



# Introduction to MLOps

Christian Johnson



# Agenda

1. Personal Introduction
2. Introduction to MLOps
  - a. What is MLOps?
  - b. Key Concepts
  - c. How to Set Up an MLOps Project
    - i. Important Tools
    - ii. Minimal Project
    - iii. Intermediate Project
  - d. Additional Resources
3. Q&A

# Personal Introduction

# Personal Introduction

- Background in Computational Linguistics
- Worked as Data Analyst / Cloud Solutions Architect
- Recent experience working heavily with Language Models for various use-cases, necessitating attention to scalability and automation



**Christian Johnson**

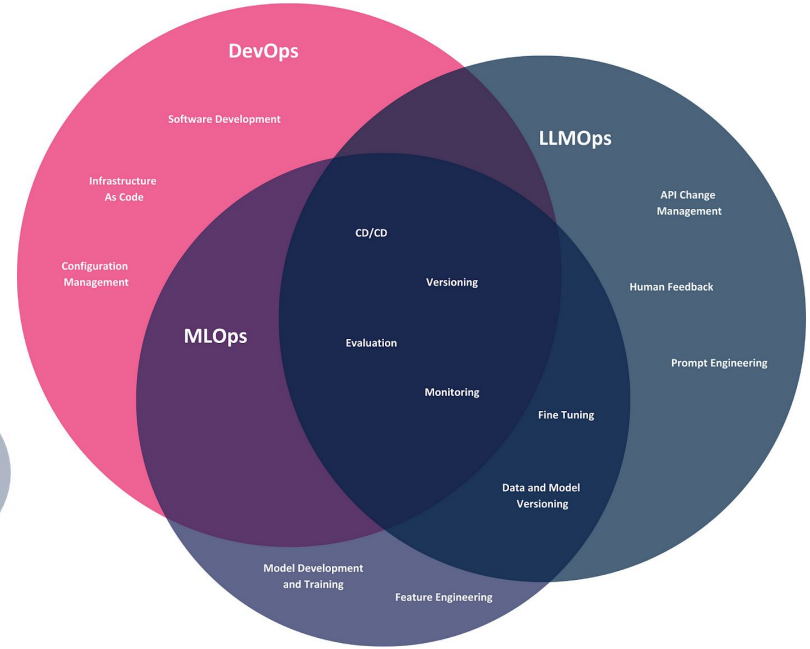
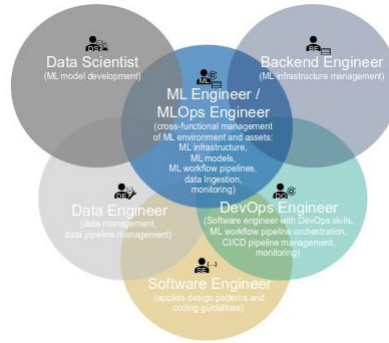
MLOps Engineer



# Introduction to MLOps

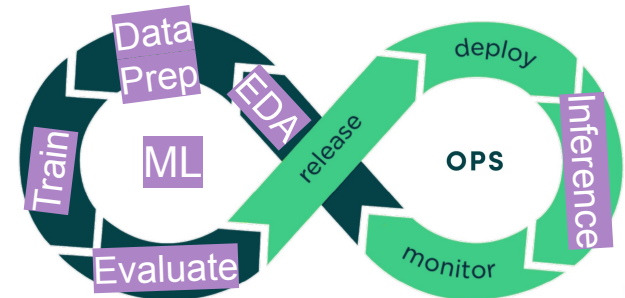
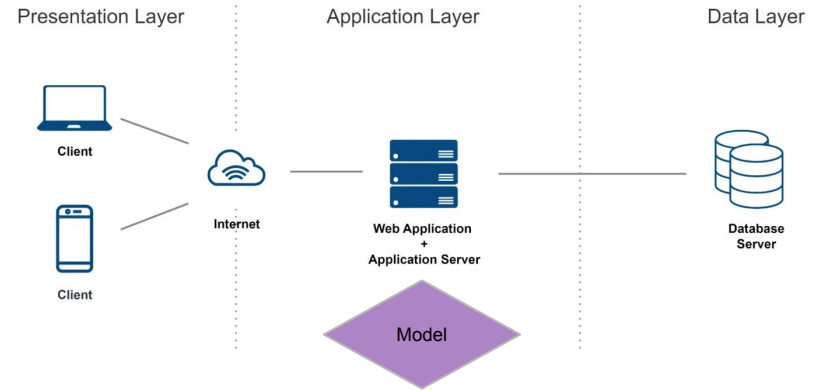
# What is MLOps?

