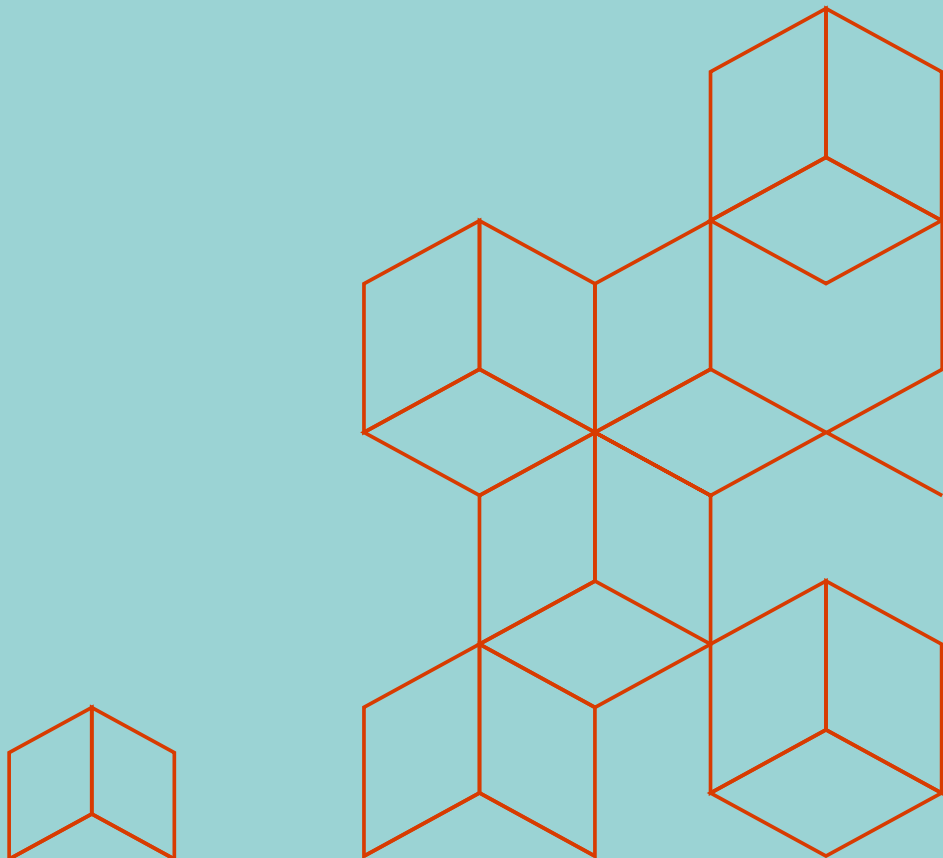
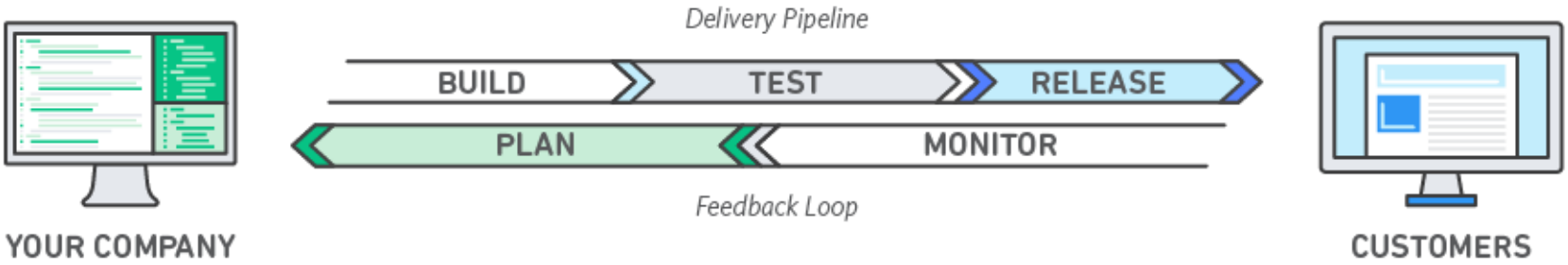
MLOps

Shinyoung Joo

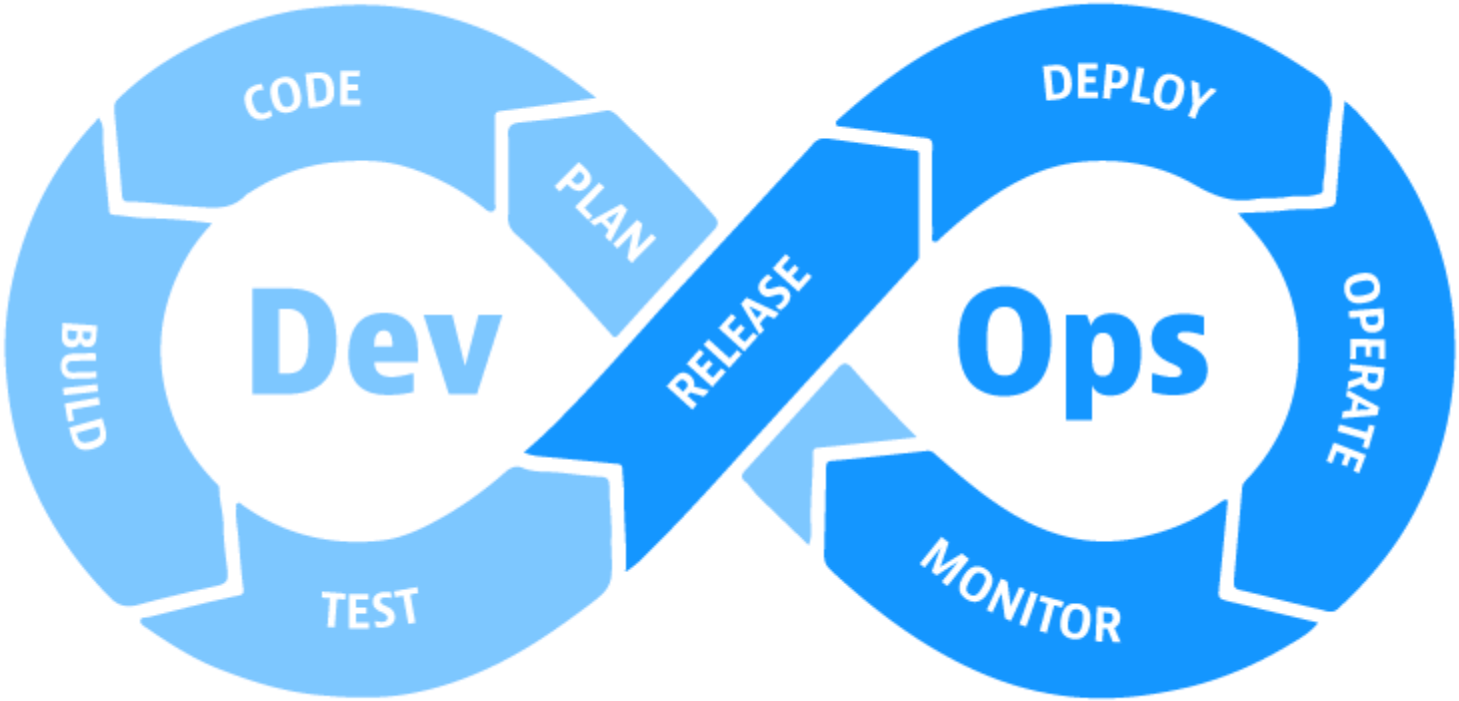
bit1010@live.com

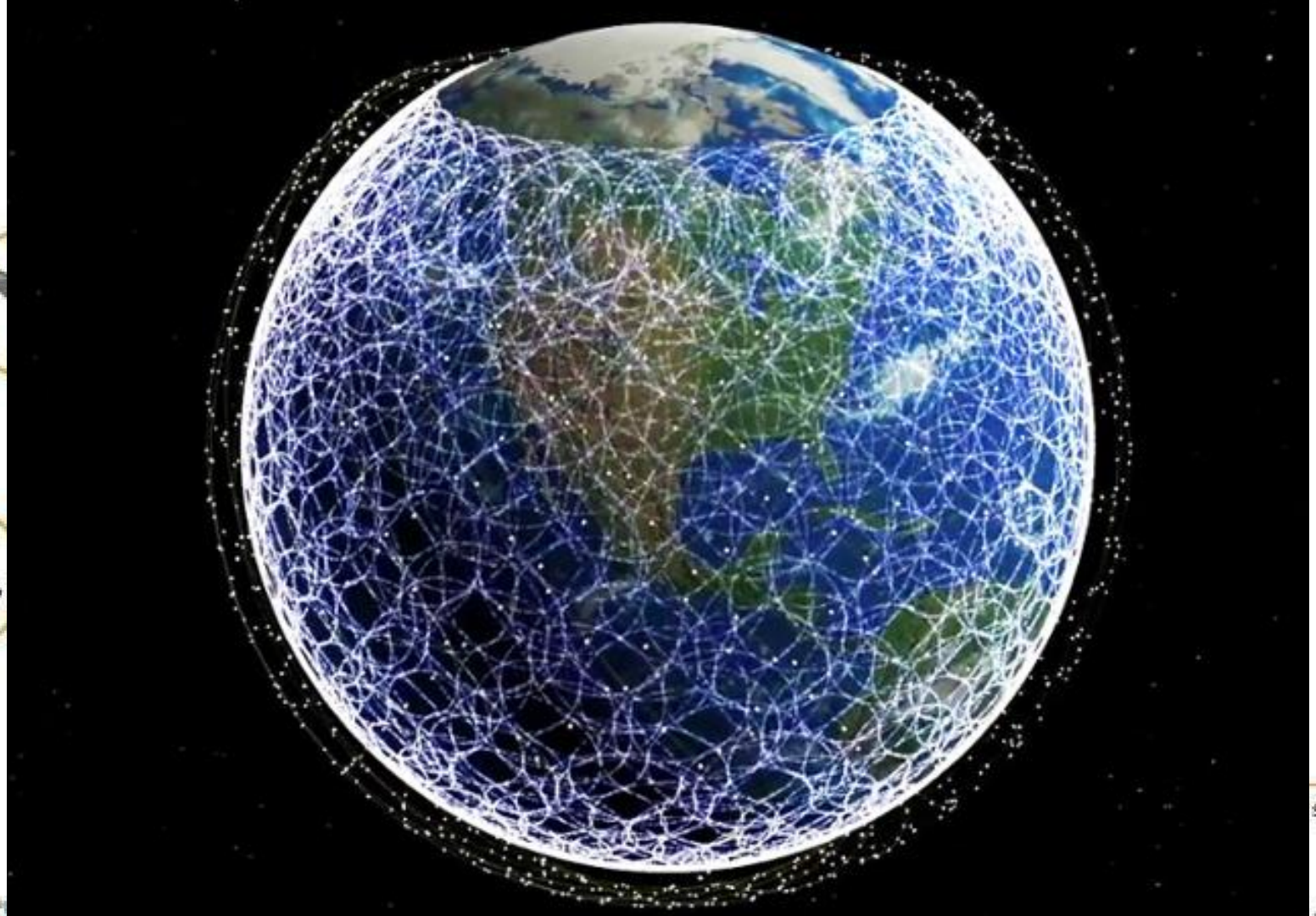
MLOps 란?




Delivery Pipeline
BUILD
TEST
RELEASE
PLAN
MONITOR
Feedback Loop
YOUR COMPANY
CUSTOMERS

DevOps = Development + Operations





Valeo

sensors

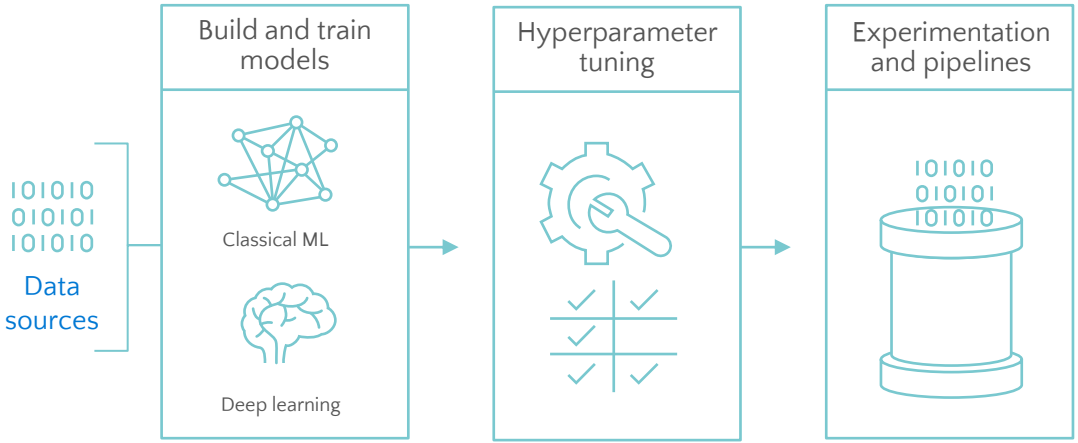

인간을 집에 가둬라!

