Introduction to fully automated MLOps

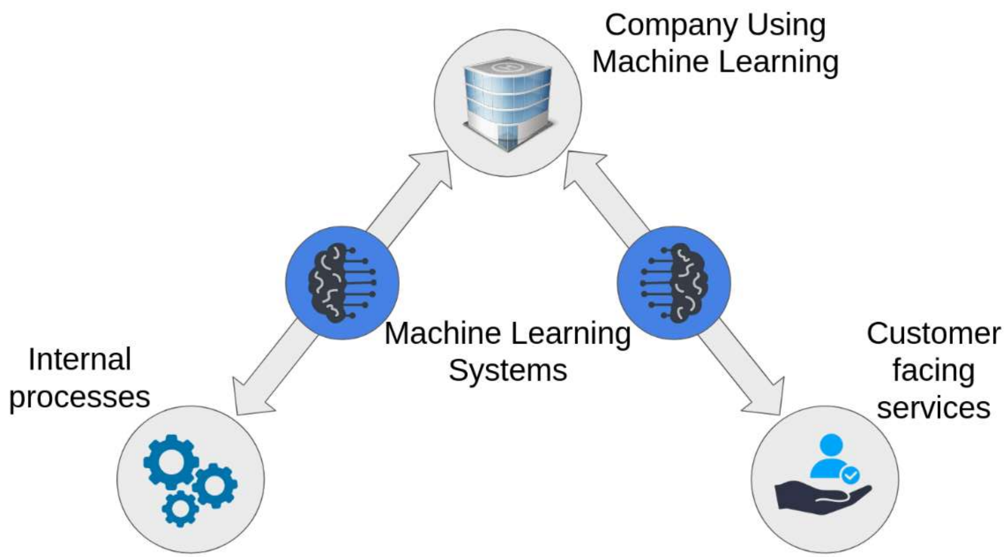
FULLY AUTOMATED MLOPS

Arturo Opsetmoen Amador
Senior Consultant - Machine Learning

MLOps in an industrial setting

What are common goals companies have when using machine learning?

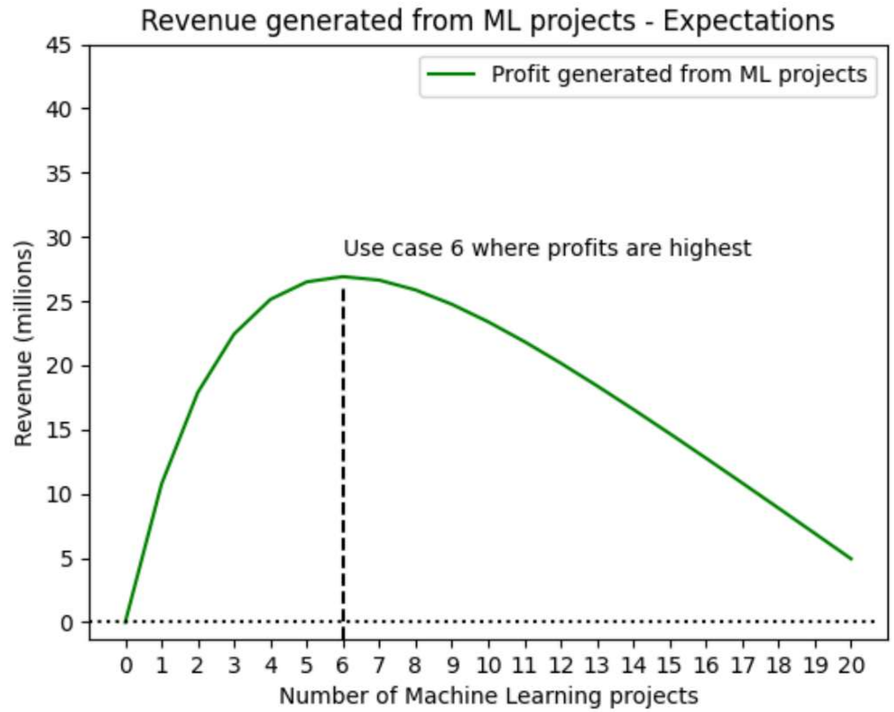
- Develop ML tools and products that use data to
  - Better serve customers
  - Optimize processes



Optimizing for value generation

Companies aim to maximize profits

- Machine learning can be used to increase profitability
- By analyzing costs and revenues, a company can estimate potential profits
- Deploying the right number of ML use-cases can result in profits



Costs in software development

Expected costs in a traditional software development project:

- Development costs
- Project management
- UI/UX design
- Quality assurance

Technical debt in software development

Technical debt or design debt:

- Cost of rework caused by poor design

[1] https://vincentdnl.com/drawings/

Hidden technical debt in ML systems

Machine Learning: "The high-interest credit card of technical debt"[1].

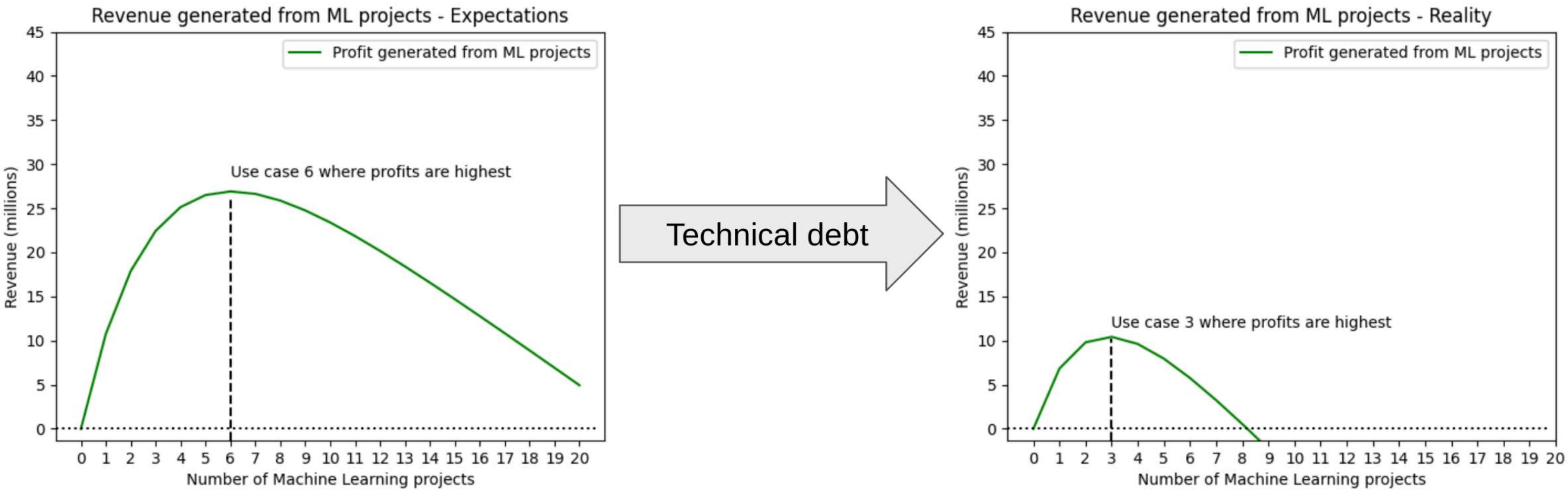
Hidden technical debt can be related to:

1. The data used to train the ML models
2. The models powering the ML system
3. The infrastructure used by the ML system
4. The monitoring of the ML system

[1] https://research.google/pubs/pub43146/

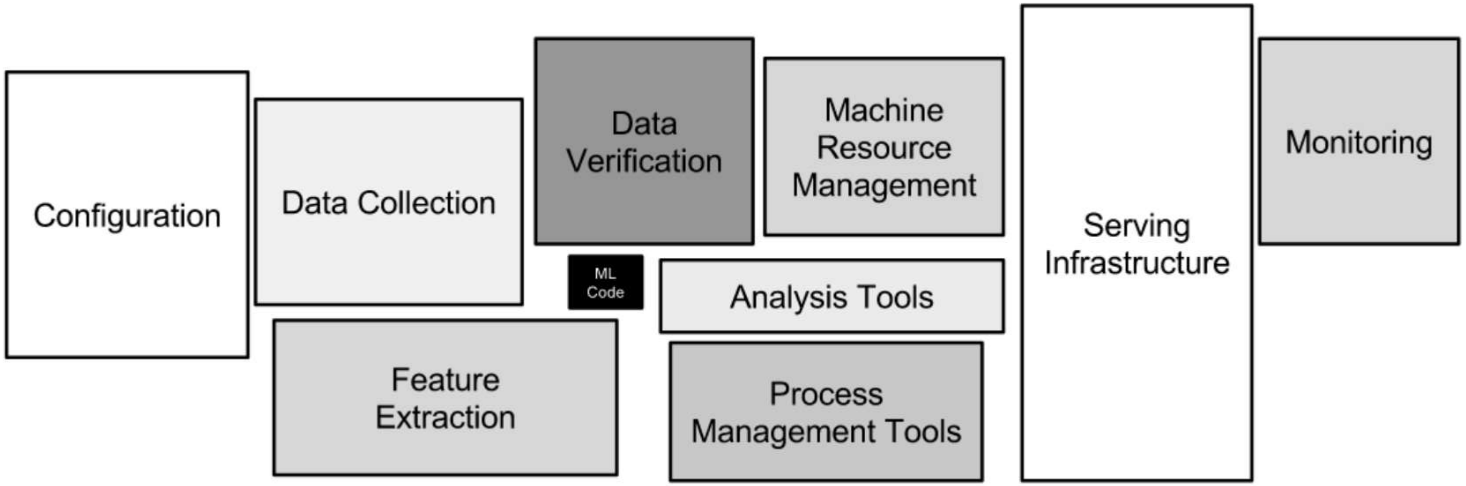
Costs of machine learning projects

Reduced profit due to technical debt in ML systems:



The high-interest credit card of technical debt

ML systems can be complex and become unruly.



MLOps: The best-known way to pay

If ML is the high-interest credit card of technical debt, MLOps is the best way to pay for it.

MLOps can include:

- Automated testing
- Automated experiment tracking
- Automated monitoring

To keep the technical debt to a minimum

Let's practice!

