BADM 4830 / BAIM 4200 Advanced Business Analytics

Module 4 - MLOps – What it is , Why MLOps

Ivan Portilla
Jesus.Portilla@Colorado.edu
Portilla@gmail.com
github.com/jiportilla/giveback

Objectives of This Module

Upon completion of this module, you will understand:

MLOps - What & Why

- Definition & People of MLOps - https://ml-ops.org/
- Key MLOps Features
  - Model Development
  - Monitoring
  - Productionalization & Deployment
  - Iteration & Lifecycle
  - Governance
- Lab: Intro to MLOps

Fall 2021
BADM 4830 / BAIM 4200 Advanced Business Analytics

- This course will give students the language, knowledge, and actionable methods to work alongside technical and non-technical members of your team to create AI solutions.

- Students will explore what it means to design artificial intelligence systems as a team, guided by a clear intent and a focus on people. This course will give you the framework and tools you need to recognize responsible AI design, align your team, and work with data sources to start building AI solutions.

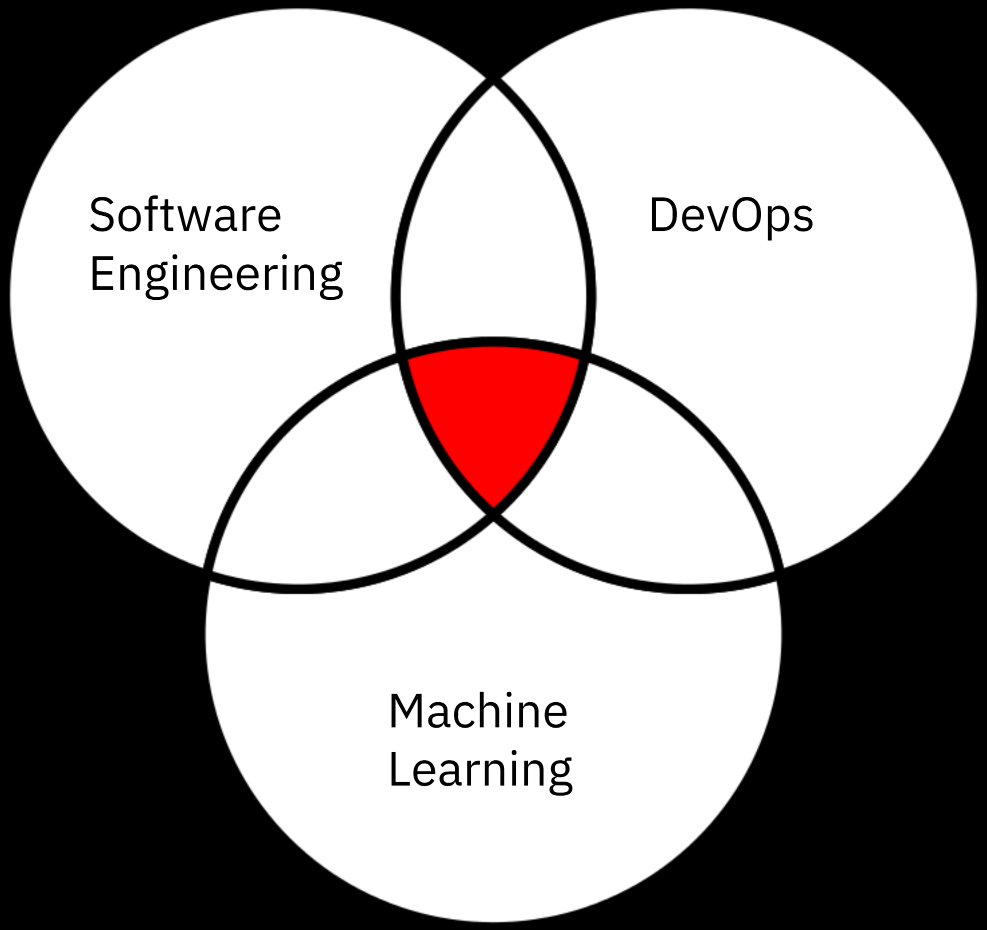
- Students will learn the tools, technology, and practices that enable cross-functional AI teams to efficiently deploy, monitor, retrain, and govern models in production systems.