머신러닝 정말 재미있어요!

우리가 알지 못하는 것들을 예측할 수 있다는 두근거림?

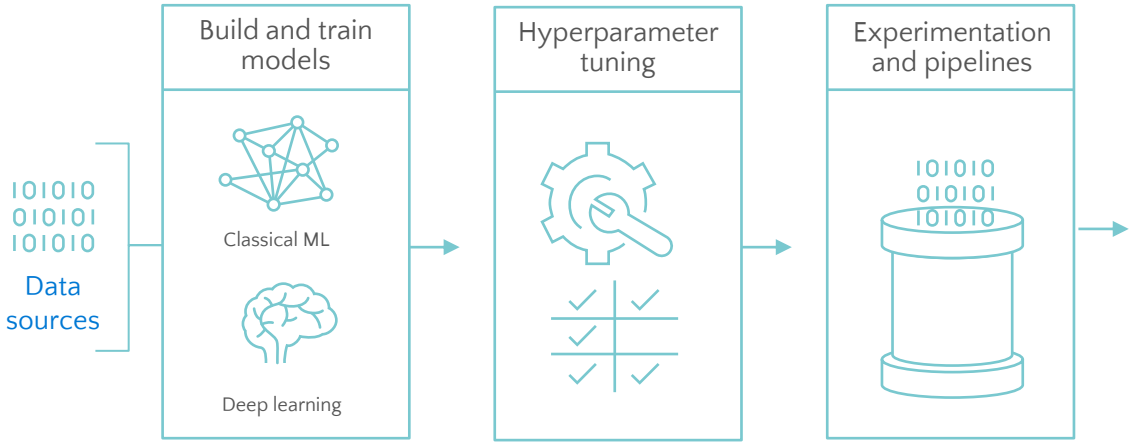
(실제론 모르는데 머리가 아는척함)

Machine Learning Process



자기 컴퓨터
서버 등에서 고정된 데이터로
주로 실험

Machine Learning Process

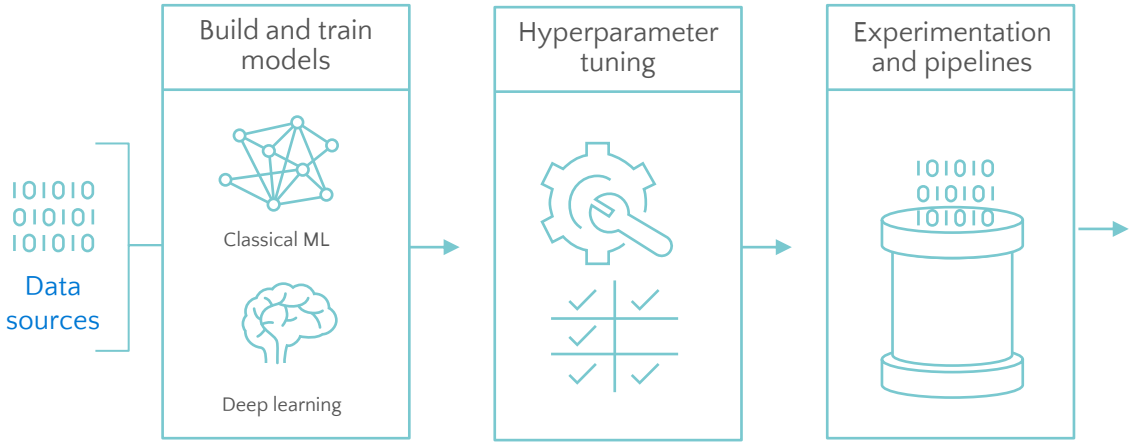


모델의 배포 및 운영 문제 고민

```python
from flask import Flask
app = Flask(__name__)

@app.route('/')
def hello_world():
    return 'Hello, World!'
```

Machine Learning Process



효과적으로 운영하기 위해
컨테이너 고민

Flask

docker

모델을 API 형태 또는 Batch 형태로 배포한 후, 생기는 일

1) (운영팀 제보) 모델의 결과값이 이상한 것 같아요..!

모델을 API 형태 또는 Batch 형태로 배포한 후, 생기는 일

1) (운영팀 제보) 모델의 결과값이 이상한 것 같아요..!

(당연히) 결과값을 어딘가 저장해서 재현 가능

모델을 API 형태 또는 Batch 형태로 배포한 후, 생기는 일

1) (운영팀 제보) 모델의 결과값이 이상한 것 같아요..!

1-1) Input 데이터는 0 ~ 23까지의 데이터가 들어가야 하는데 77이란 데이터가 들어갔네..?
1-2) Float 타입으로 들어가야 하는데 Int로 들어왔네..?

모델을 API 형태 또는 Batch 형태로 배포한 후, 생기는 일

2) 모델 성능 어떤가요? 새롭게 배포한 모델은 어떤가요?
특정 카테고리에서 잘 맞추나요?

모델을 API 형태 또는 Batch 형태로 배포한 후, 생기는 일

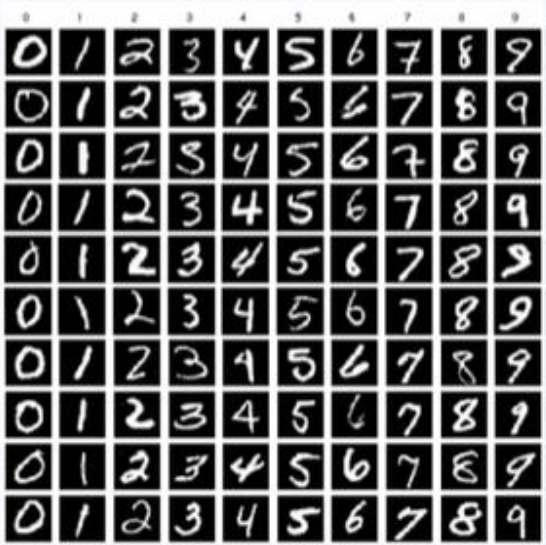
2) 모델 성능 어떤가요? 새롭게 배포한 모델은 어떤가요?
특정 카테고리에서 잘 맞추나요?

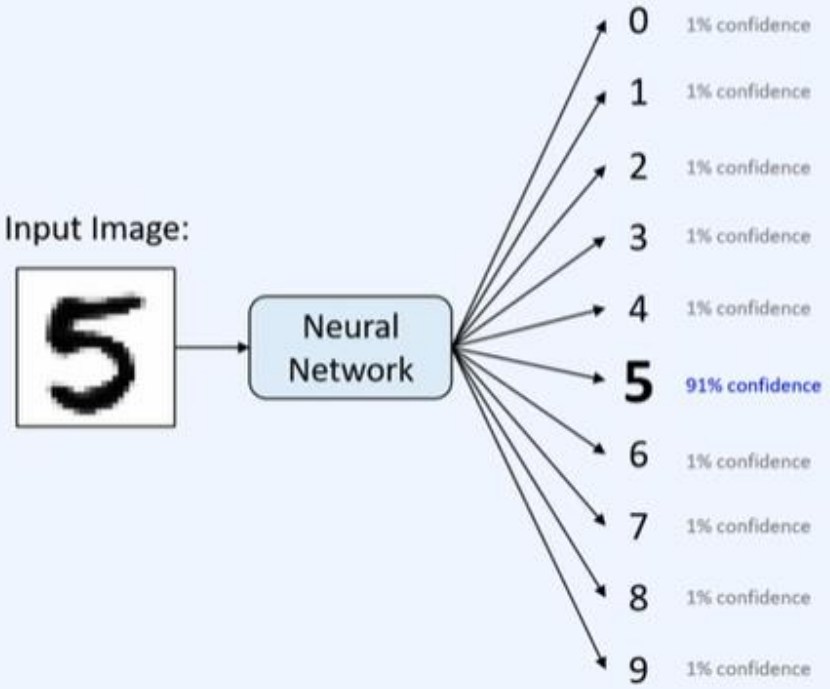
다행히 모델 모니터링 대시보드 미리 만들어서 확인
Research 환경에서 더 좋았던 새로운 모델.
Production 환경에선 저번 모델이 더 좋다..?

휴 이제 모델링에 집중할까 했는데, 새벽 2시네..?

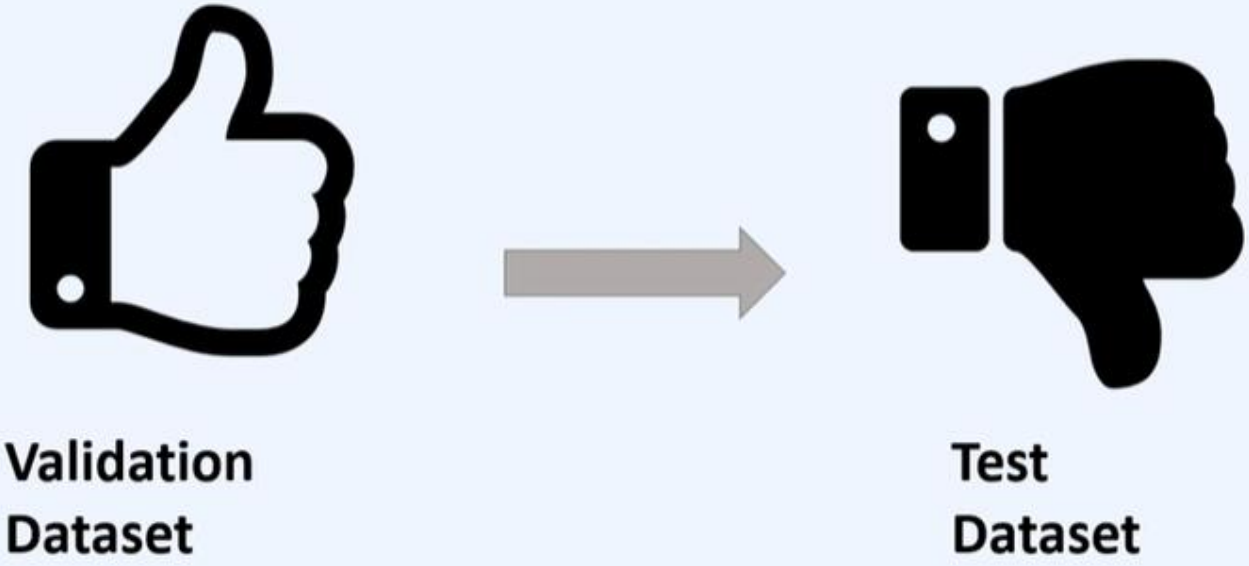
모델링에 집중하고 싶은데 할 일이 너무 많다..

MNIST 데이터 분류 모델 중
성능이 가장 좋았던 모델의 정보가 기억나시나요?