- Relationship to DevOps and LLMOps
- Unique challenges of MLOps
  - Model versioning
  - Deployment and monitoring



# What is MLOps?

- High level concepts / system architecture
  - Experiment tracking
  - Workflow orchestration
  - Model registry
  - Model serving
  - Performance monitoring.



# Key Concepts

- Data Collection
  - Structured data collection through REST or streaming technologies may be relevant for continuous learning or dataset versioning.
- Experiment Tracking
  - System for logging experiment results and artifacts, offering additional analytics.
- Model Versioning
  - Registration and tracking of model artifacts is important for automated deployment and evaluation.
- Training Systems
  - Automated workflow orchestration leveraging compute engines or persistence layers to produce deployable models.
- Model Serving
  - Optimized serving kernels are critical especially for GPU-reliant workloads.
- Observability and Monitoring
  - Specialized monitoring and dashboarding tools may be helpful to capture the unique performance behaviours of machine learning models.





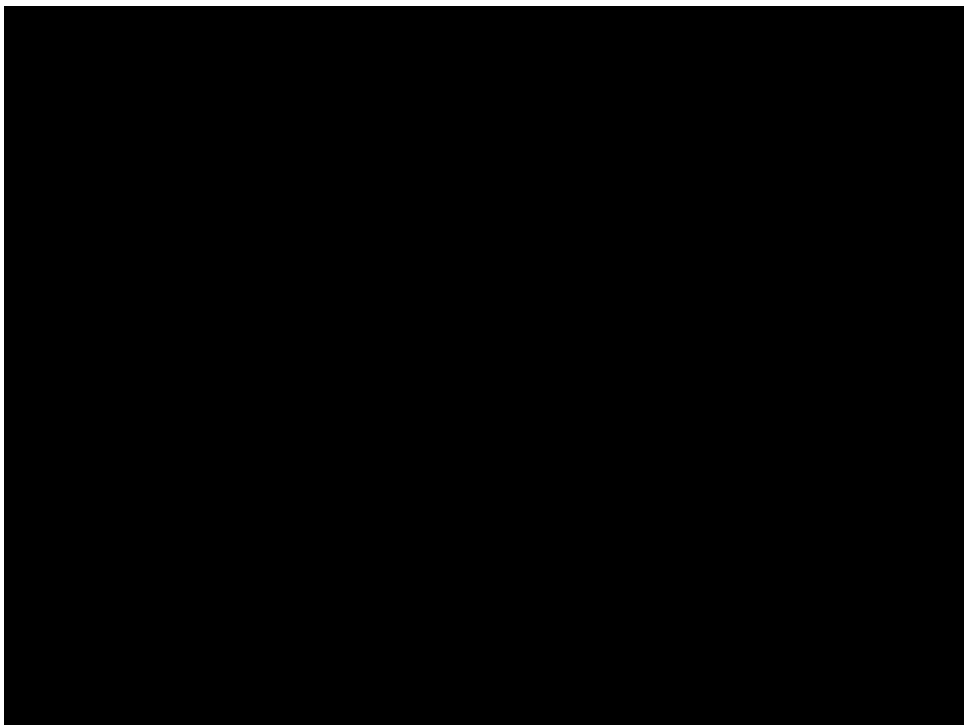
# How to Set Up an MLOps Project

- Important Tools

- MLflow
- Airflow
- MLServer

- Demo

- Minimal Project



# How to Set Up an MLOps Project

- Important Tools
  - MLflow
  - Airflow
  - MLServer
- Demo
  - Minimal Project
  - Intermediate Project: [Language Model Adapter Serving](#)

# Additional Resources

- Theory
  - [Reliable Machine Learning](#)
  - [Overview of MLOps](#)
  - [Code and Slides](#)
- Technologies
  - [MLflow](#)
  - [Airflow](#)
- Cloud Platforms
  - [Azure Machine Learning](#)
  - [Amazon Sagemaker](#)

Q&A