Agenda

MLOps What & Why

1. MLOps definition
2. MLOps principles
3. MLOPs Lifecycle
4. AI Roles
5. Operationalize AI
6. Best practices

MLOps is ...



MLOps is … Software Engineering

How do we write software:

From punch cards to multi-user UNIX systems…

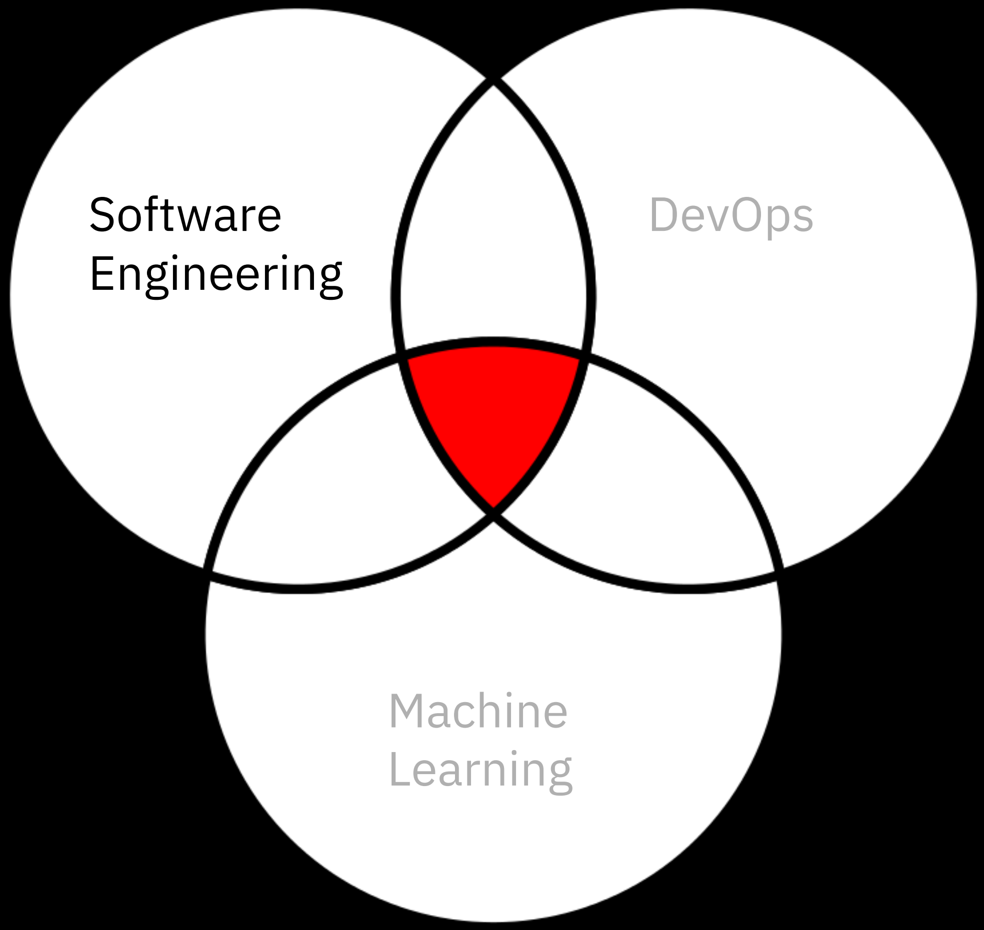
To personal computers…

To networked computers…

Emailing patch files…

Centralized version control…

Distributed version control: Git, Github



MLOps is … DevOps

How do we deploy software:

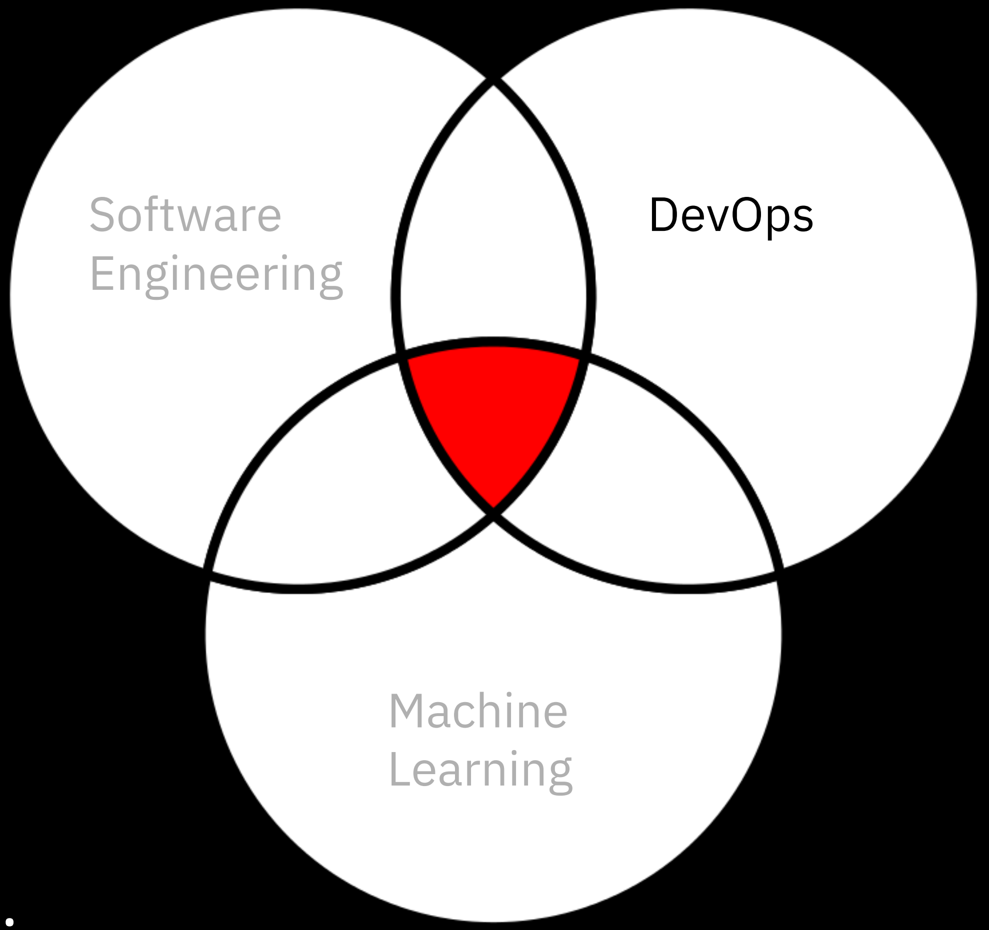
From editing code live on the server…

To building binaries & email them around…

To continuous integration…

To public cloud…

To immutable infrastructure…
Docker, Kubernetes, GitOps



MLOps is … Machine Learning

How do we use data & math to train models:

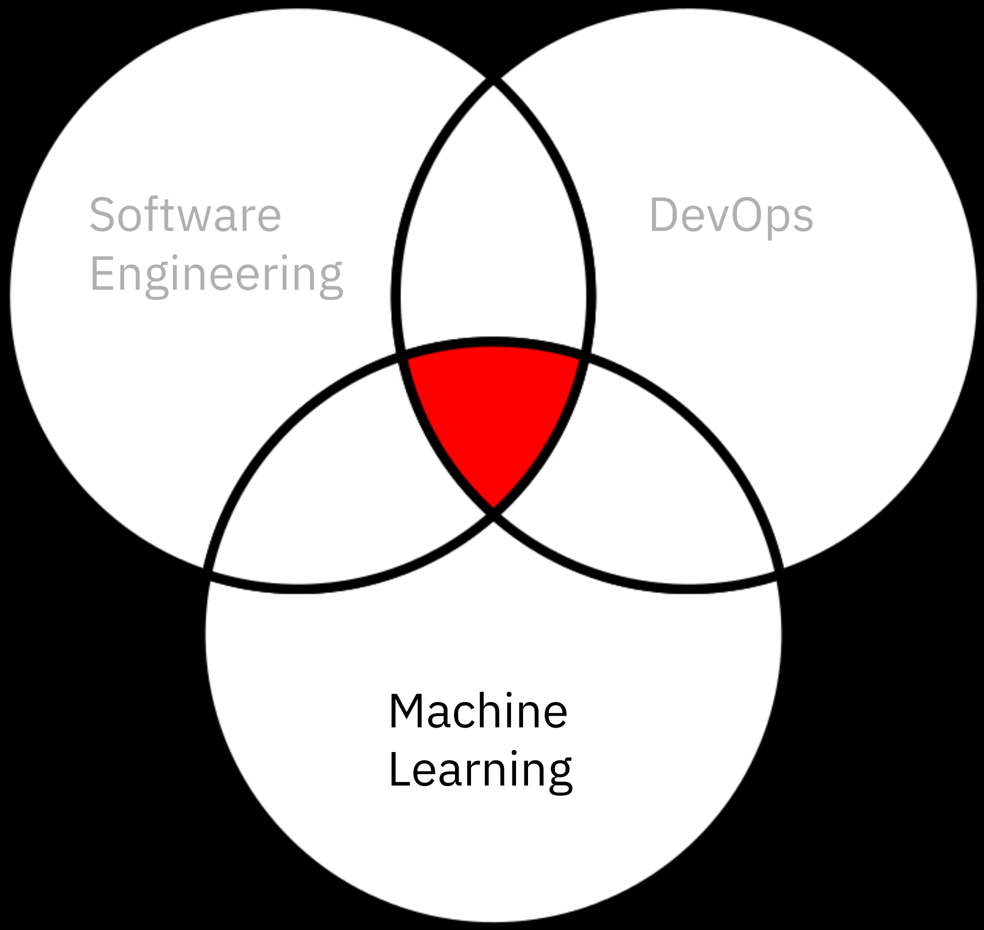
From curve fitting, linear regressions…

To Markov chains … to neural networks

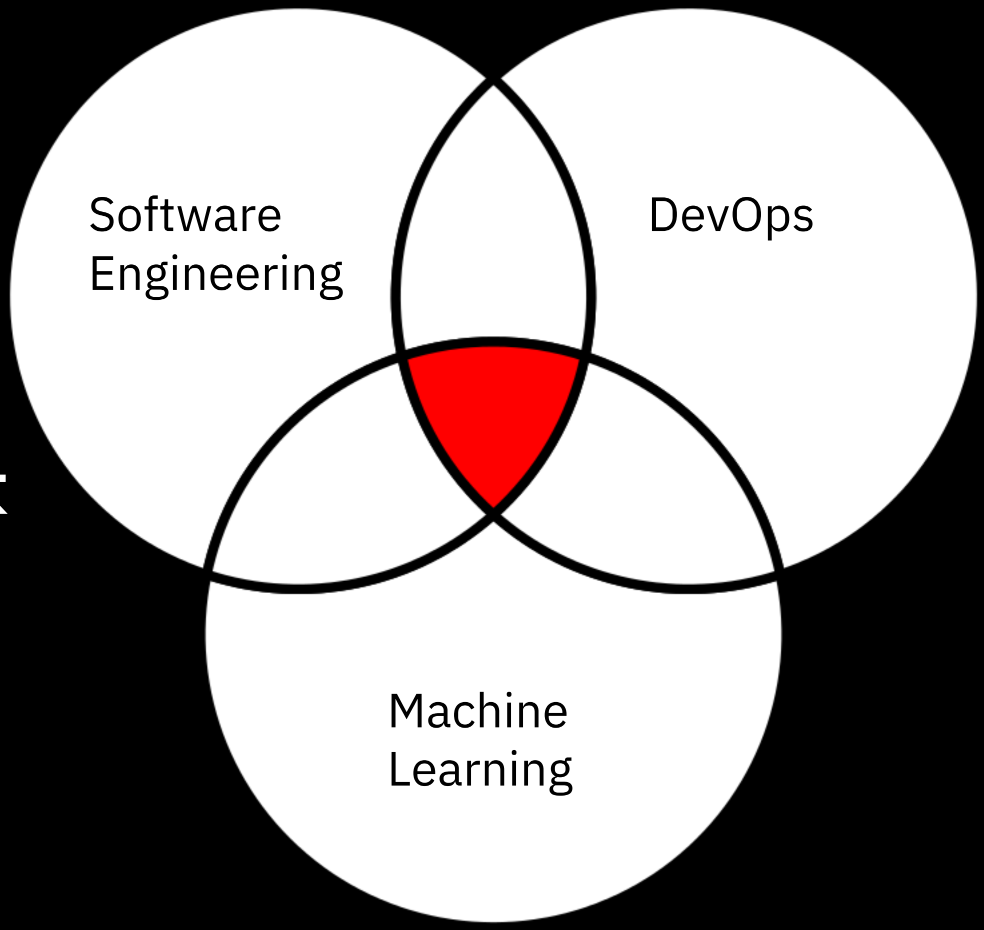
To the first AI winter…