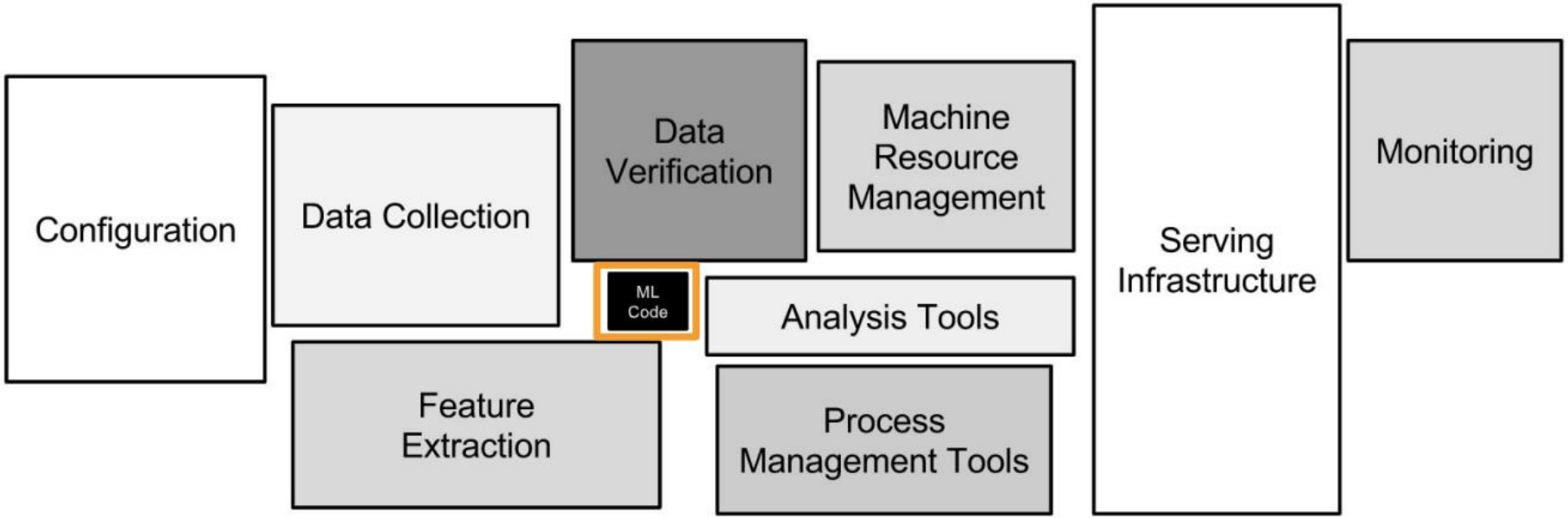
학습 환경과 배포 환경은 같지 않습니다



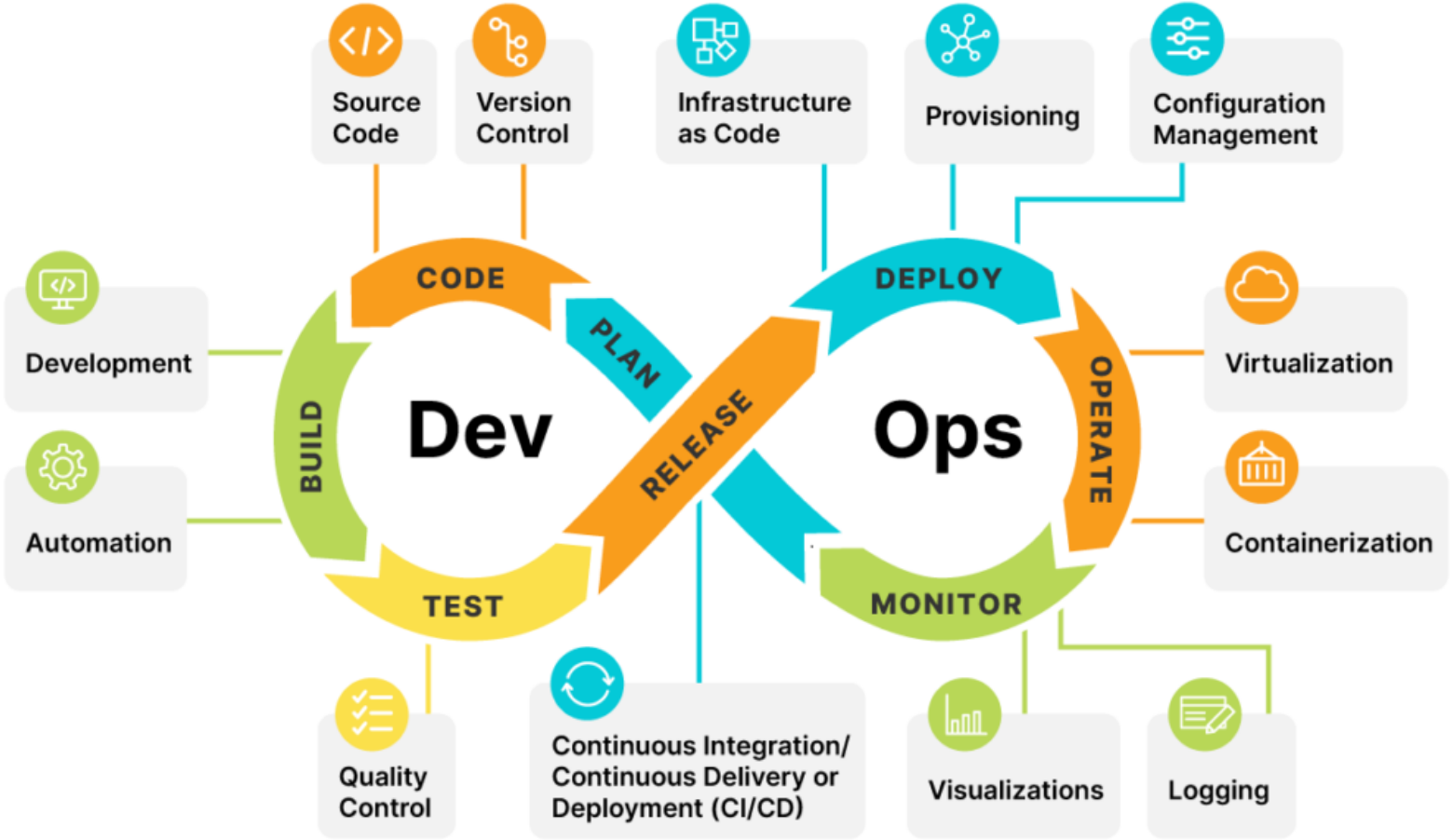

PyTorch

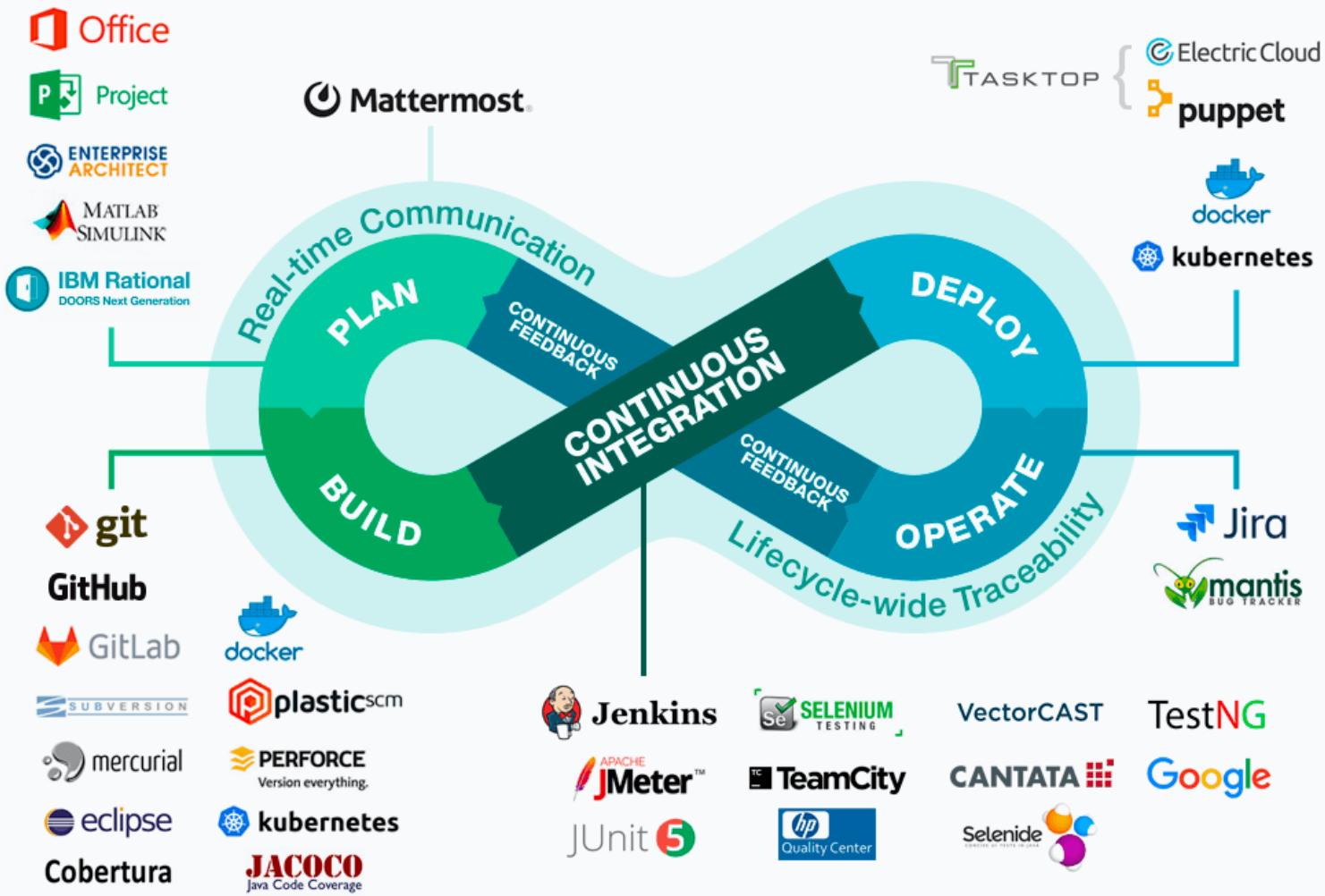

TensorFlow


ubuntu



“Hidden Technical Debt in Machine Learning Systems”


Source Code
Version Control
Infrastructure as Code
Provisioning
Configuration Management
Development
Automation
CODE
PLAN
BUILD
TEST
RELEASE
DEPLOY
OPERATE
MONITOR
Dev
Ops
Virtualization
Containerization
Quality Control
Continuous Integration/ Continuous Delivery or Deployment (CI/CD)
Visualizations
Logging


Office
Project
ENTERPRISE ARCHITECT
MATLAB SIMULINK
IBM Rational DOORS Next Generation
Mattermost
TASKTOP
Electric Cloud
puppet
docker
kubernetes
Real-time Communication
PLAN
CONTINUOUS FEEDBACK
CONTINUOUS INTEGRATION
DEPLOY
BUILD
CONTINUOUS FEEDBACK
OPERATE
Lifecycle-wide Traceability
Jira
mantis BUG TRACKER
git
GitHub
GitLab
SUBVERSION
mercurial
eclipse
Cobertura
docker
plasticscm
PERFORCE Version everything.
kubernetes
JACOCO Java Code Coverage
Jenkins
SELENIUM TESTING
VectorCAST
TestNG
APACHE JMeter™
TeamCity
CANTATA
Google
JUnit 5
hp Quality Center
Selenide