FULLY AUTOMATED MLOPS

datacamp

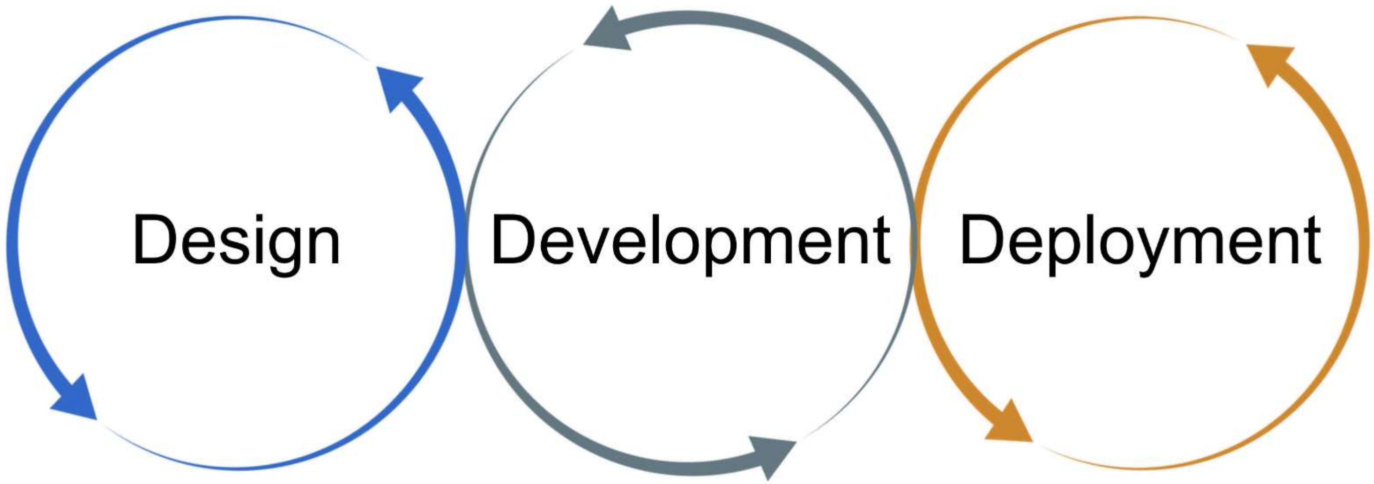
MLOps lifecycle stages

FULLY AUTOMATED MLOPS

Arturo Opsetmoen Amador
Senior Consultant - Machine Learning

The MLOps lifecycle

The MLOps lifecycle includes three core stages:

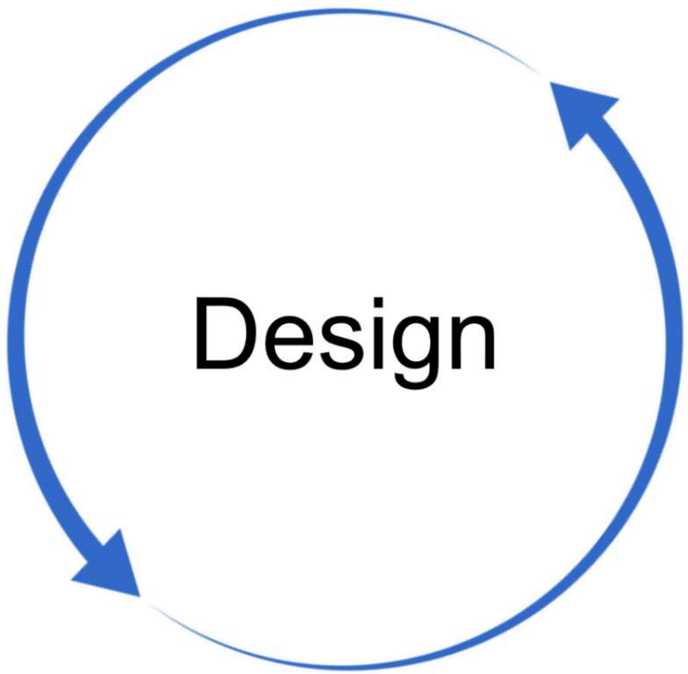


- The three stages are iterative
- The three stages are interconnected and rely on each other
- It is normal to go back and forth between stages

MLOps in the ML lifecycle - Design

The Design stage includes:

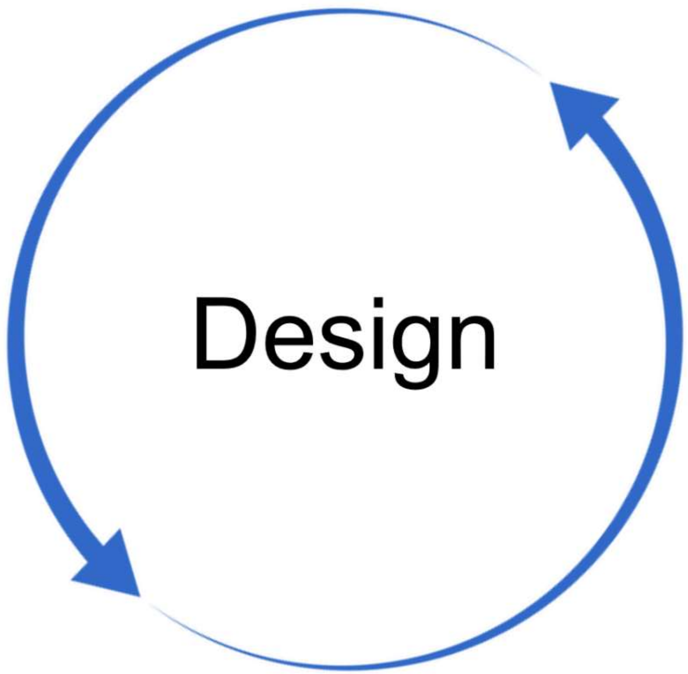
- BUSINESS UNDERSTANDING
- DATA UNDERSTANDING
- DESIGNING THE ML SOLUTION



MLOps in the ML lifecycle - Design

The Design stage includes:

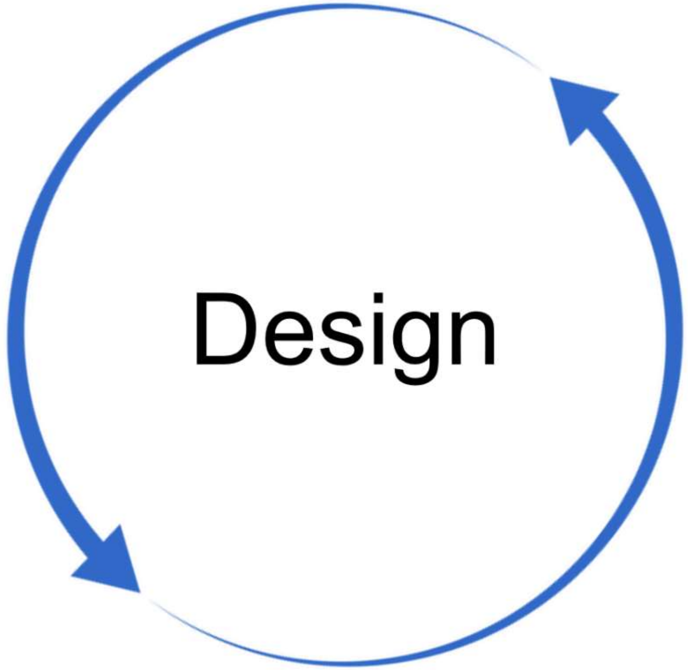
- **BUSINESS UNDERSTANDING**
  - Understanding the business context
  - Establishing the business goals



MLOps in the ML lifecycle - Design

The Design stage includes:

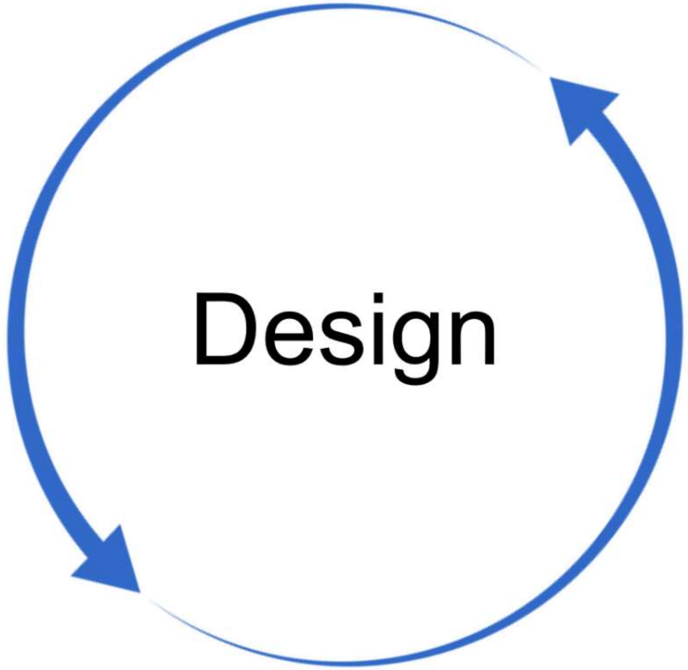
- **DATA UNDERSTANDING**
  - Data exploration
  - Data visualization



MLOps in the ML lifecycle - Design

The Design stage includes:

- **SYSTEM DESIGN**
  - System architecture design
    - Data security
    - Data privacy

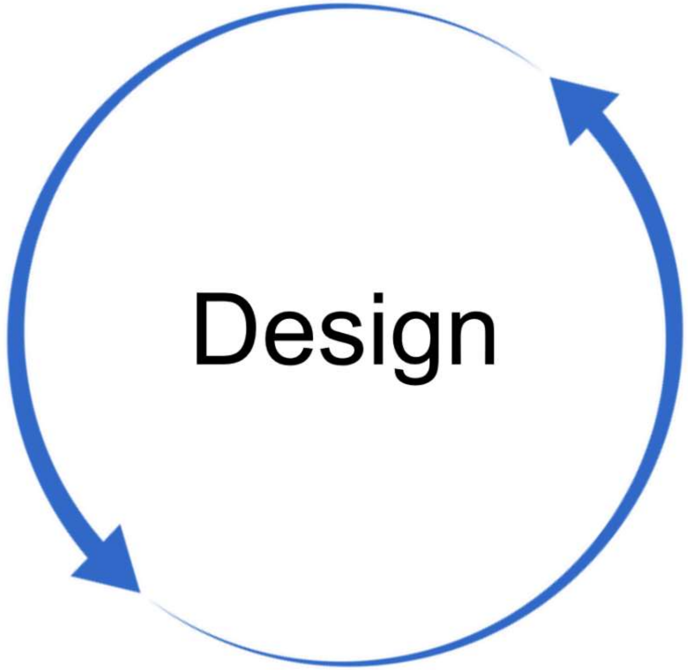


MLOps in the ML lifecycle - Design

The Design stage includes:

- BUSINESS UNDERSTANDING
- DATA UNDERSTANDING
- DESIGNING THE ML SOLUTION

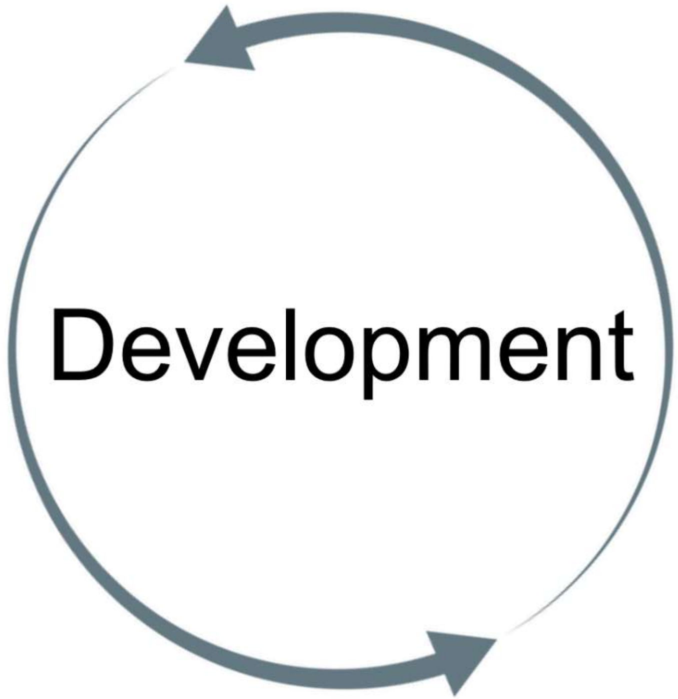
Much of the design phase can't be automated



MLOps in the ML lifecycle - Development

The ML experimentation and development stage includes:

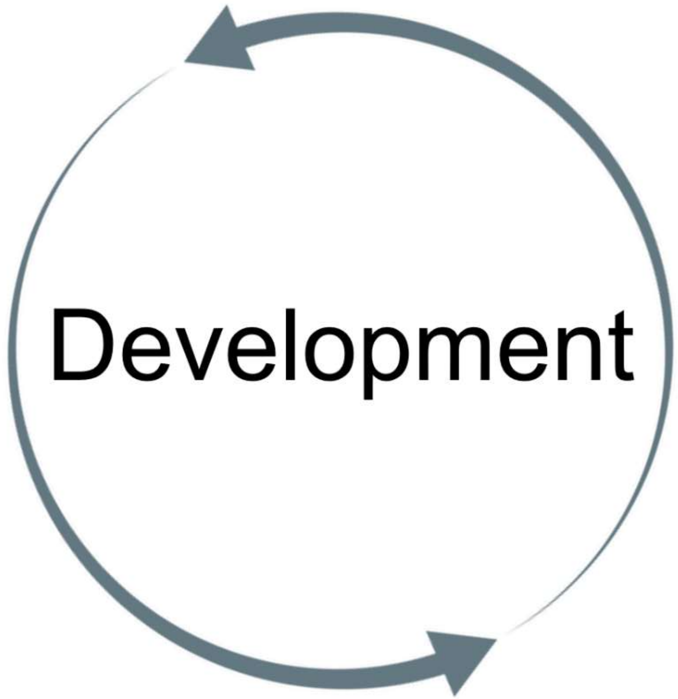
- DEVELOPING PROOF-OF-CONCEPTS (POCS)
- DATA ENGINEERING
- MODEL DEVELOPMENT



MLOps in the ML lifecycle - Development

The ML experimentation and development stage includes:

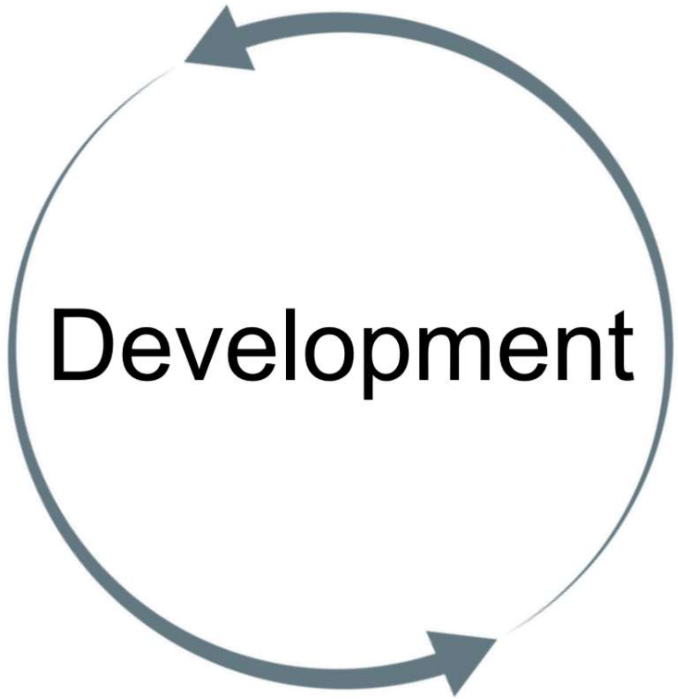
- DEVELOPING PROOF-OF-CONCEPTS (POCS)
  - Combination of process & automation
  - Use of frameworks



MLOps in the ML lifecycle - Development

The ML experimentation and development stage includes:

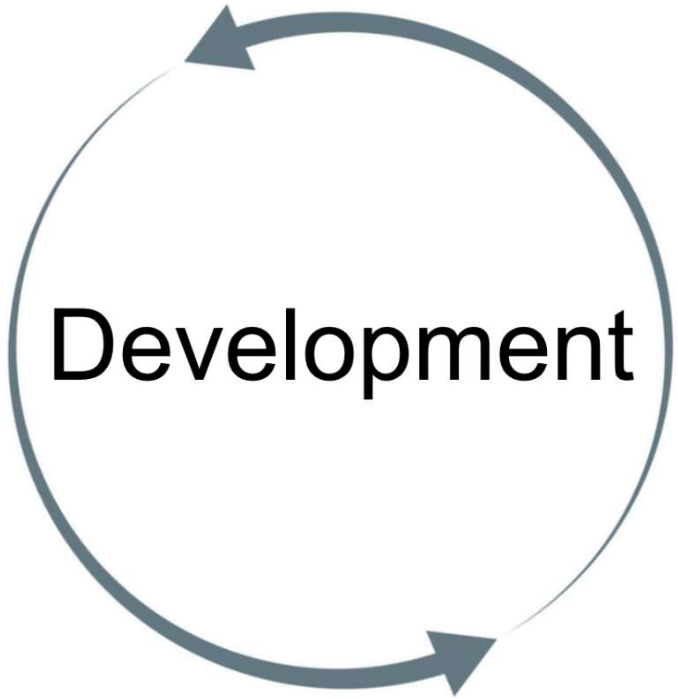
- **DATA ENGINEERING (POCS)**
  - Use of automation --> Quality assurance



MLOps in the ML lifecycle - Development

The ML experimentation and development stage includes:

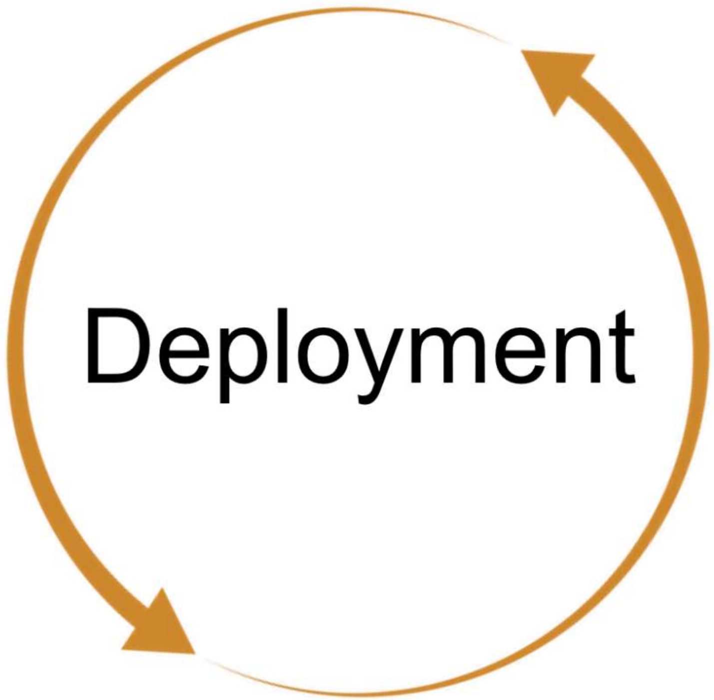
- MODEL DEVELOPMENT (POCS)
  - Experiment tracking
  - Automated training
  - Automated hyperparameter tuning



MLOps in the ML lifecycle - Deployment

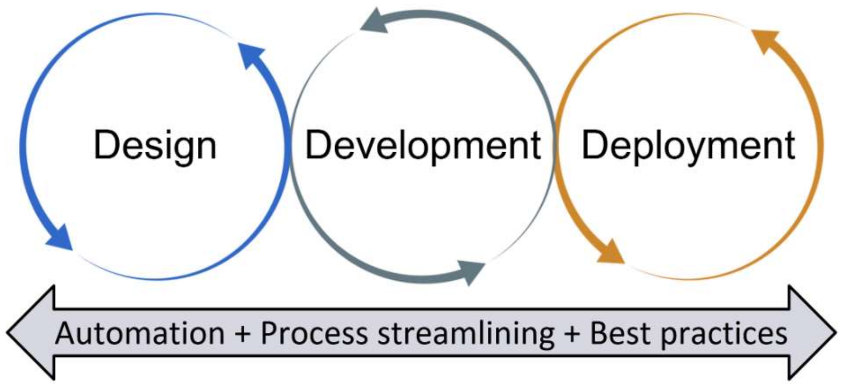
The ML deployment and operations phase include:

- Productionization
  - Testing
  - Versioning
  - Continuous delivery
  - Monitoring



Building for scale: Automation first

We use process streamlining, best practices, and automation



When no automation, process streamlining:

- CRISP-DM
- TDSP [1]

[1] https://www.ibm.com/docs/en/spss-modeler/saas?topic=dm-crisp-help-overview

Process streamlining and best practices

DESIGN PHASE

- Best practices
  - Include domain expertise
  - Involve business stakeholders
  - Get feedback from end-users

DEVELOPMENT PHASE

- Best practices
  - Write clean code
  - Document our work

Let's practice!

FULLY AUTOMATED MLOPS

Reference architecture: Fully automated MLOps

FULLY AUTOMATED MLOPS

Arturo Opsetmoen Amador
Senior Consultant - Machine Learning

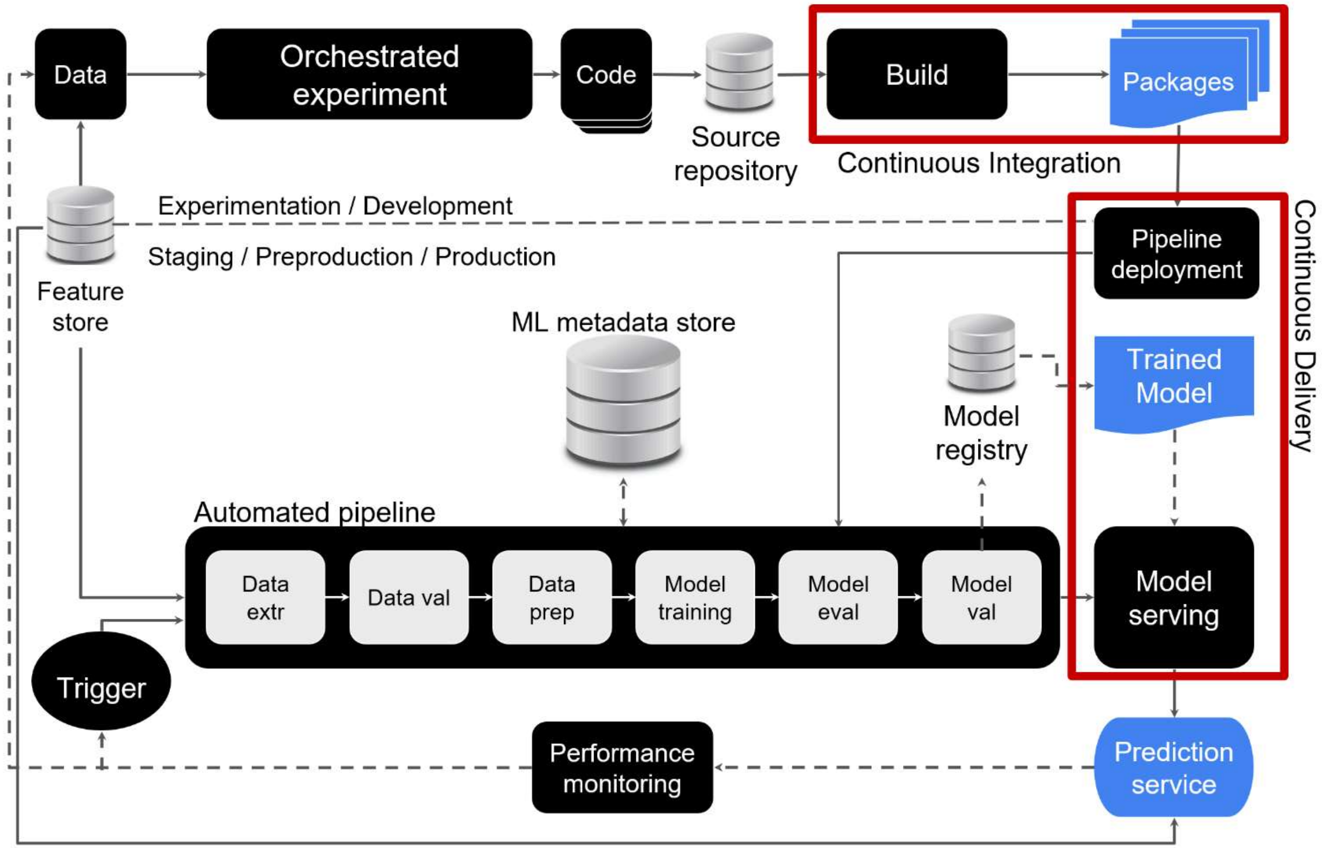
What is a reference architecture?

