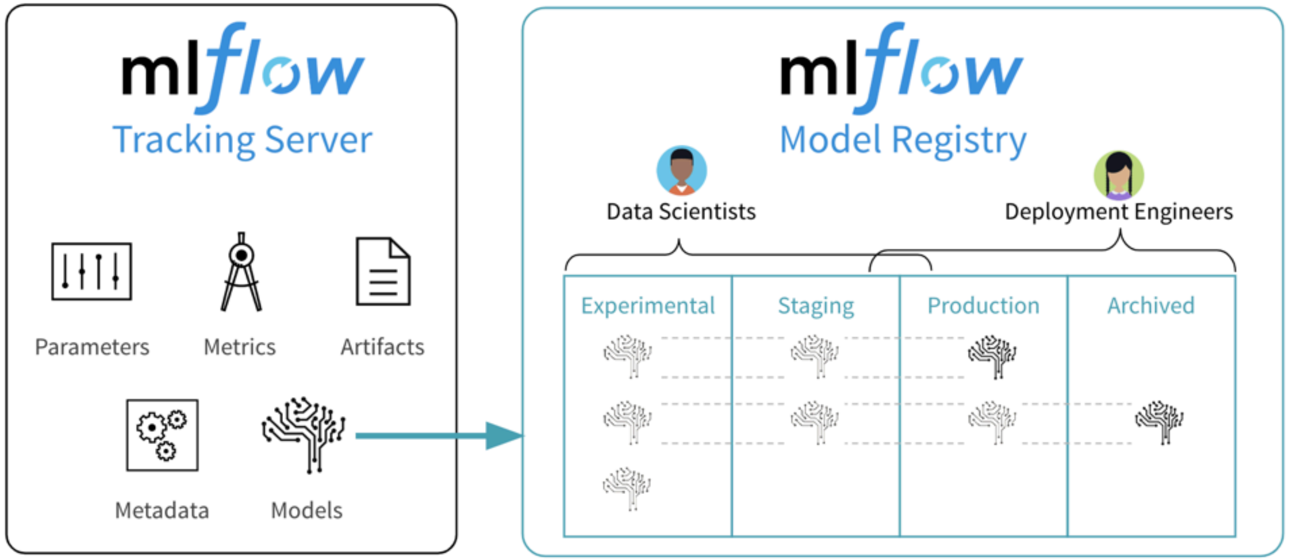
Model lifecycle management

Collaboratively manage the full lifecycle of a model

MLflow – Model Registry

Managing Models Collaboratively

- Model administration and review
  - Sharing, versioning, approval workflow
- Integrate with MLflow tracking
- Centralized activity logs and comments
- Integration with CI/CD



MLflow

- Demo
  - Deploy model to a REST endpoint
    - `mlflow models serve`
  - Simple application to call REST endpoint to perform prediction
    - Flask application