MLOps = 머신러닝 엔지니어링
+ 데이터 엔지니어링
+ 인프라

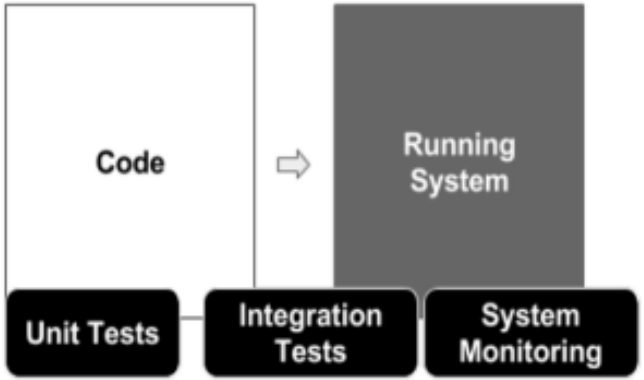
ML + DevOps

머신러닝 모델 개발(ML)과 머신러닝 모델 운영(Ops)에서 사용되는 문제, 반복을 최소화하고 비즈니스 가치를 창출하는 것이 목표

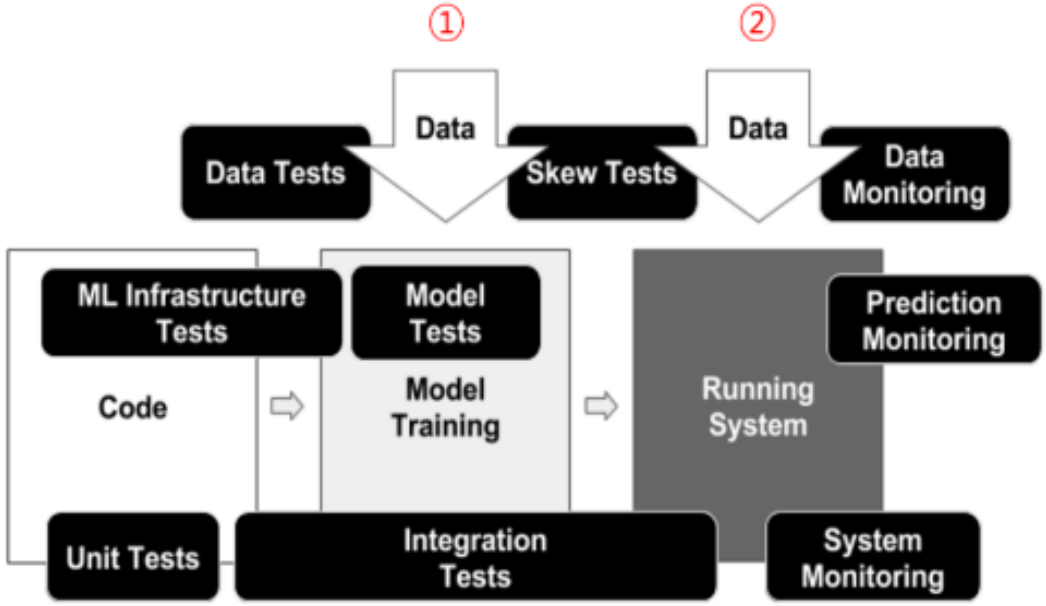
모델링에 집중할 수 있도록 관련된 인프라를 만들고, 자동으로 운영되도록 만드는 일

예 : API 형태로 서버 만들기, 실험 파라미터와 결과 저장하기, 모델 결과 자동화하기, 데이터 Validation 등

ML 기반 시스템의 테스트와 모니터링

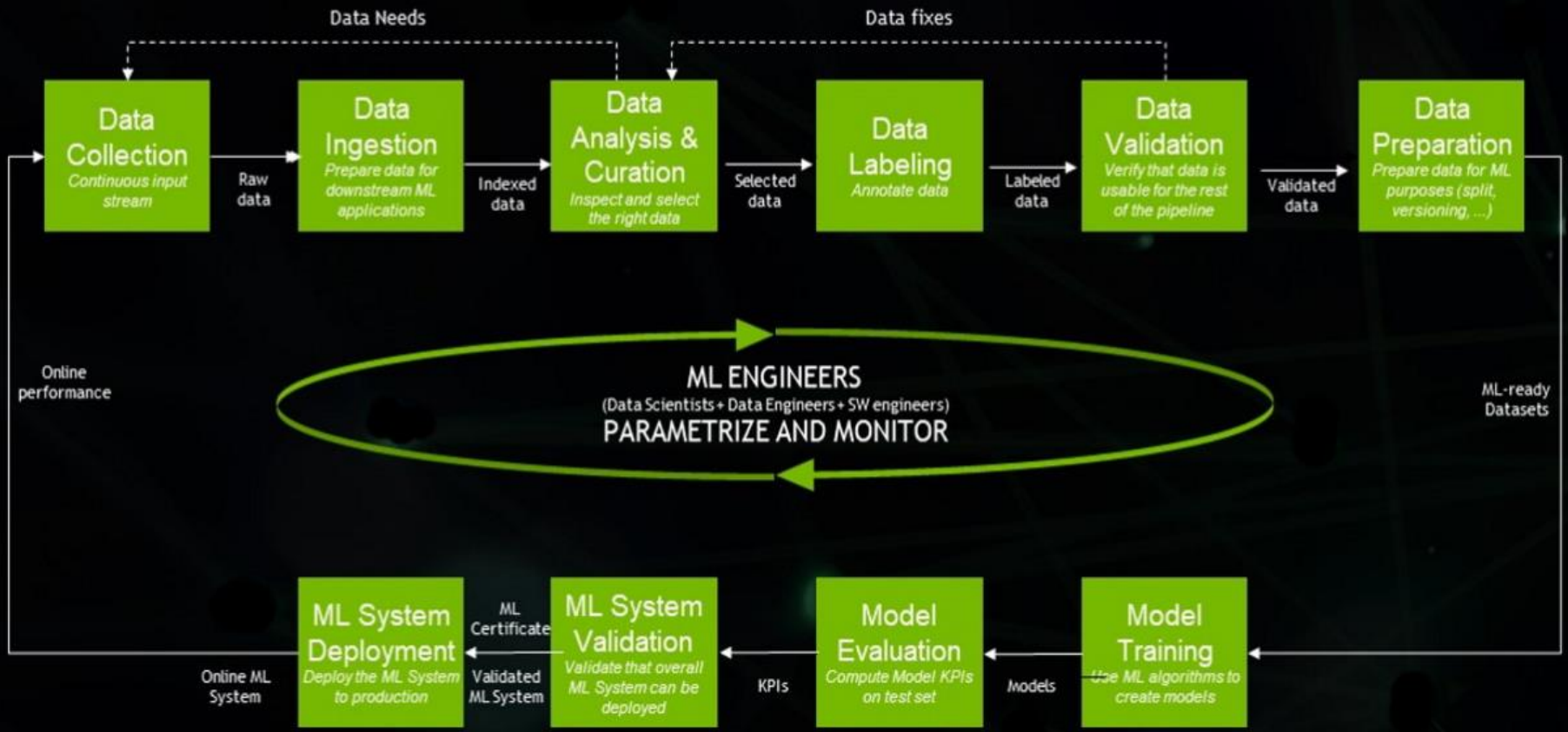


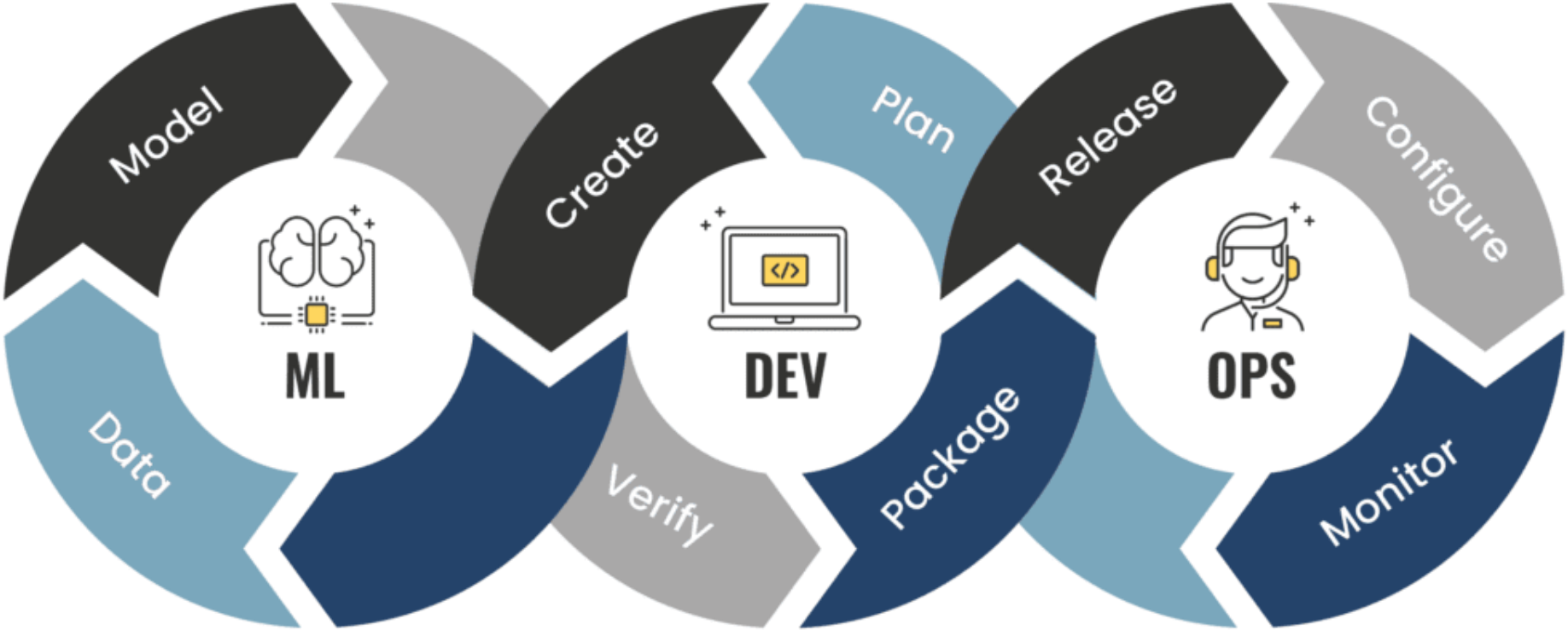
Traditional System Testing and Monitoring