A reference architecture is a template we can use to design solutions in different IT domains.

A set of documents with:

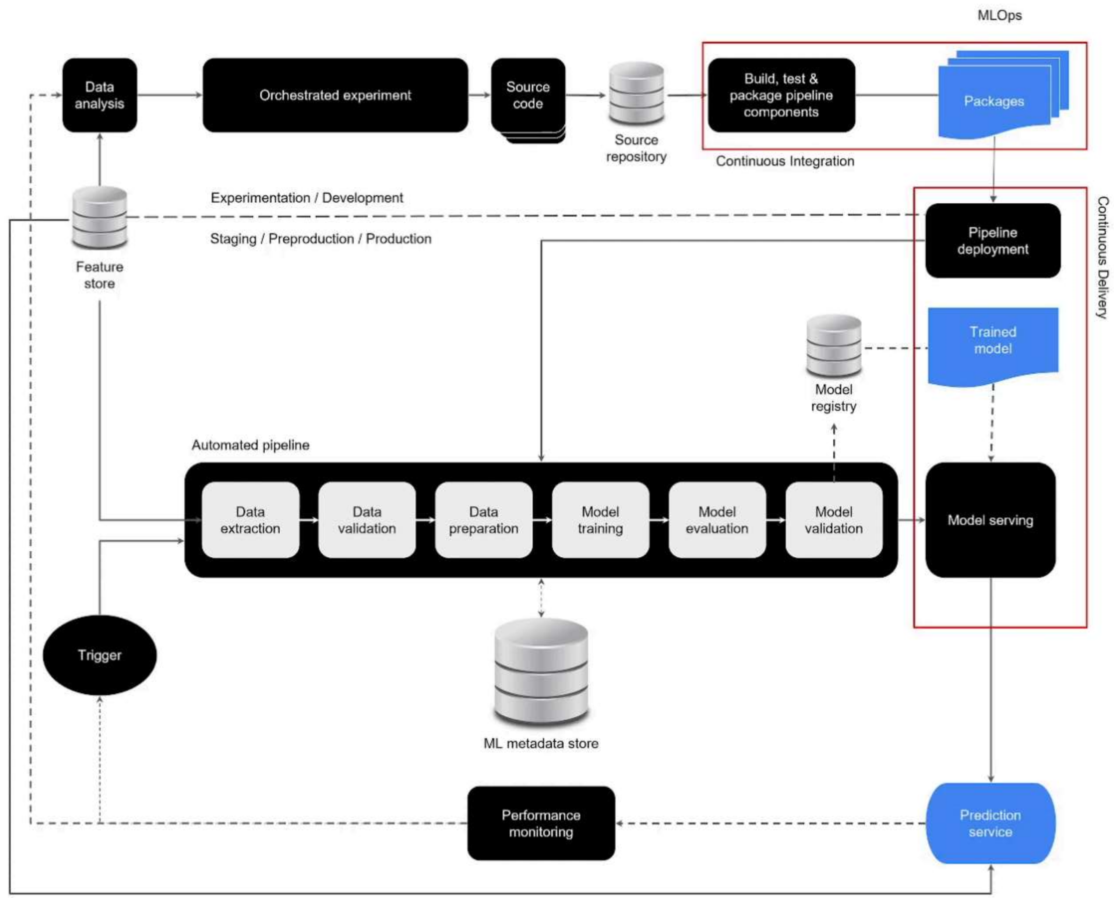
- Structures and integrations of IT elements
- Patterns commonly present in IT systems, including ML systems
- Leverage experience and best practice from industry players

Fully automated MLOps architecture

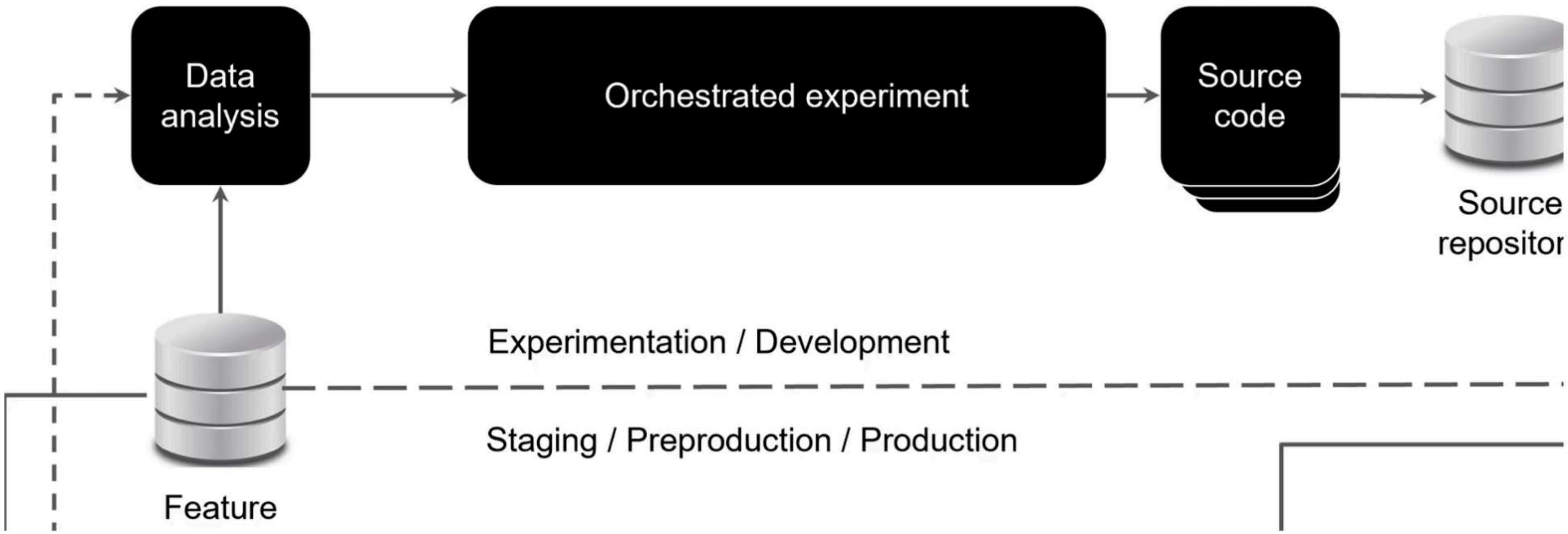


[1] https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

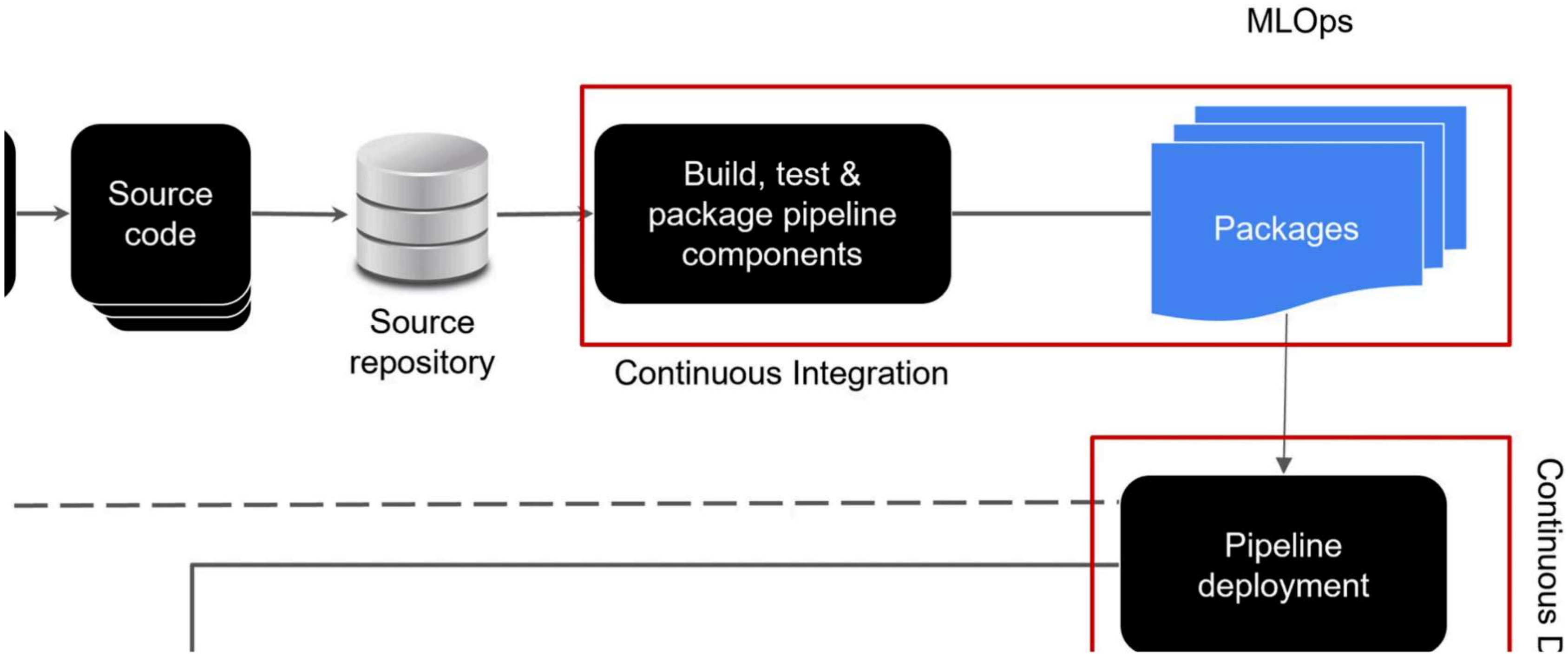
Reference architecture - Orchestrated experiments



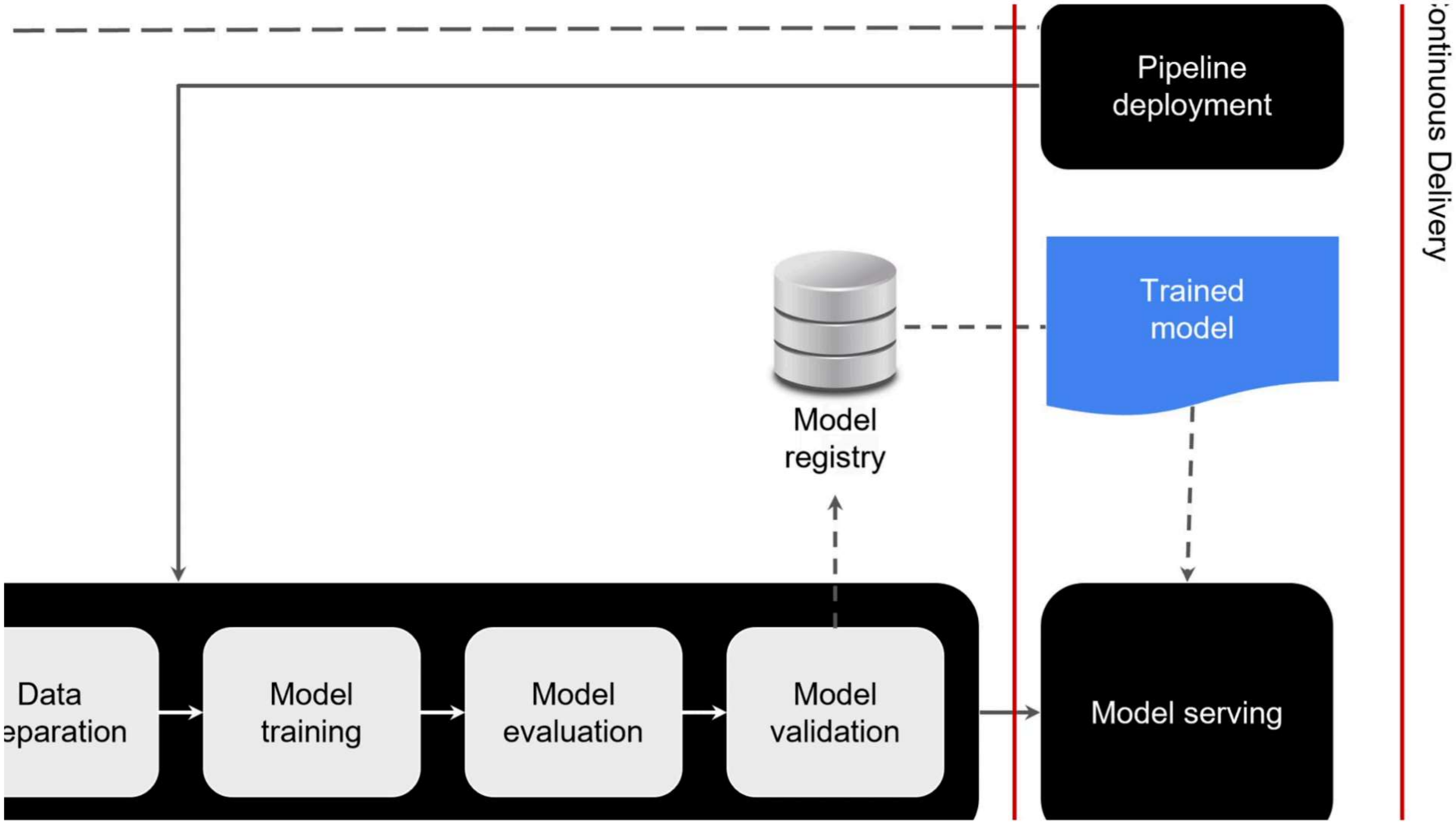
Reference architecture - Source code & CI