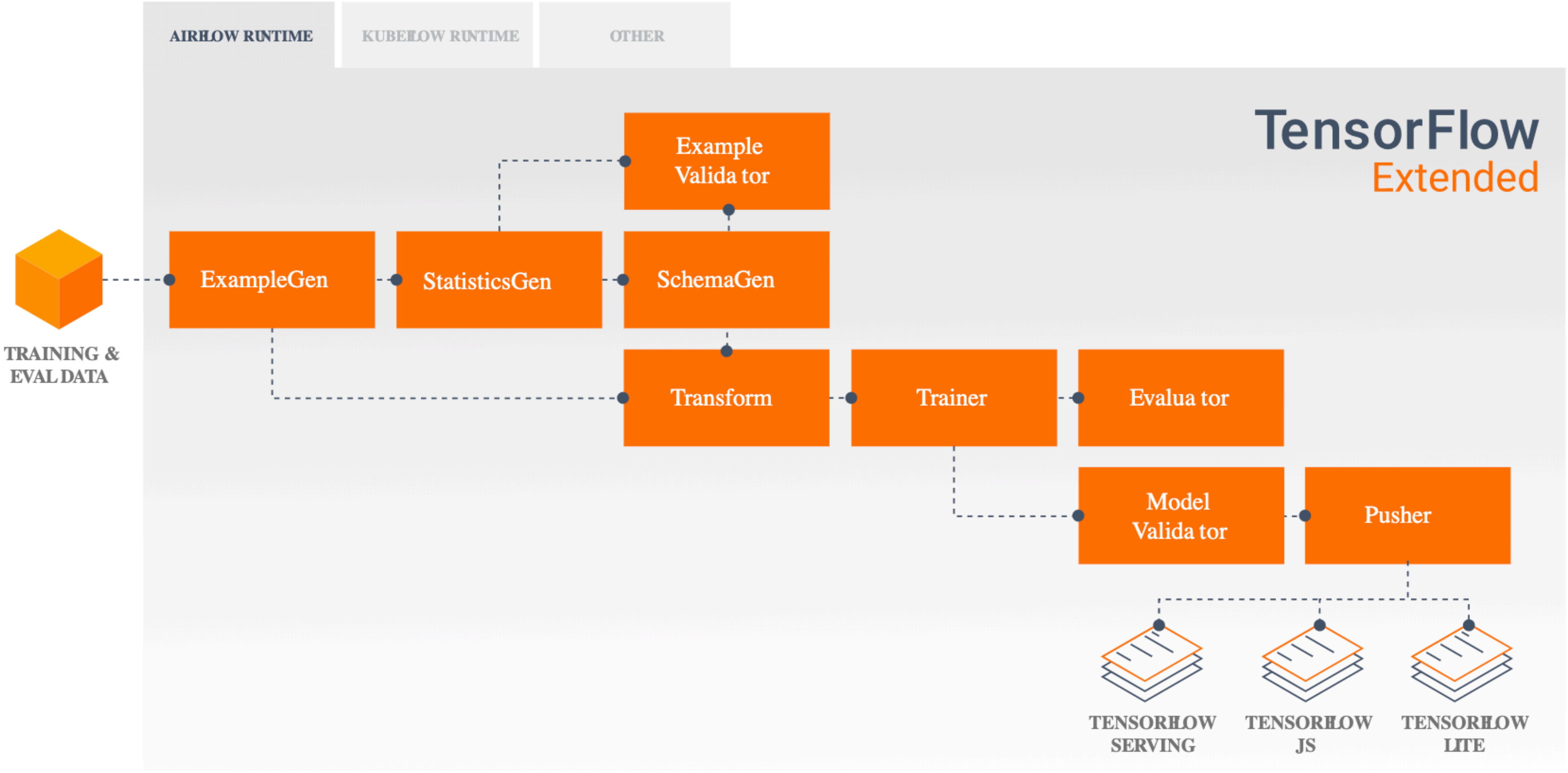
Summary

- Interesting & unique challenges with ML development process
- In-house & cloud-based ML infrastructures
- MLflow – open source ML platform
  - Tracking
  - Project
  - Model
  - Model registry

Q&A

Machine Learning Platform Tour

Google - TFX



ML Development Process Overview

People

Data Engineer | Data Scientist | Software Engineer

Technologies

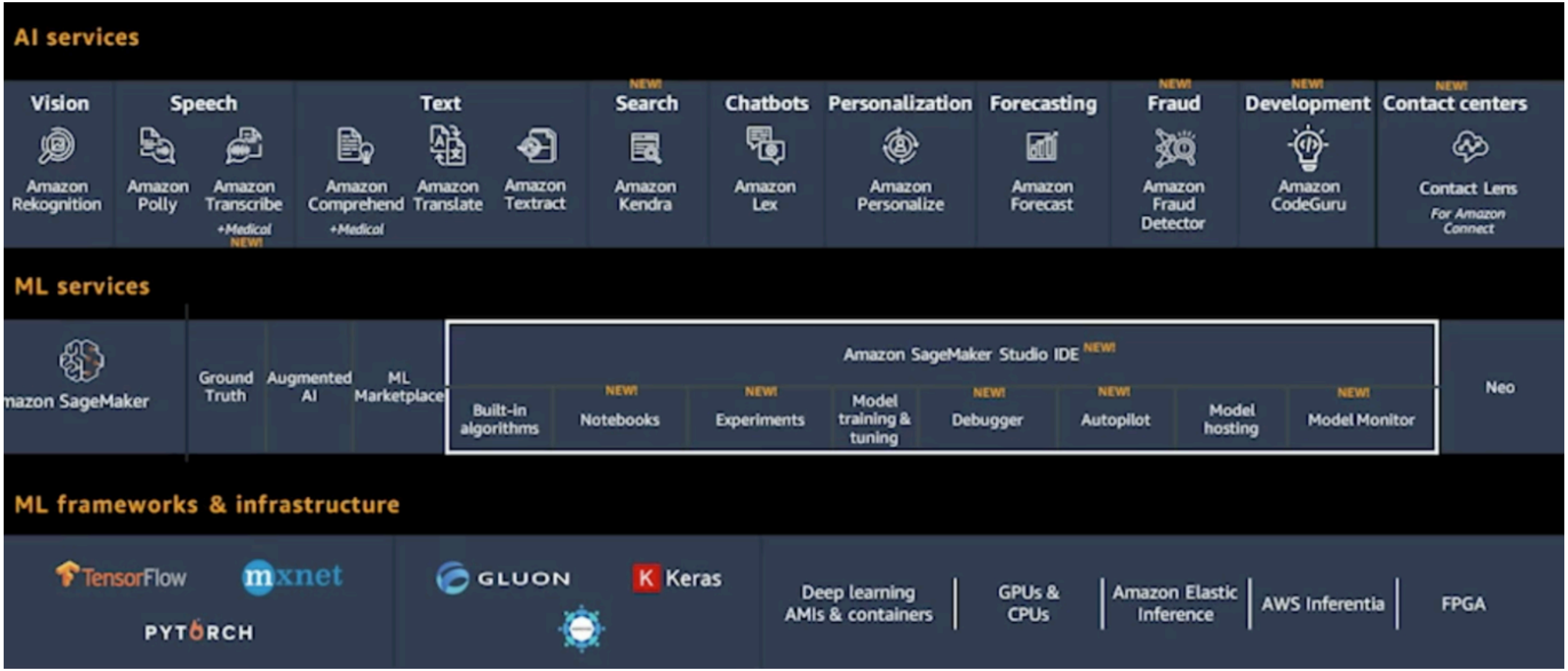
PYTORCH | TensorFlow | APACHE Spark | dmlc XGBoost