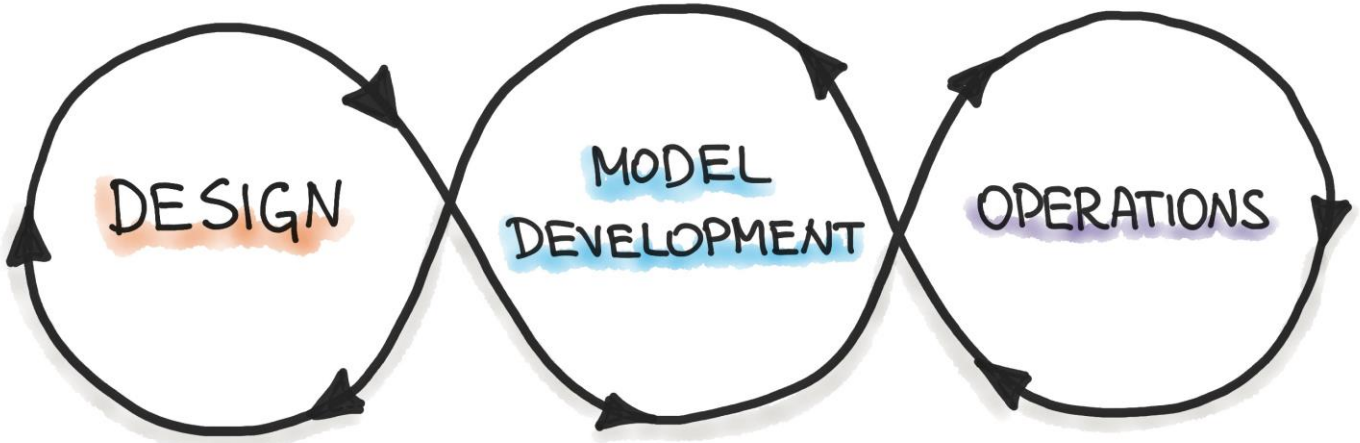
ML-Based System Testing and Monitoring

MLOPS: THE AI LIFECYCLE FOR IT PRODUCTION




Model
Create
Plan
Release
Configure
ML
DEV
OPS
Data
Verify
Package
Monitor

MLOps



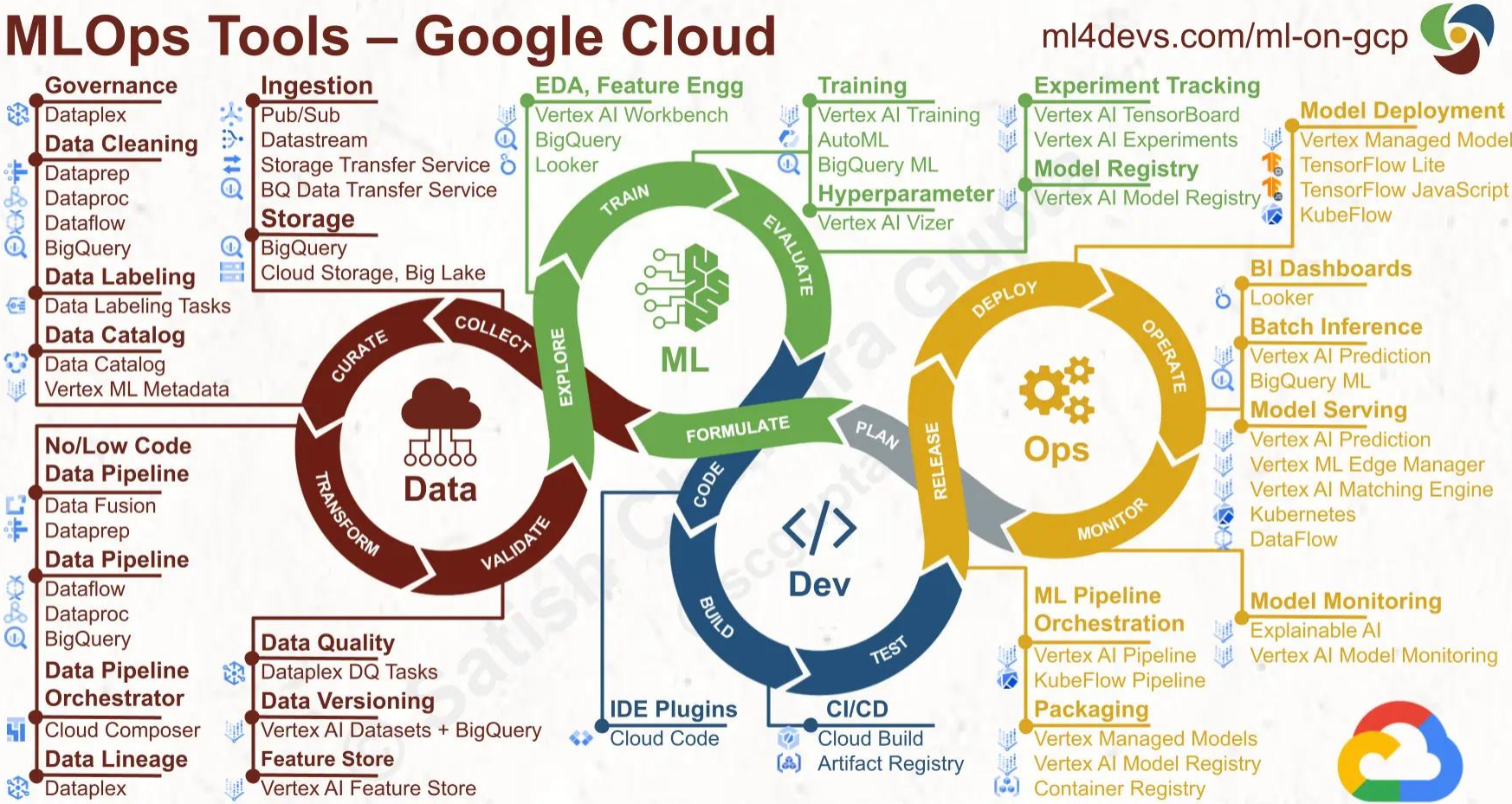
- Requirements Engineering
- ML Use-Cases Priorization
- Data Availability Check

- Data Engineering
- ML Model Engineering
- Model Testing & Validation

- ML Model Deployment
- CI/CD Pipelines
- Monitoring & Triggering

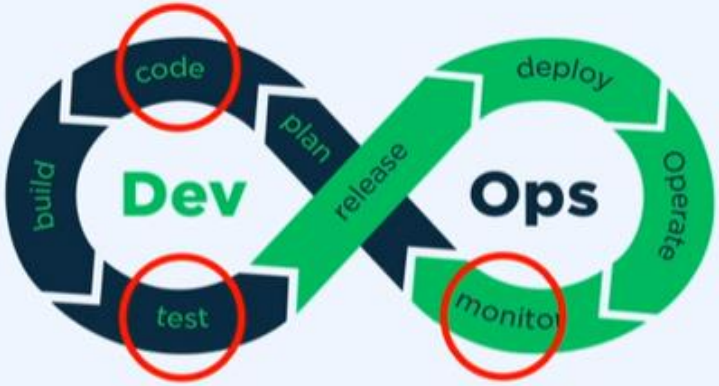
MLOps Tools – Google Cloud

ml4devs.com/ml-on-gcp



scgupta linkedin.com/in/scgupta

ML 프로젝트와 SW 프로젝트의 유사성



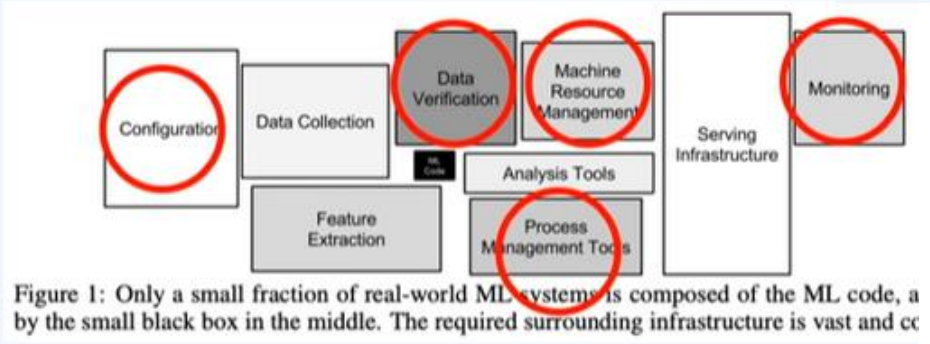


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, a
by the small black box in the middle. The required surrounding infrastructure is vast and co

버전 관리

데이터 버전 관리
모델 버전 관리

테스트 자동화

모델 학습 자동화
모델 성능 평가 자동화

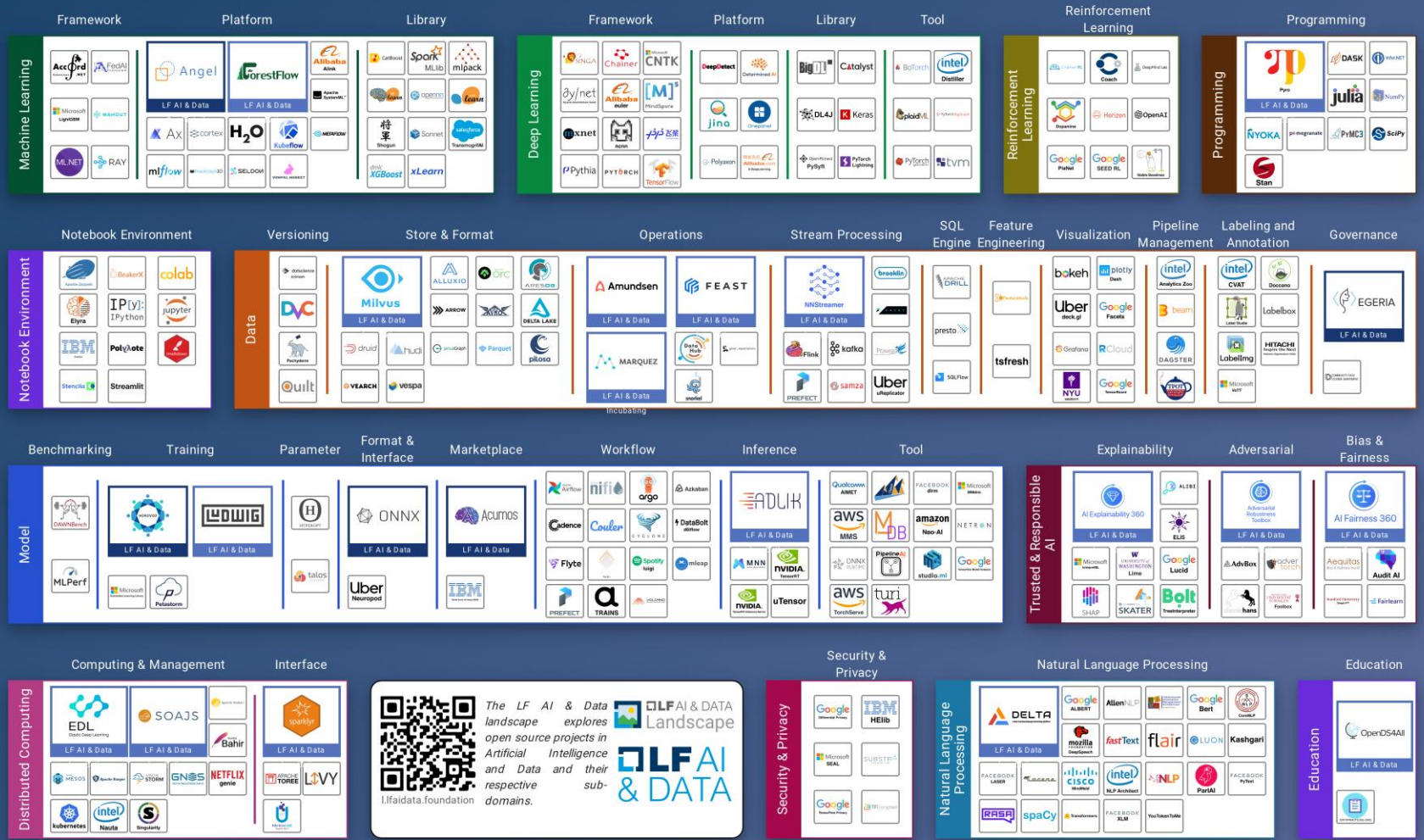
모니터링

서빙 모델 모니터링
데이터 변화 모니터링
시스템 안정성 모니터링

Linux Foundation AI Landscape
2020-11-01T21:39:53Z 83fd88a

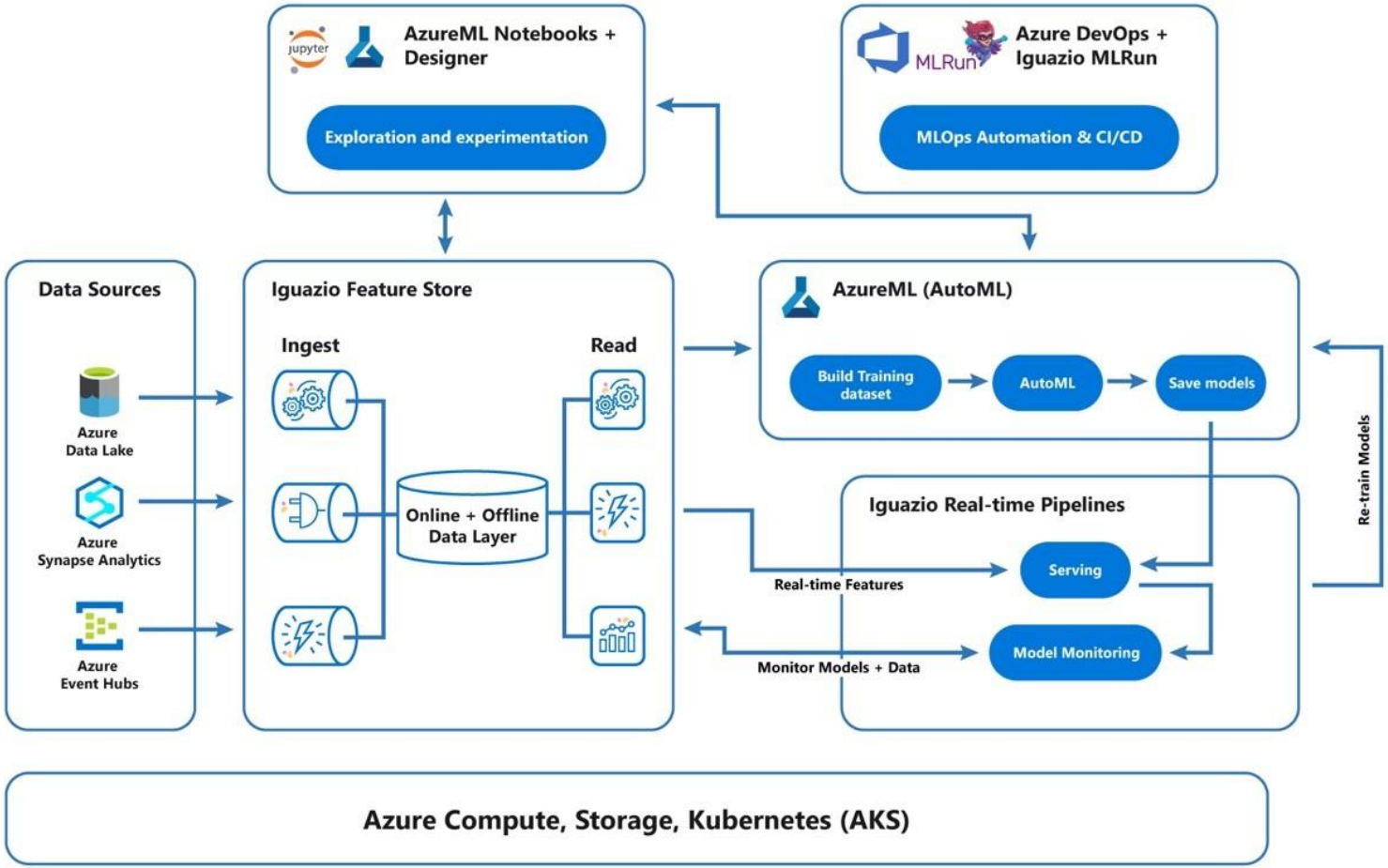
See the interactive landscape at l.lfai.foundation