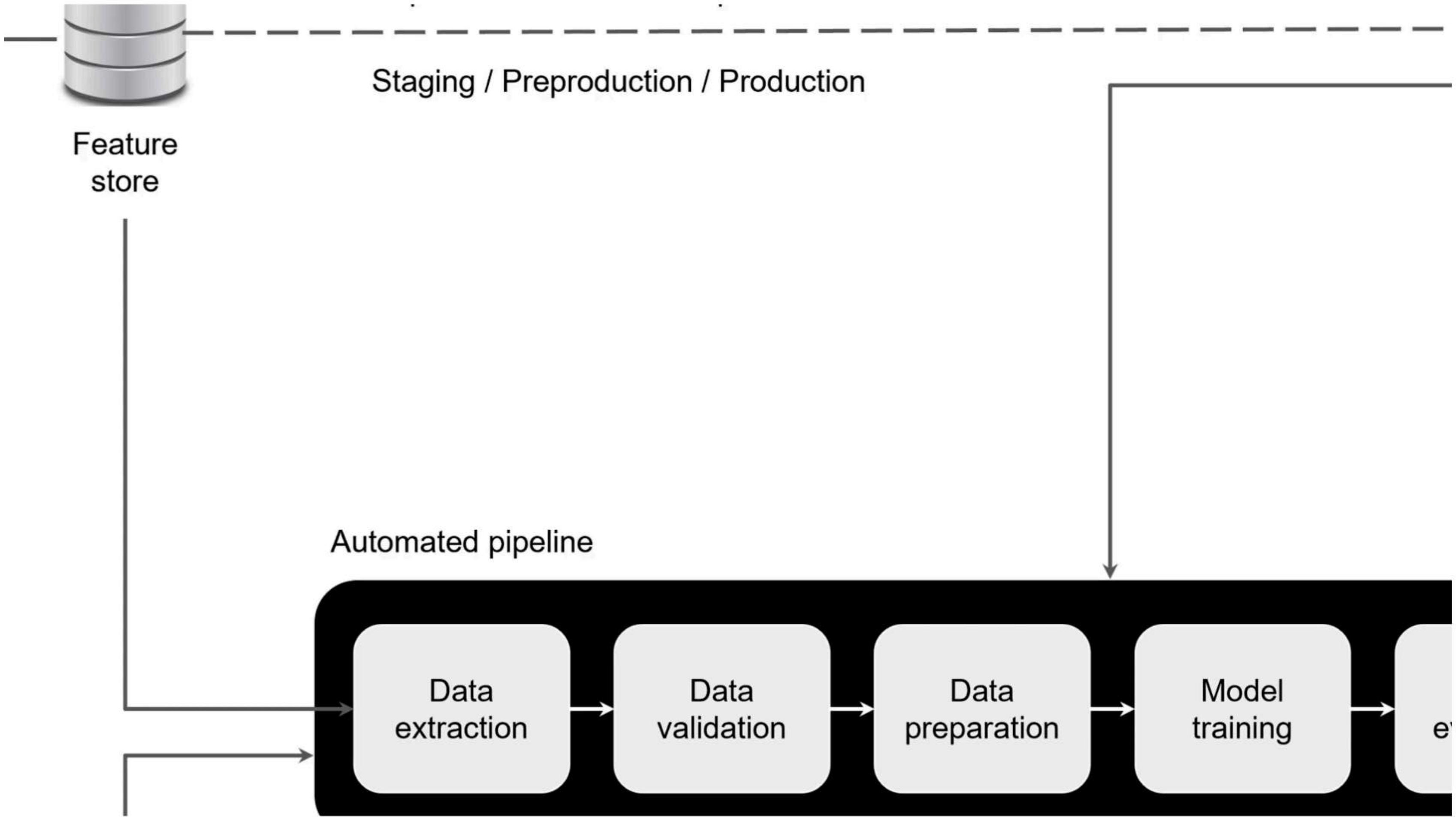
Reference architecture - Artifacts & CD



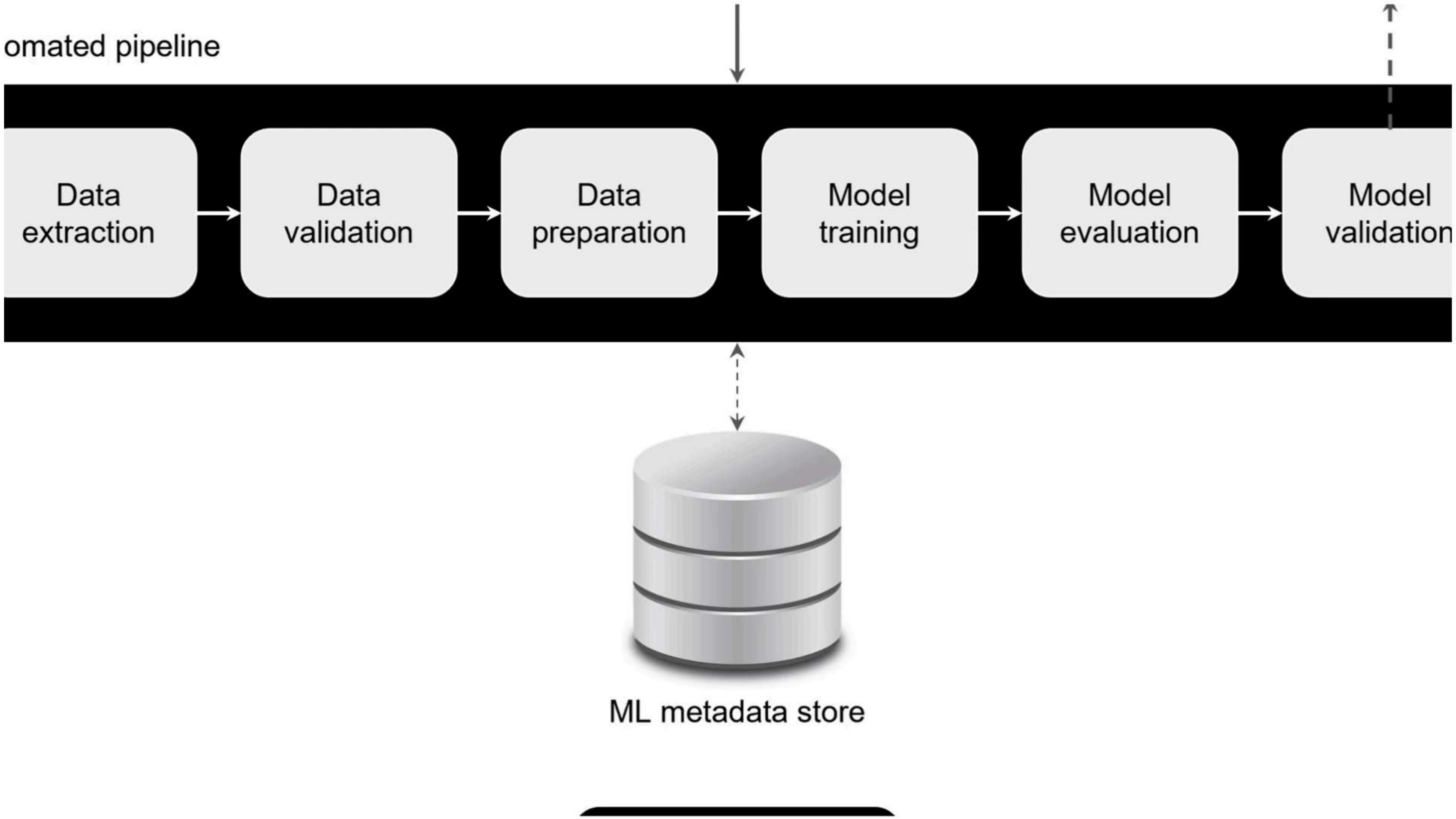
Reference architecture - ML pipeline deployment



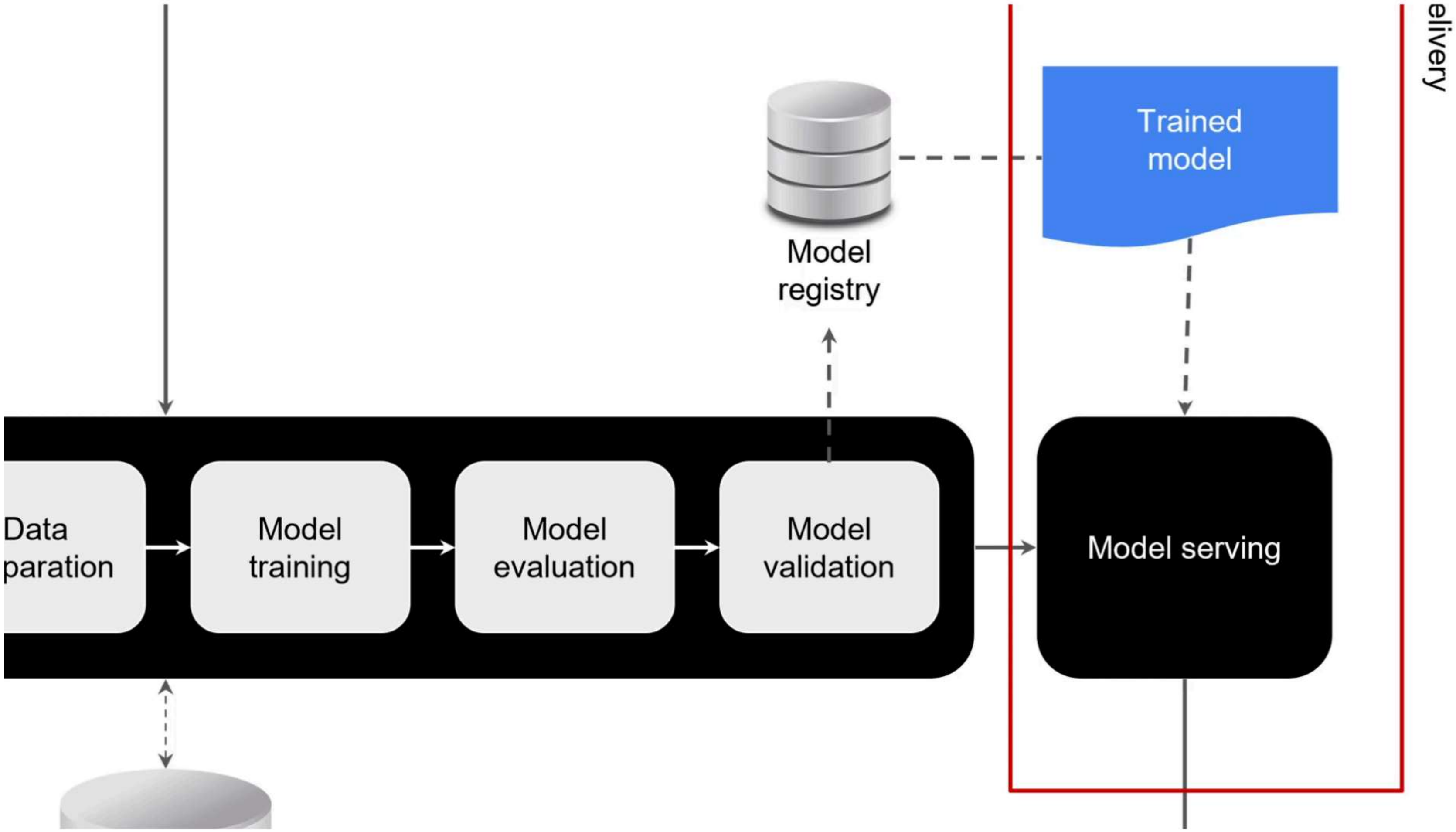
Reference architecture - The metadata store