Environments

docker | databricks

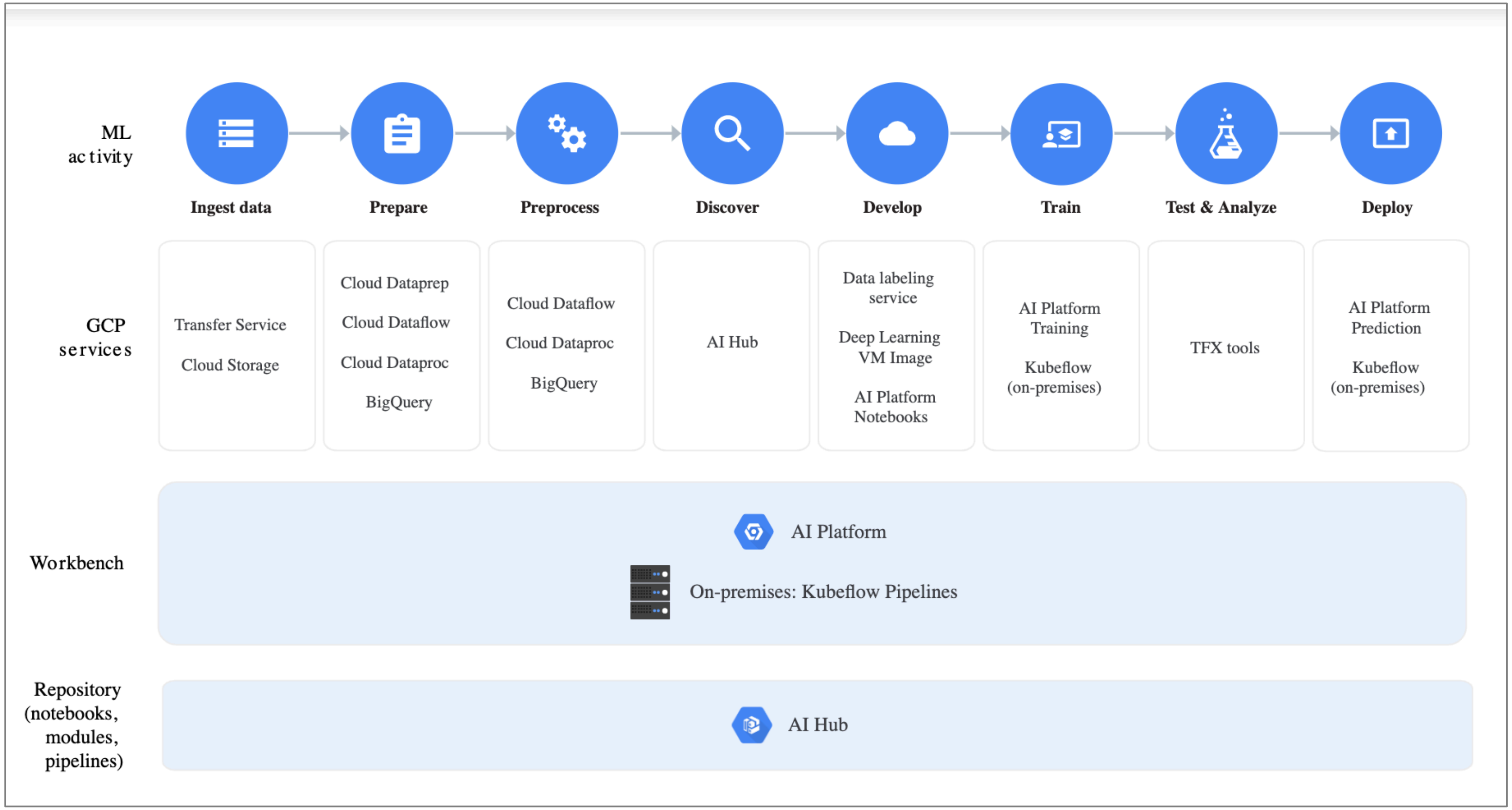
Machine Learning Platform Tour

AWS AI Platform



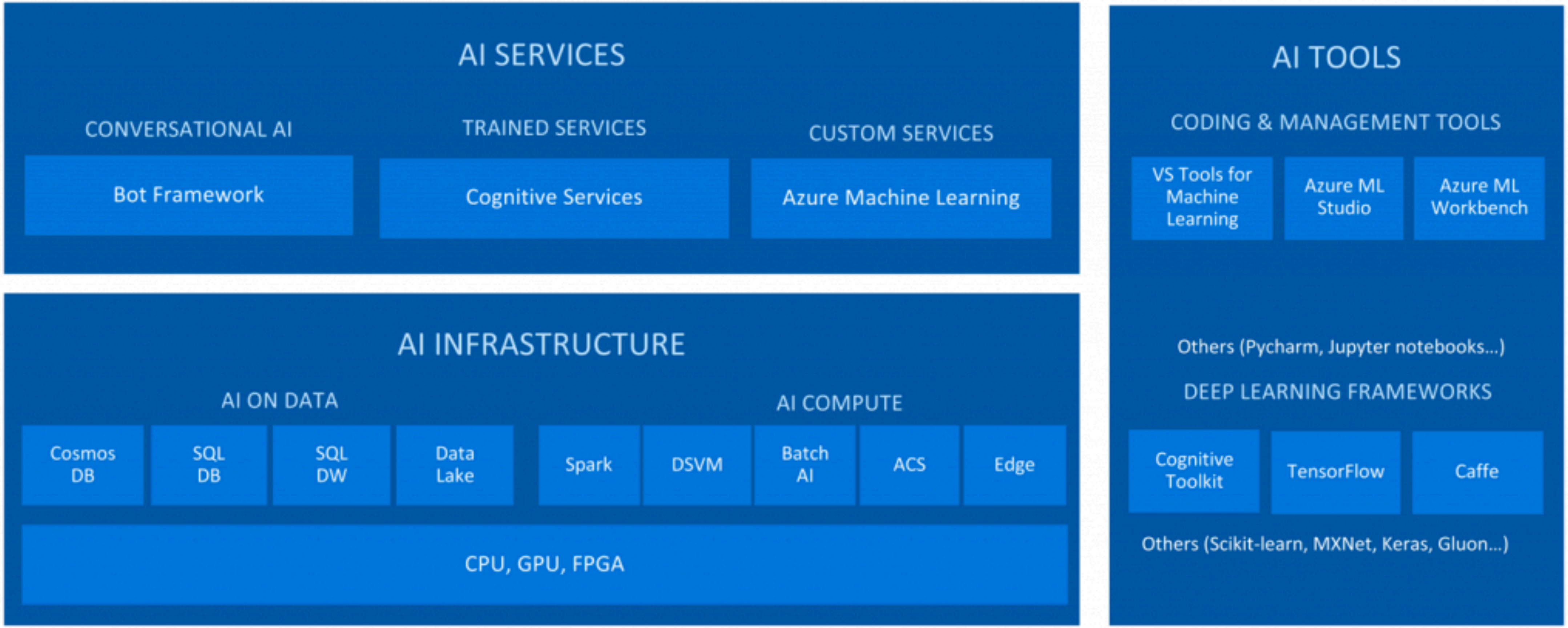
Machine Learning Platform Tour

Google AI Platform



Machine Learning Platform Tour

Azure AI Platform



MLflow – Models

Models

General format for sending models to diverse deploy tools

- Packaging format for ML Model
  - Directory structure
  - Different flavors
    - (keras, pytorch, sklearn, spark, tensorflow)
- Define dependencies
  - Easy reproducibility & deployment
- Model creation utilities
  - To save model in MLFlow format
- Deployment APIs
  - CLI – Python,R,Java

MLflow – Projects

Projects

Packaging format for reproducible runs on any platform

- Packaging format for reproducible for ML runs
  - Training code folder or Github repository
  - Project configuration file in YMAL format
- Define dependencies for <u>reproducibility</u>
  - Using Conda to specifying and managing dependencies
- Execution API for running projects
  - Local or remote execution
  - CLI – Python,R,Java
    - `mlflow run`