Greyed logos are not open source



Microsoft

iguazio



Research와 Production의 머신러닝

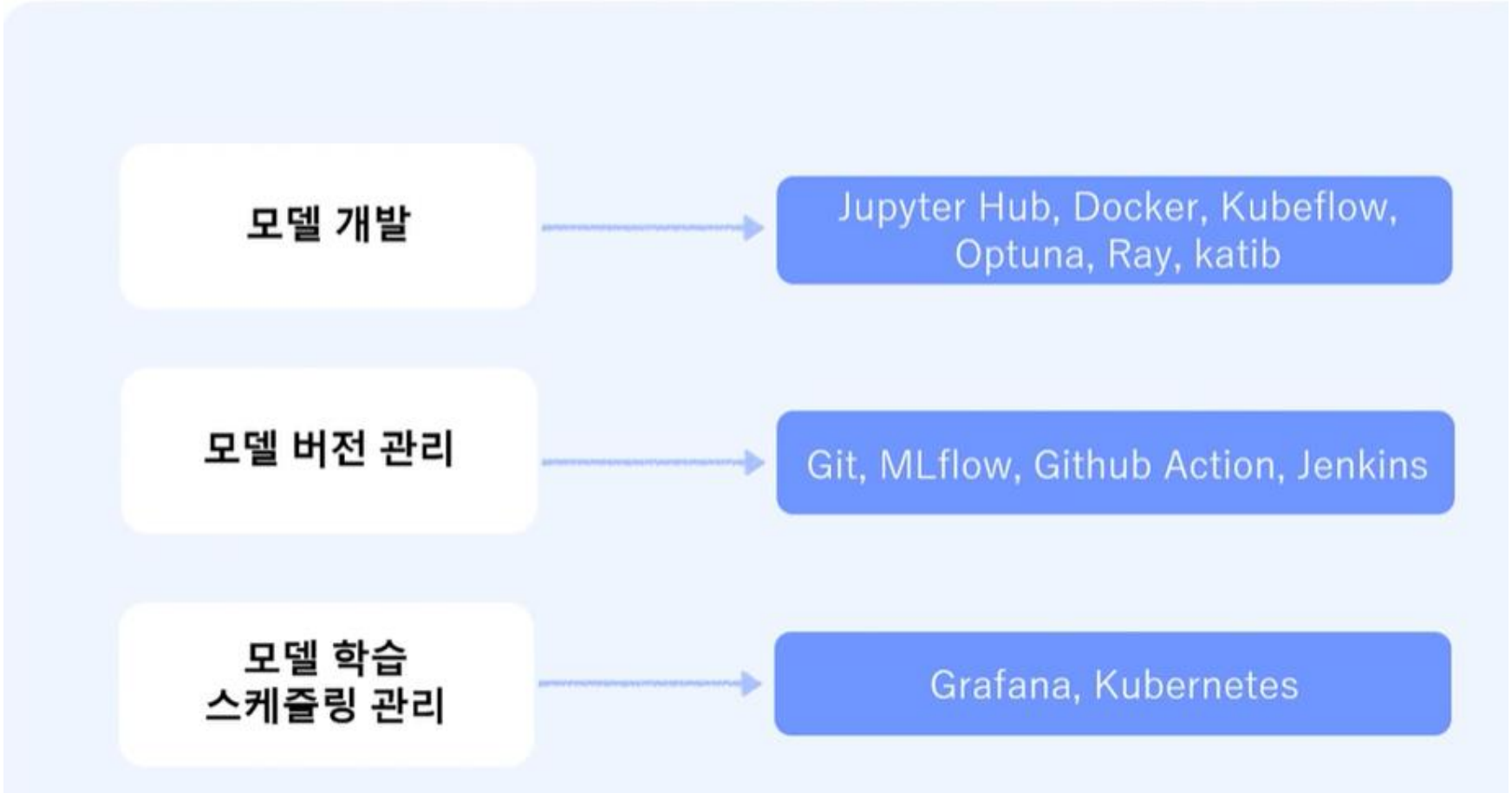
	Research ML	Production ML
데이터	고정(Static)	계속 변함(Dynamic - Shifting)
중요 요소	모델 성능(Accuracy, RMSE 등)	모델 성능, 빠른 Inference 속도, 해석 가능함
도전 과제	더 좋은 성능을 내는 모델, 새로운 구조의 모델	안정적인 운영, 전체 시스템 구조
학습	데이터가 고정이라 모델구조, 파라미터 기반 재학습	시간의 흐름에 따라 데이터가 변경되어 재학습
목적	논문 출판	서비스에서 문제 해결
표현	Offline	Online

요즘 MLOps

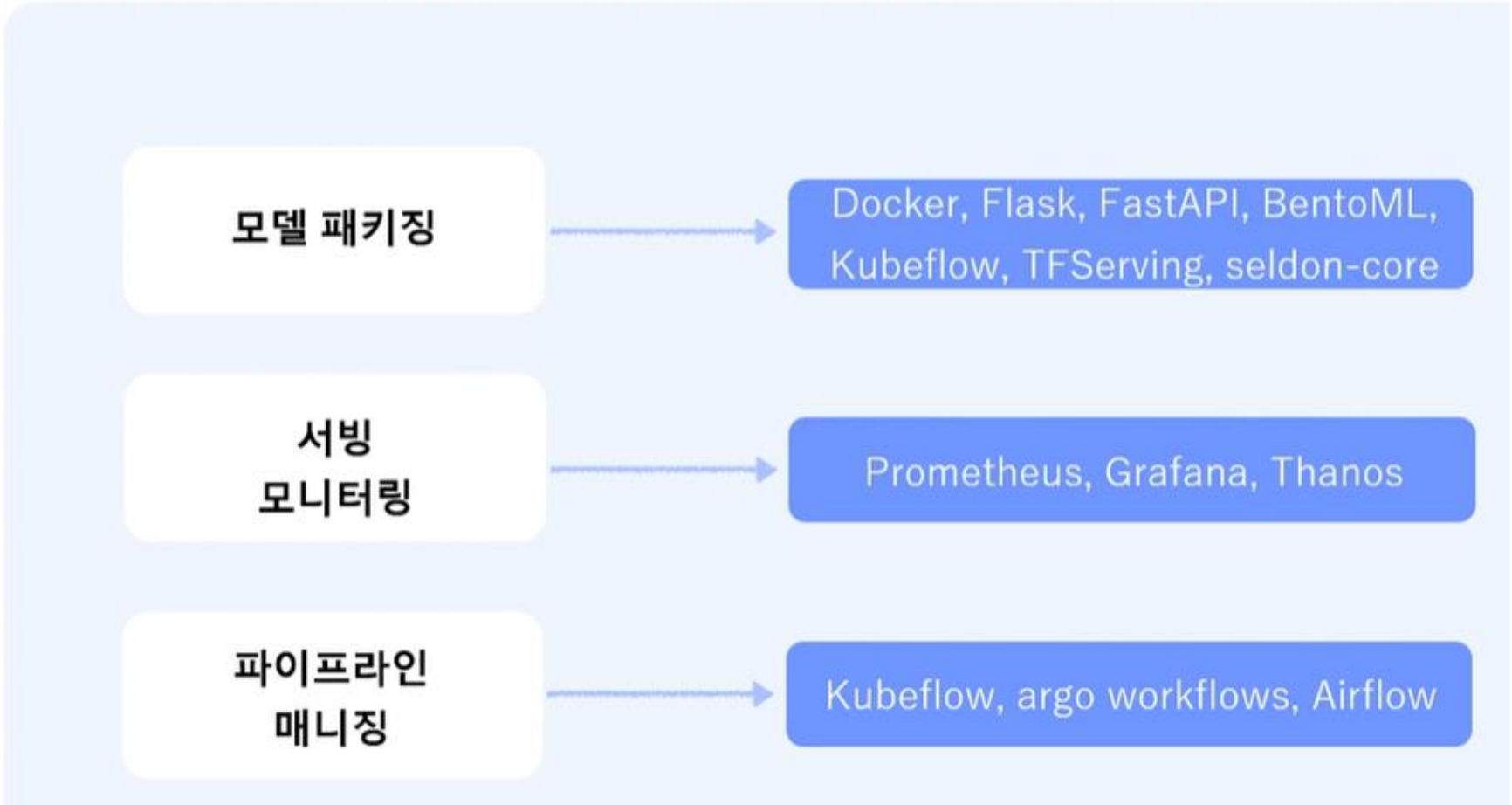
데이터


데이터 수집
파이프라인
Sqoop, Flume, Kafka, Flink,
Spark Streaming, Airflow
데이터 저장
MySQL, Hadoop, Amazon S3, MinIO
데이터 관리
TFDV, DVC, Feast, Amundsen

모델


모델 개발
Jupyter Hub, Docker, Kubeflow, Optuna, Ray, katib
모델 버전 관리
Git, MLflow, Github Action, Jenkins
모델 학습 스케쥴링 관리
Grafana, Kubernetes

서빙


모델 패키징
Docker, Flask, FastAPI, BentoML,
Kubeflow, TFServing, seldon-core
서빙
모니터링
Prometheus, Grafana, Thanos
파이프라인
매니징
Kubeflow, argo workflows, Airflow

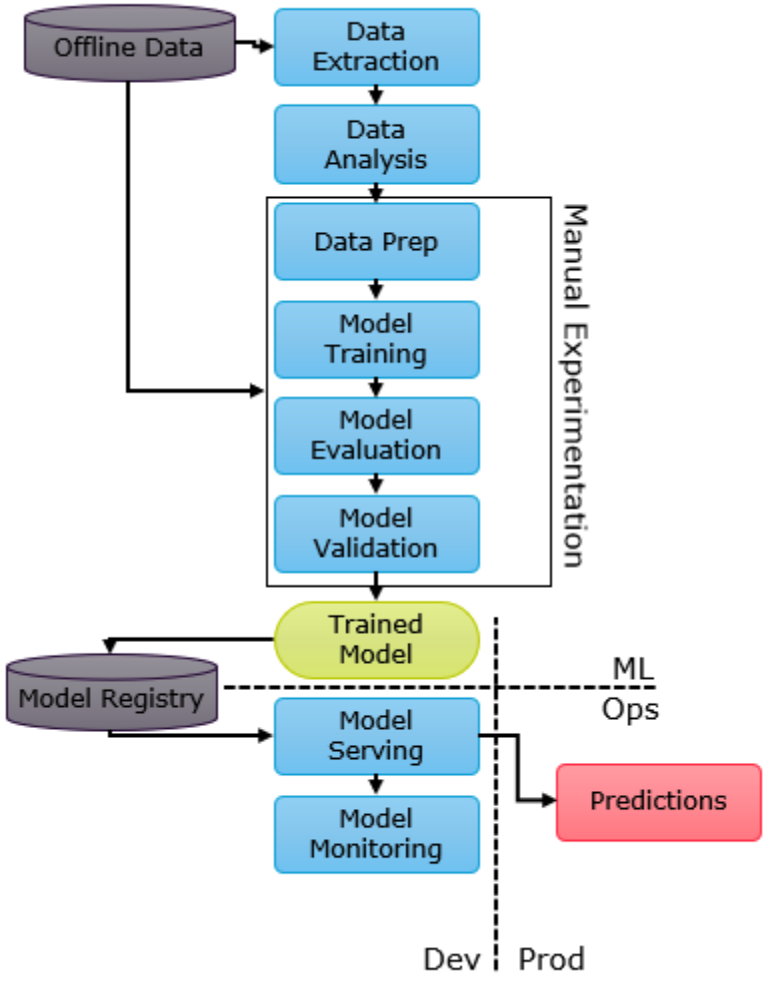
MLOps Components

Serving
Experiment, Model Management
Feature Store
Continuous Training
Model Analysis
Auto ML
MLOps Infra Management(GPU 등)
Monitoring
……