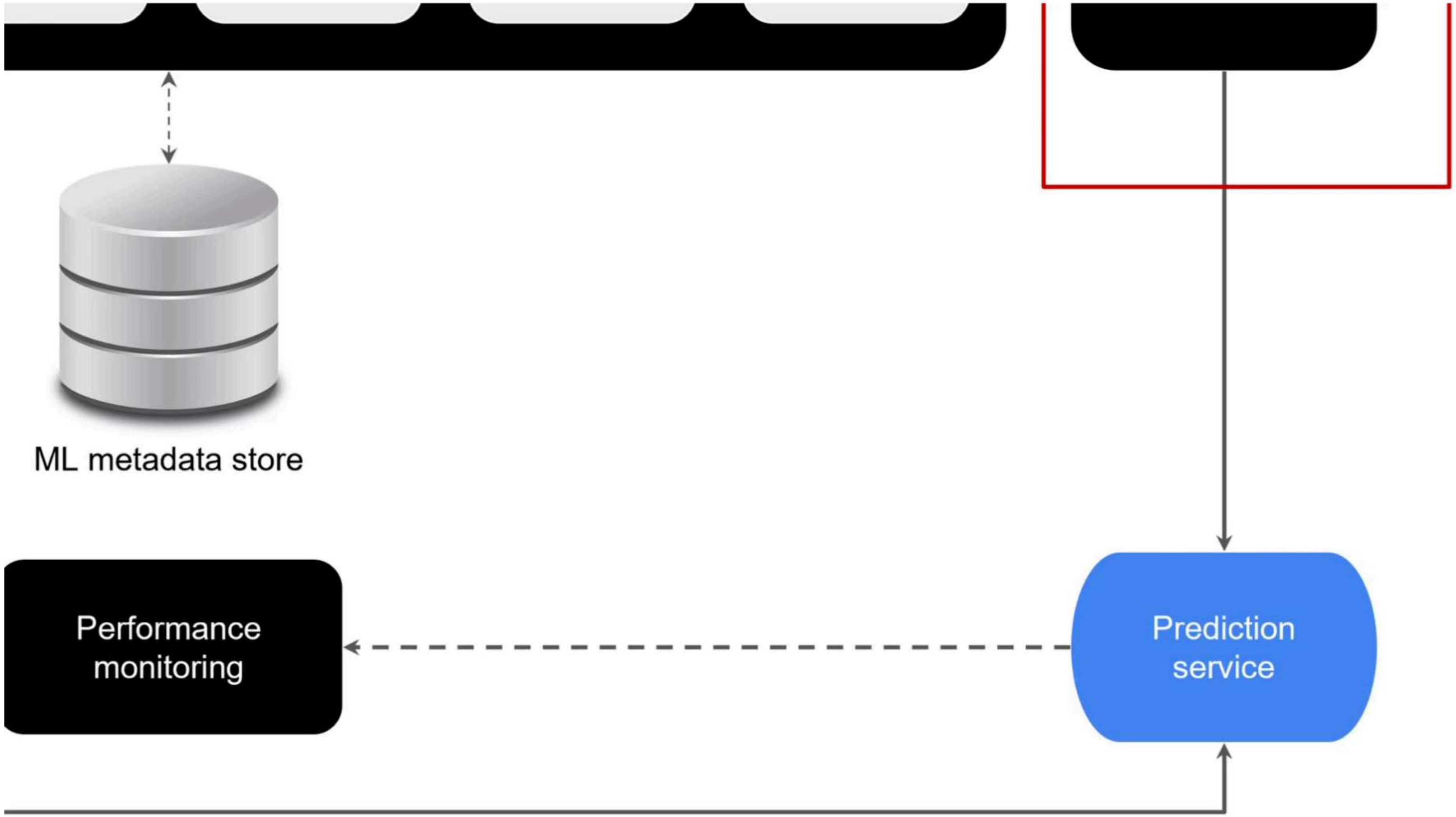
Reference architecture - The model registry



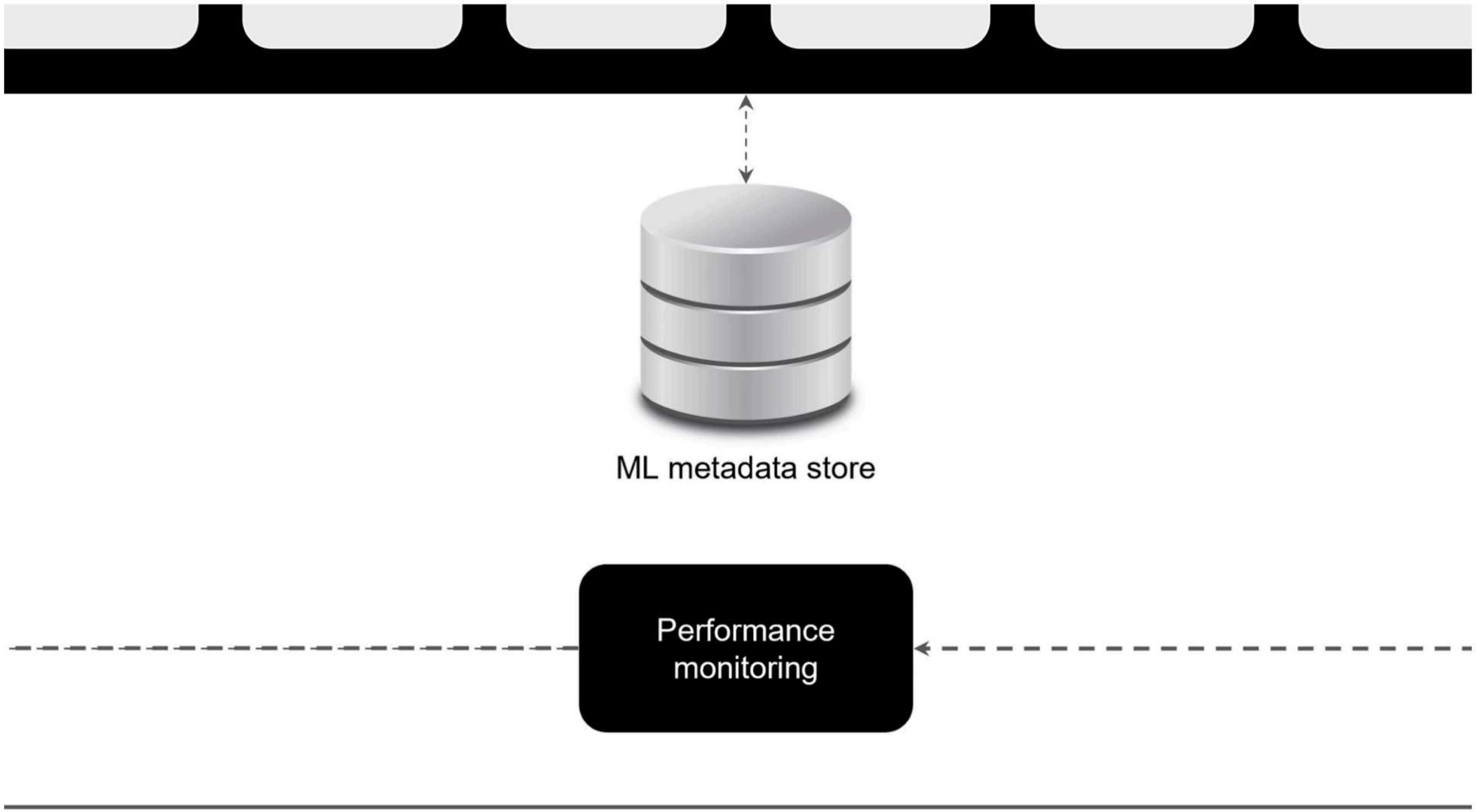
Reference architecture - Prediction services



Reference architecture - Continuous monitoring



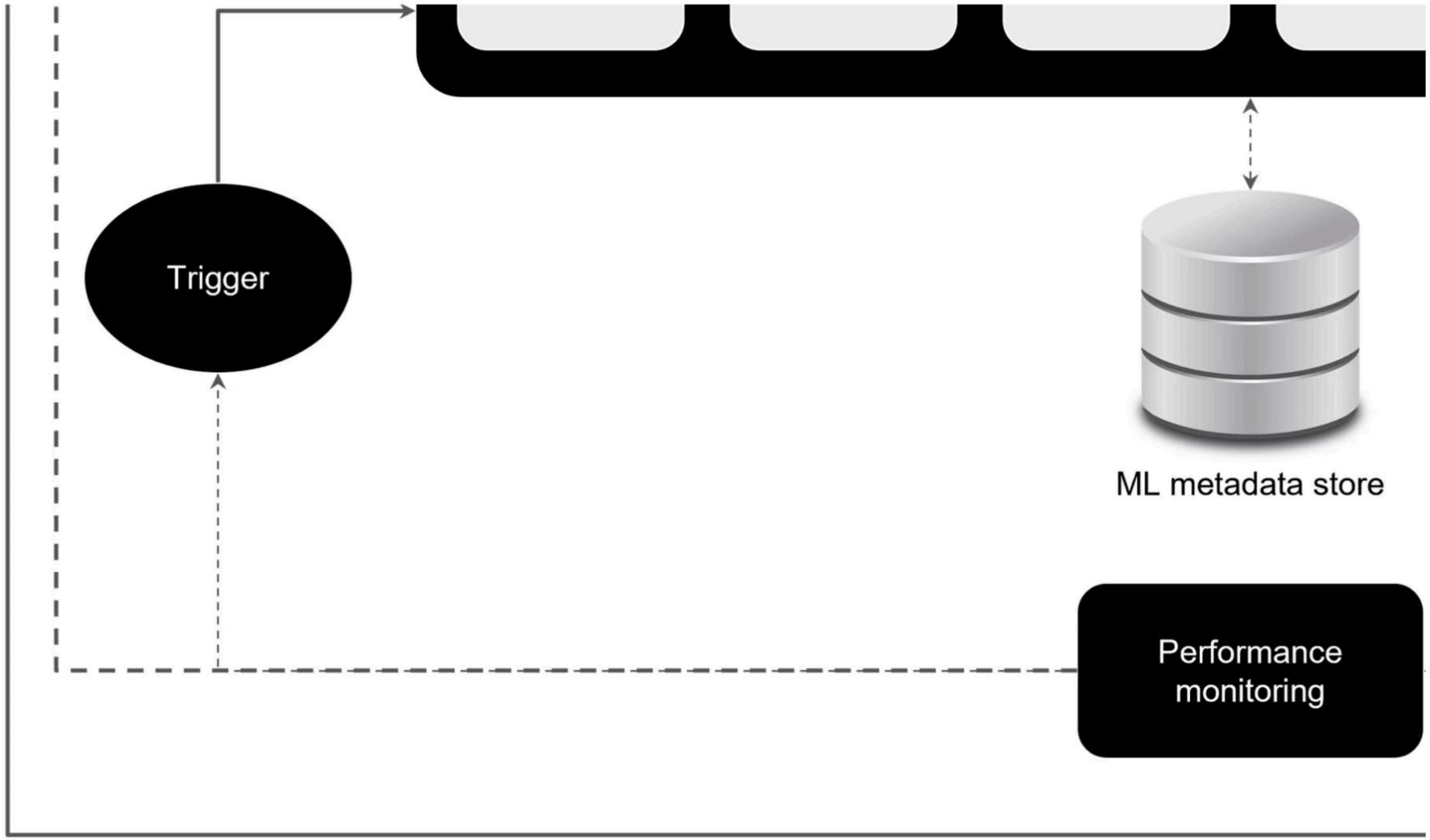
Reference architecture - Automated trigger