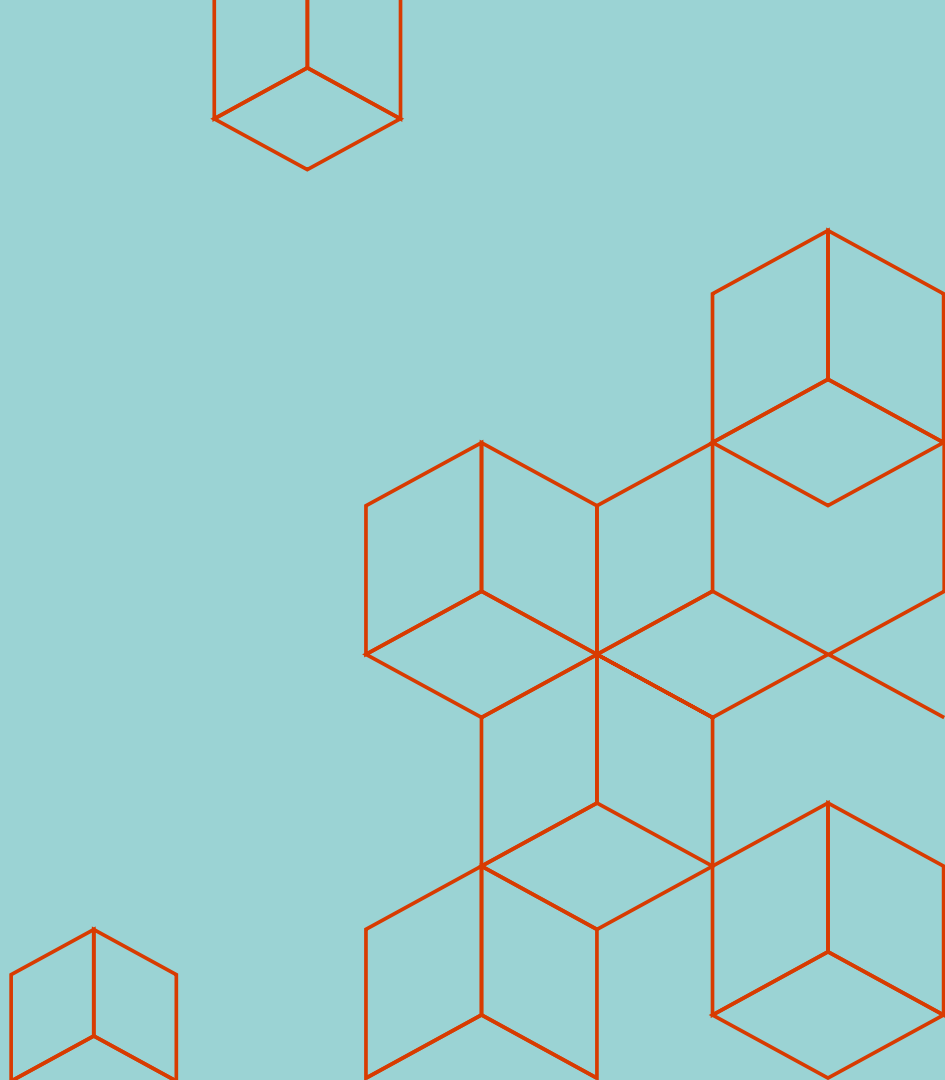
MLOps Components

Serving
Experiment, Model Management
Feature Store
Continuous Training
Model Analysis
Auto ML
MLOps Infra Management(GPU 등)
Monitoring
……

Flask
MLflow
Git
DVC
Feature Store
Jenkins
Git Action
Kubeflow
Docker
Kubernetes
Prometheus
Grafana


Offline Data
Data Extraction
Data Analysis
Data Prep
Model Training
Model Evaluation
Model Validation
Manual Experimentation
Trained Model
Model Registry
ML Ops
Model Serving
Predictions
Model Monitoring
Dev
Prod

Microsoft Machine Learning



Domain Specific Pretrained Models
To reduce time to market


Vision


Speech


Language


Search

Familiar Data Science Tools
To simplify model development


PyCharm


Jupyter


Visual Studio Code


Command line

Popular Frameworks
To build machine learning and deep learning solutions


PyTorch


TensorFlow


Scikit-Learn


ONNX

Productive Services
To empower data science and development teams


Azure Databricks


Azure Machine Learning


Machine Learning VMs

Powerful Hardware
To accelerate deep learning


CPU


GPU


FPGA


From the Intelligent Cloud to the Intelligent Edge


Azure Machine Learning Service

Set of Azure Cloud Services

Python SDK

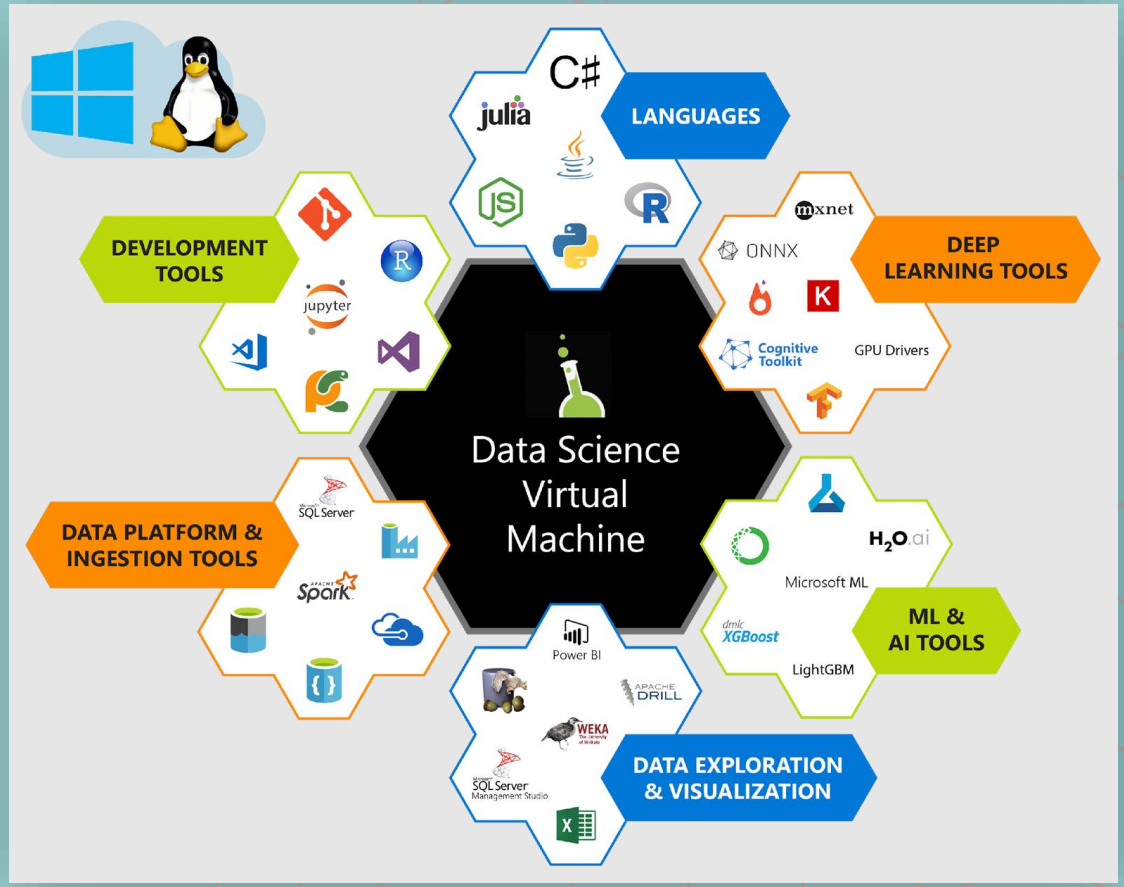
That enables you to:

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

Data Science Virtual Machines (DSVM)

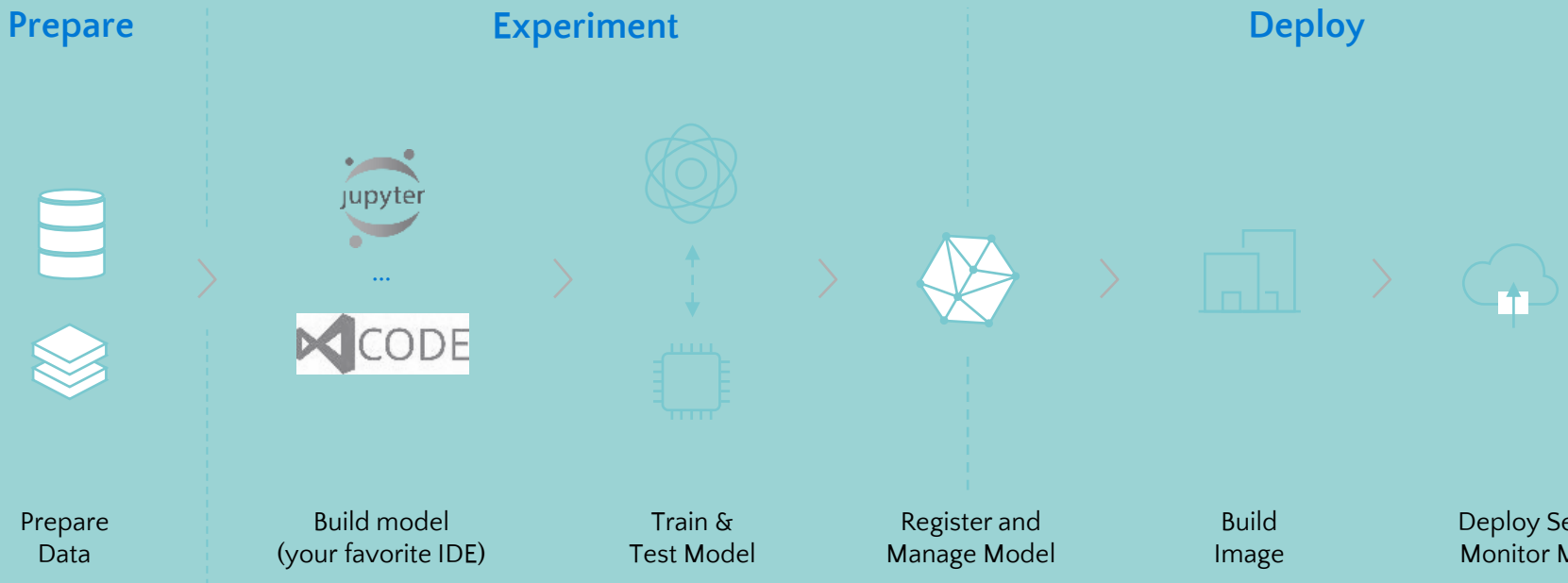
Pre-Configured environments in the cloud for
Data Science & AI Modeling, Development & Deployment.

Samples to get started

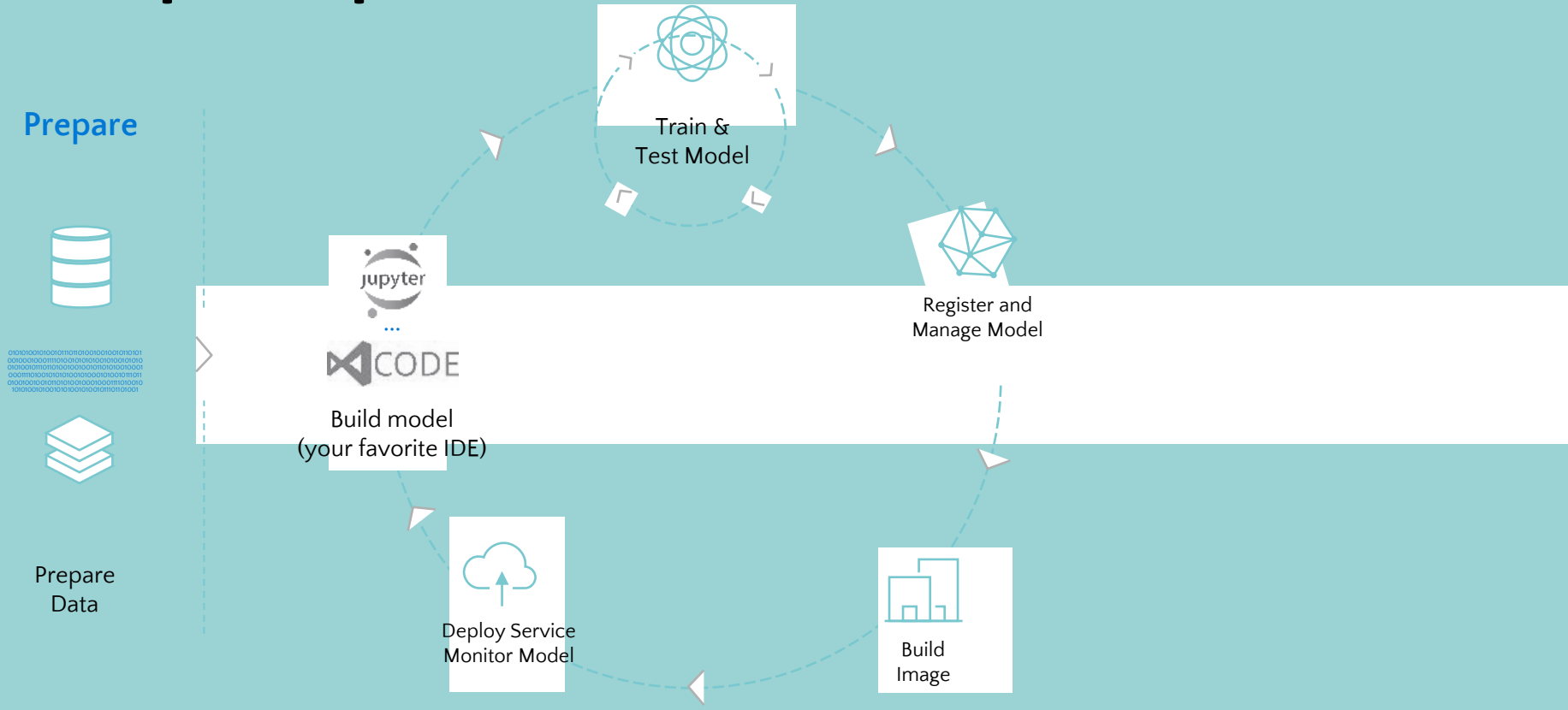


DevOps for Machine Learning

DevOps loop for data science



DevOps loop for data science



Model Management in detail

Create/Retrain Model

Enable DevOps with full CI/CD integration with VSTS

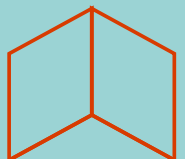
Register Model

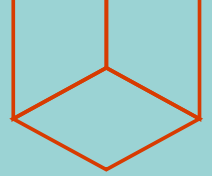
Track model versions with a central model registry

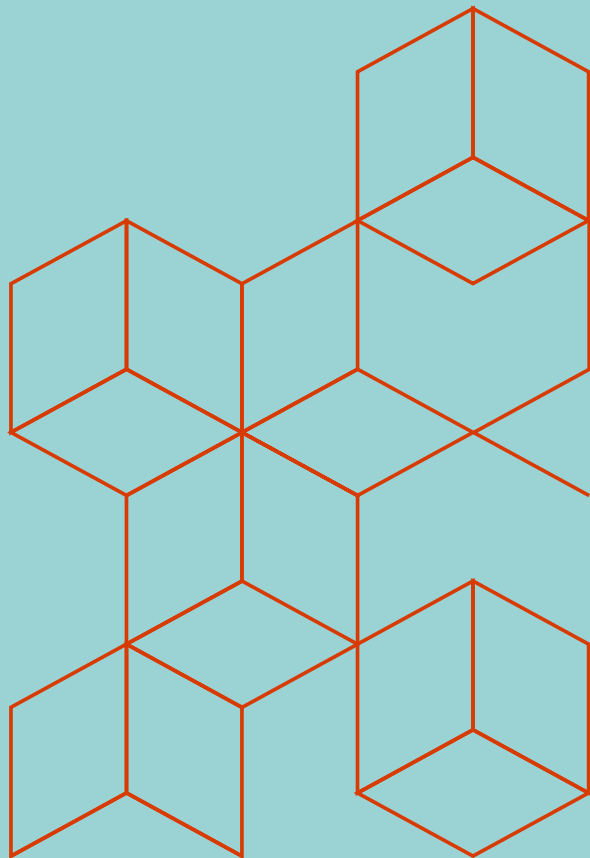
Monitor

Oversea deployments through Azure AppInsights

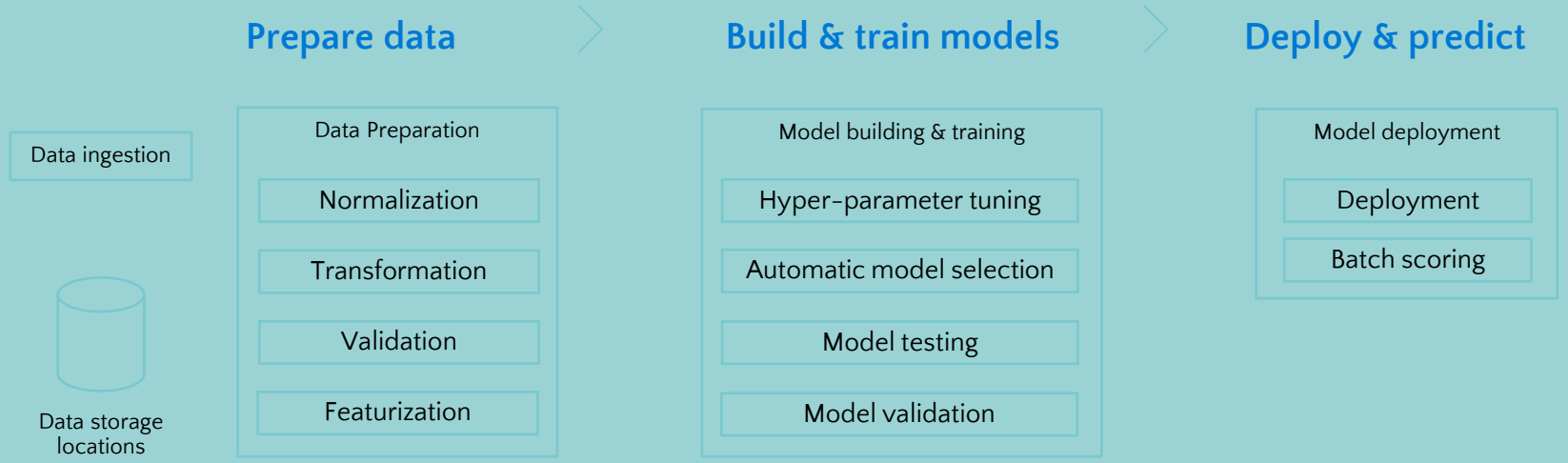
Azure Machine Learning Pipelines



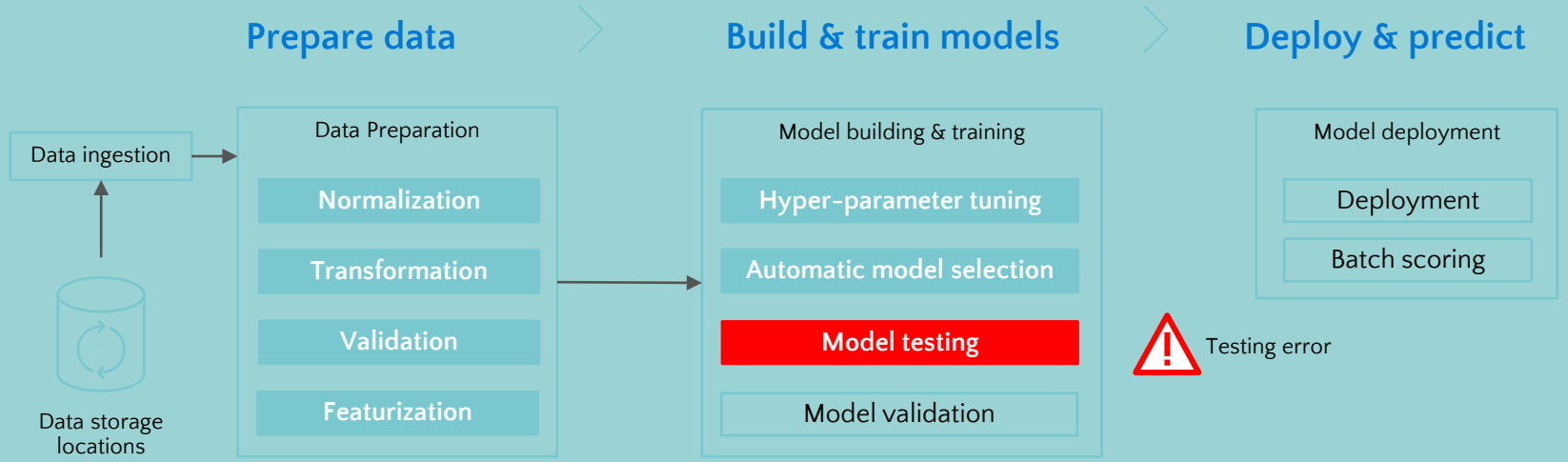




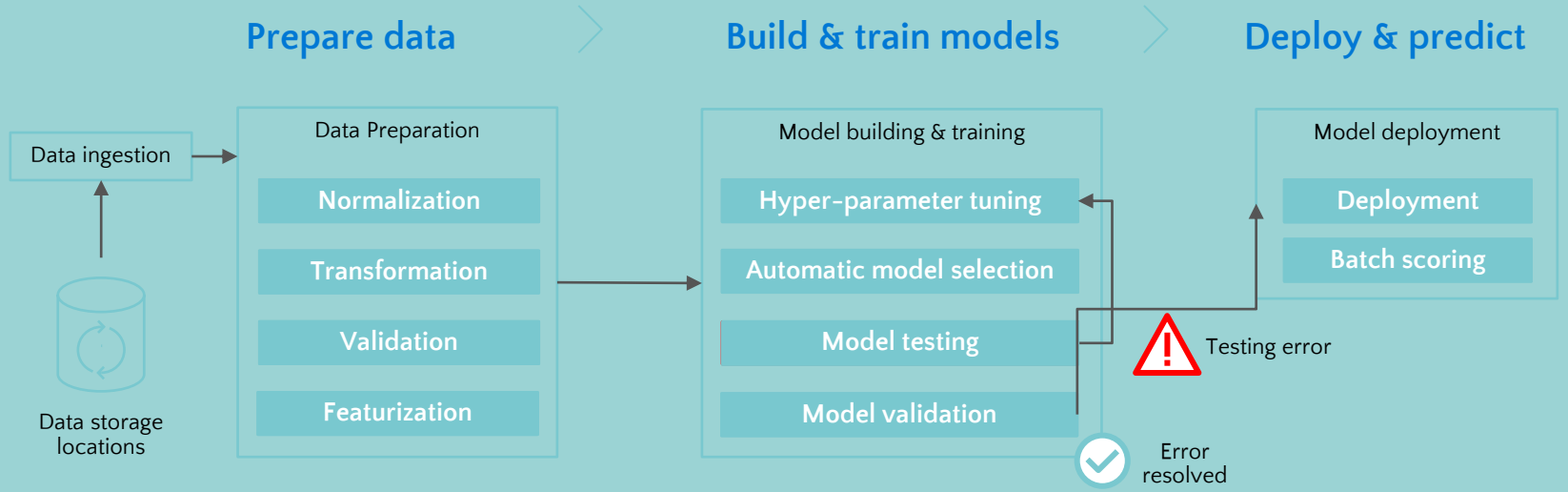
Azure Machine Learning Pipelines



Azure Machine Learning Pipelines



Azure Machine Learning Pipelines



Azure Machine Learning Pipelines with new data