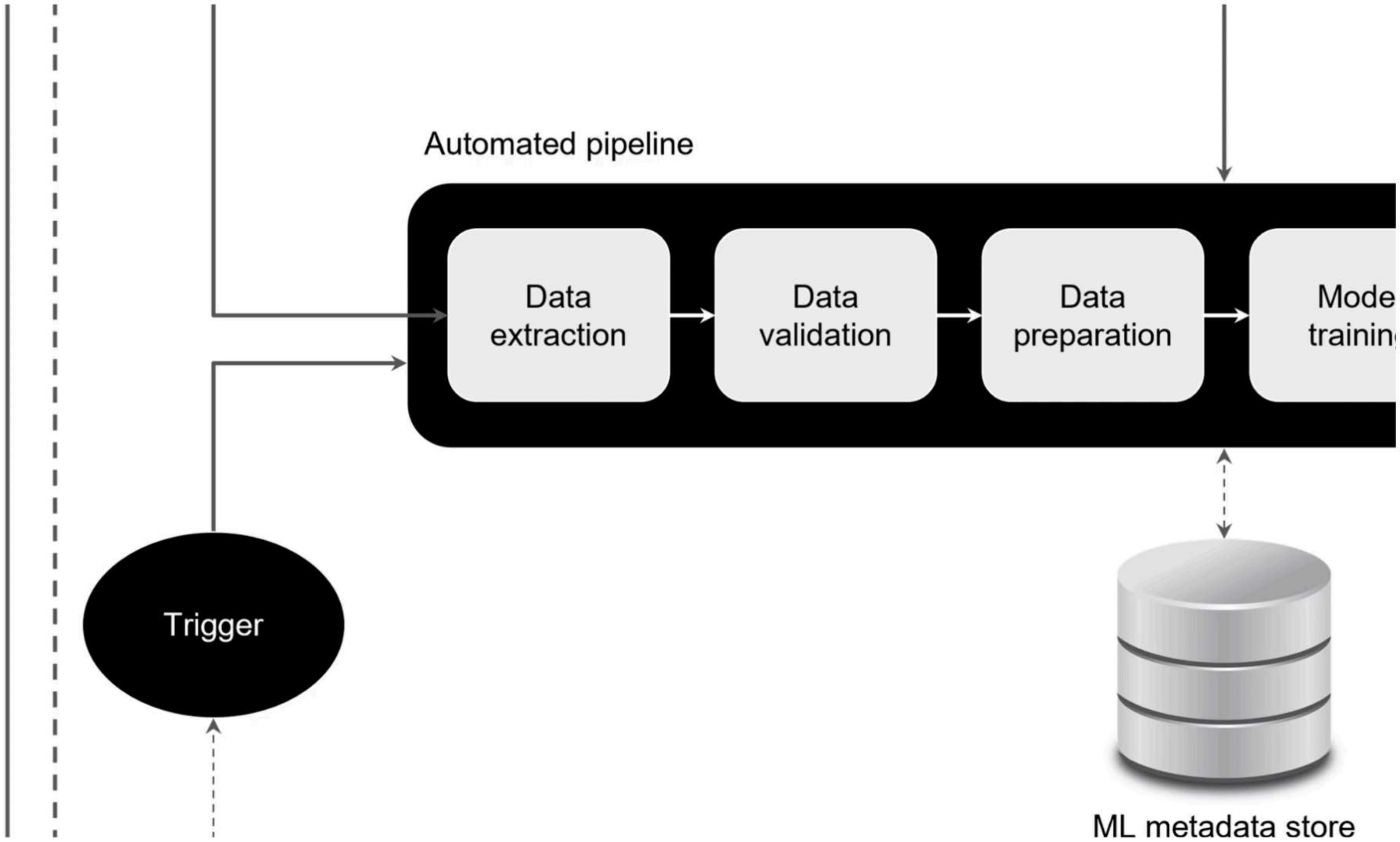
datacamp

FULLY AUTOMATED MLOPS

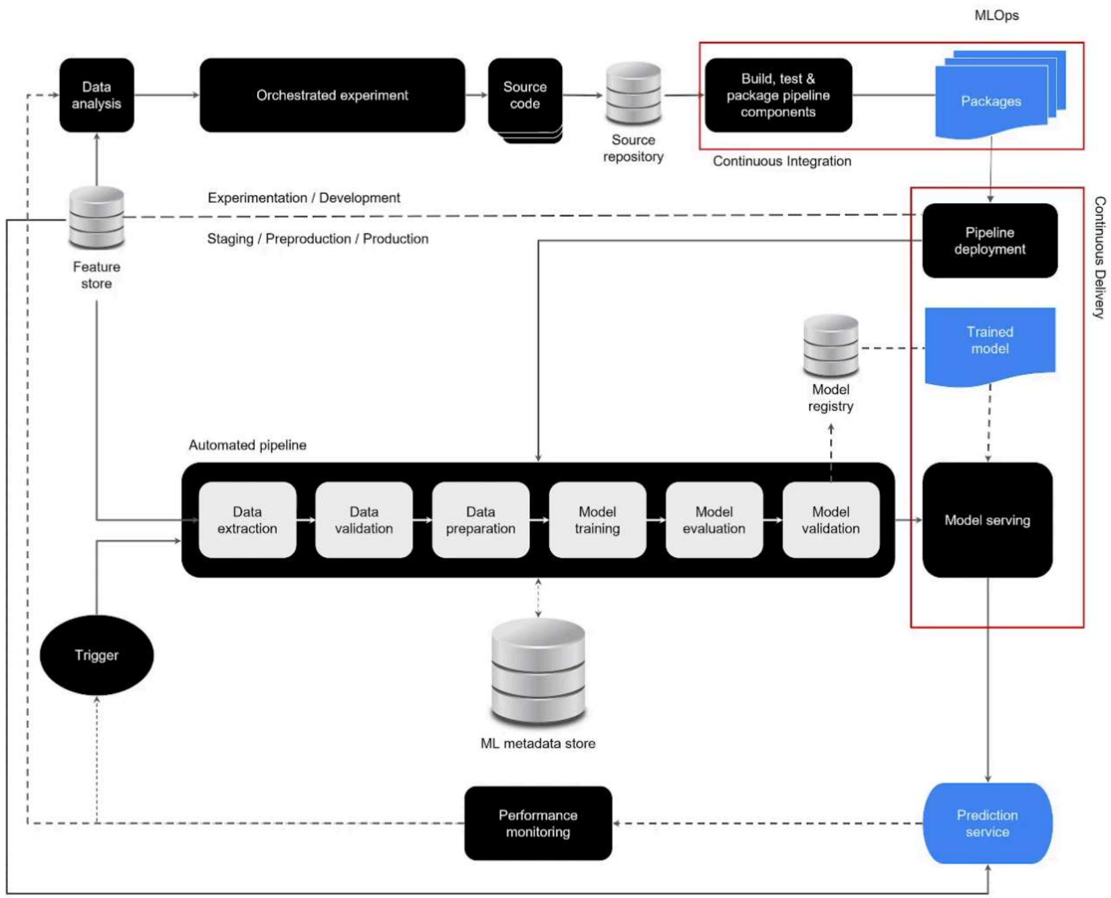
Reference architecture - Automated retraining



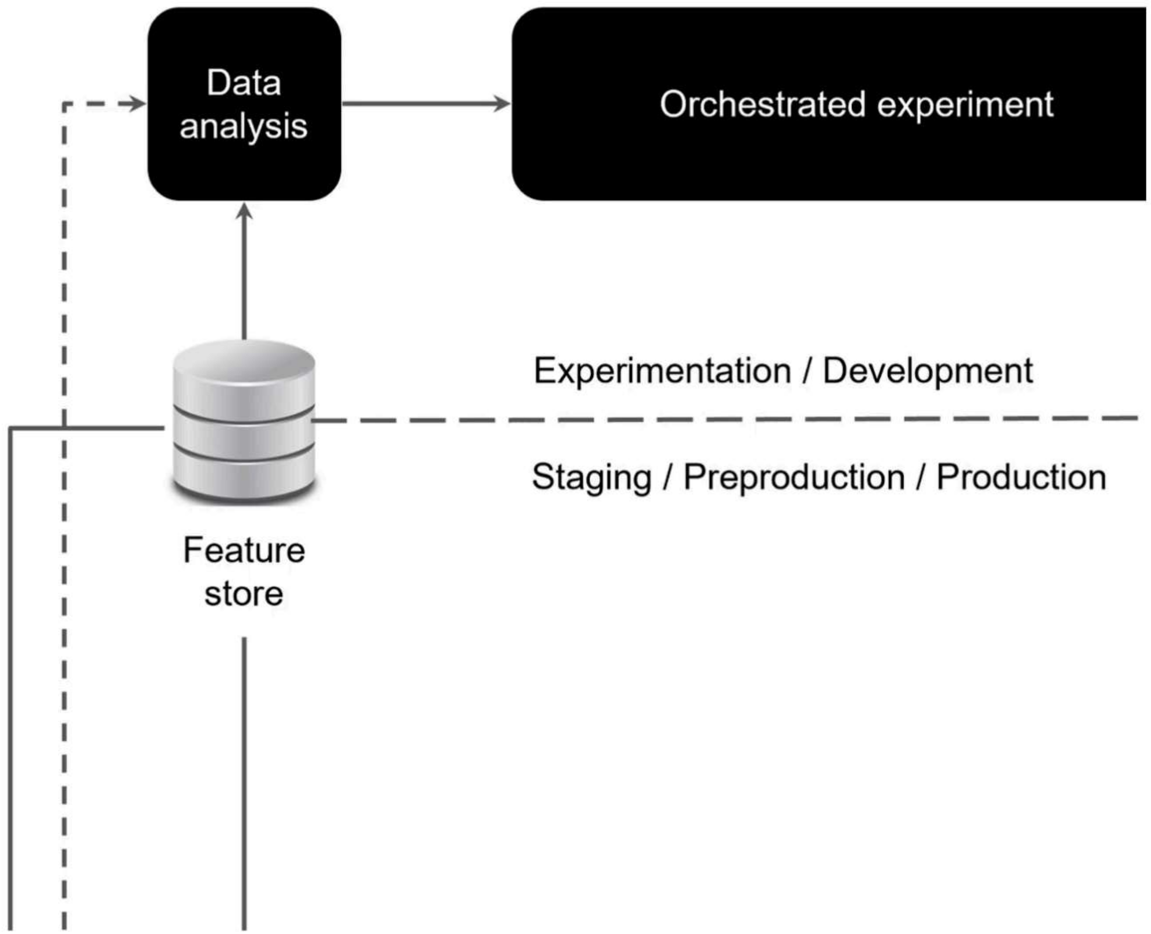
Reference architecture



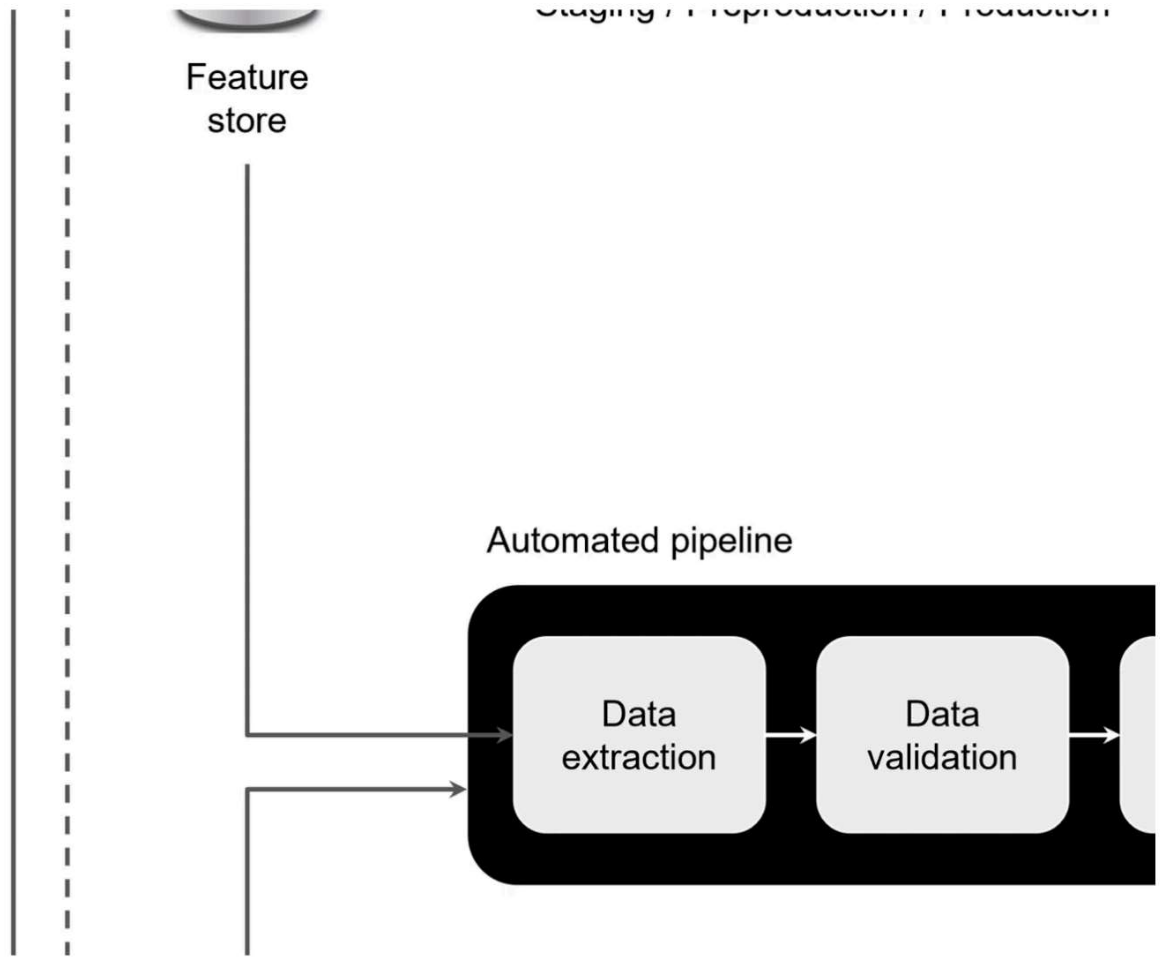
Reference architecture - The feature store



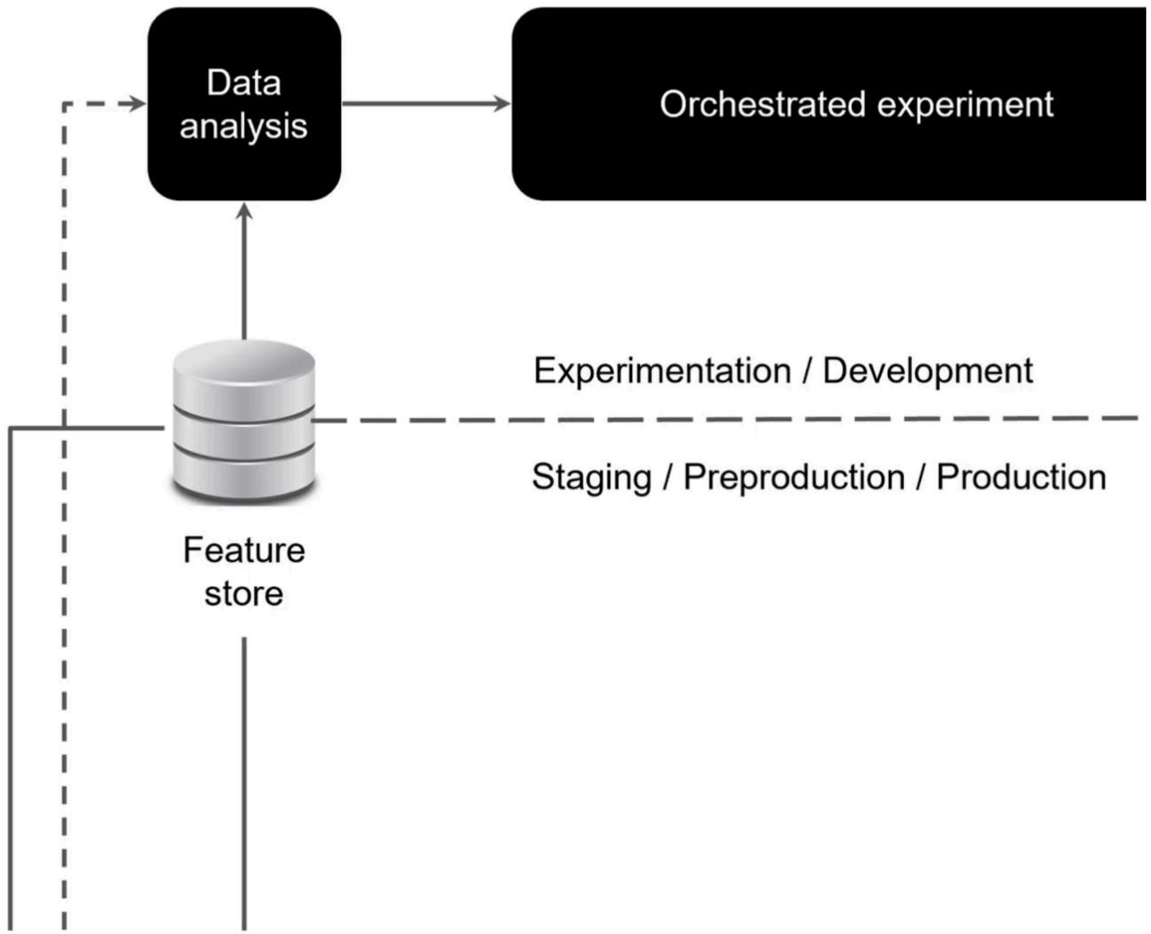
Reference architecture - The feature store



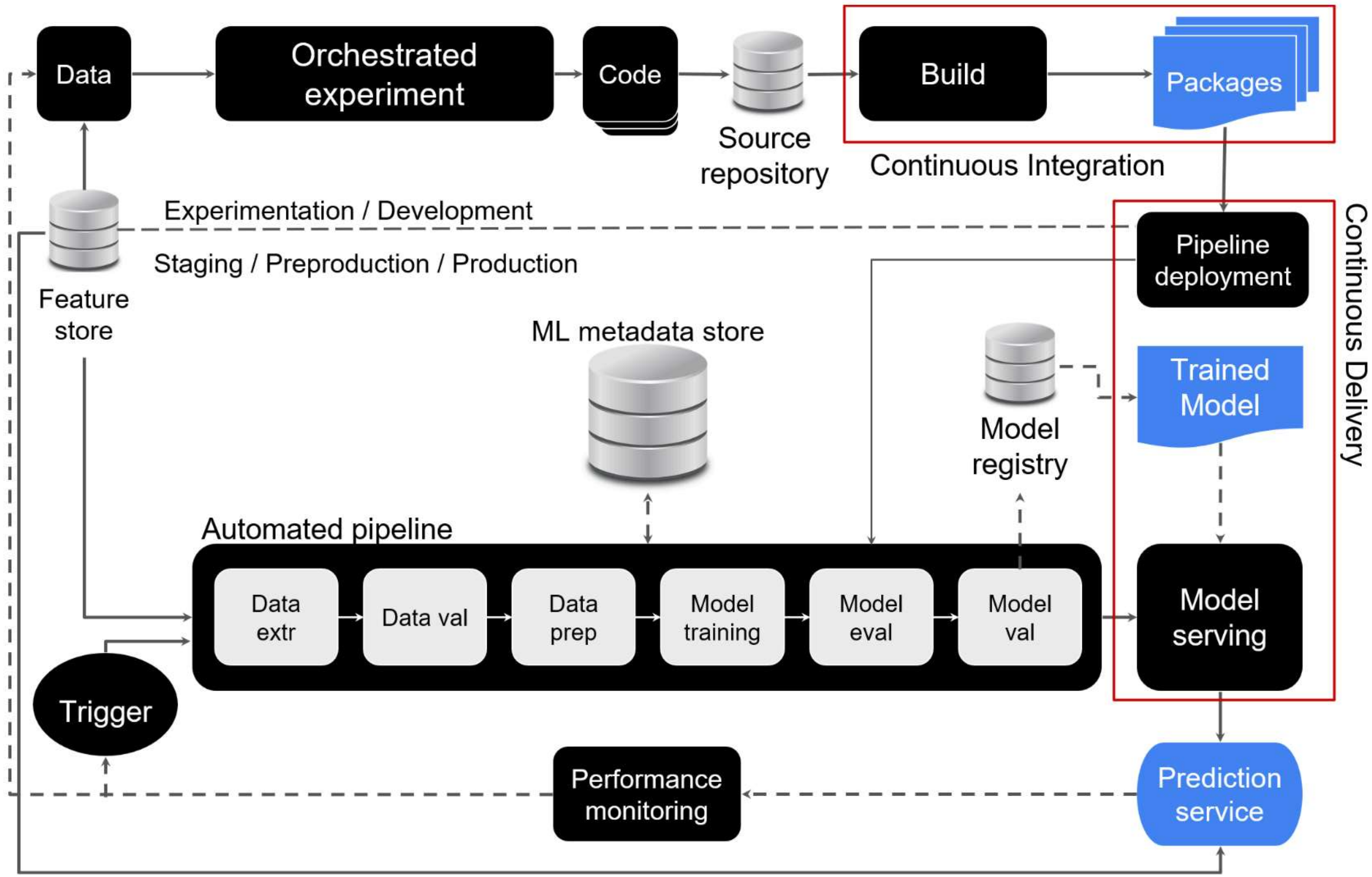
Reference architecture - The feature store



Reference architecture - The feature store



Fully automated MLOps architecture



Let's practice!

FULLY AUTOMATED MLOPS

datacamp