Rediscovery of backdrop, move to data driven approach…

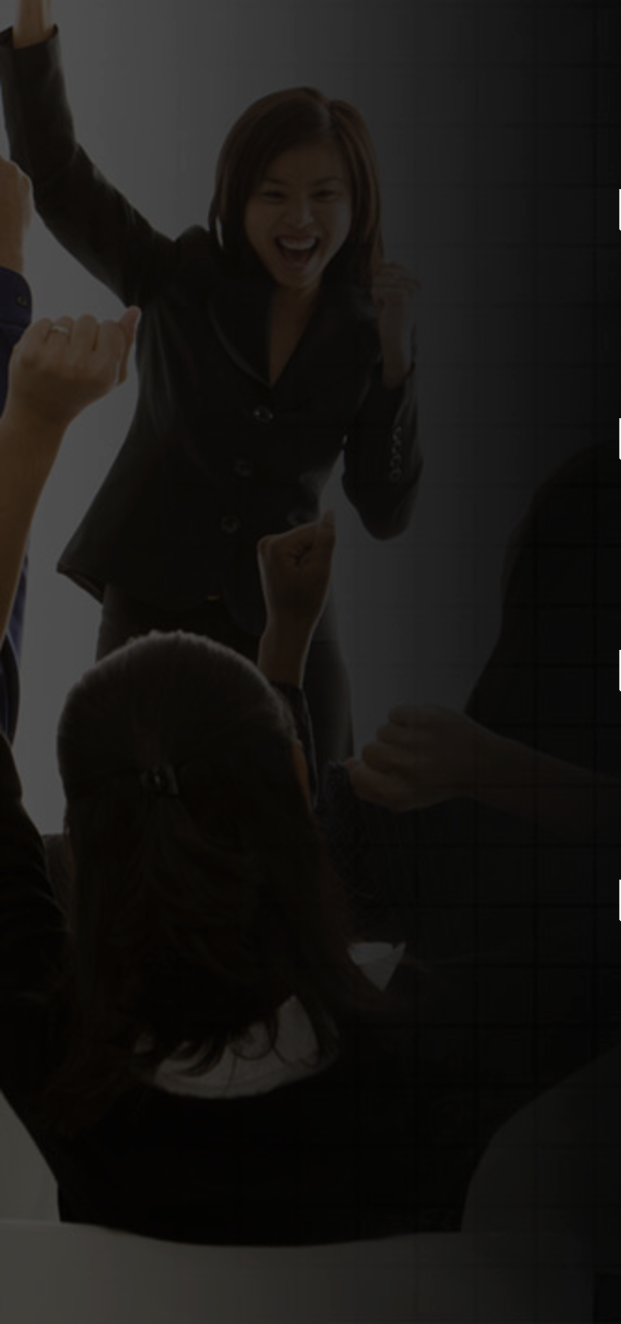
Deep learning becomes computationally feasible



MLOps is the **convergence** of Software Engineering, Machine learning & DevOps



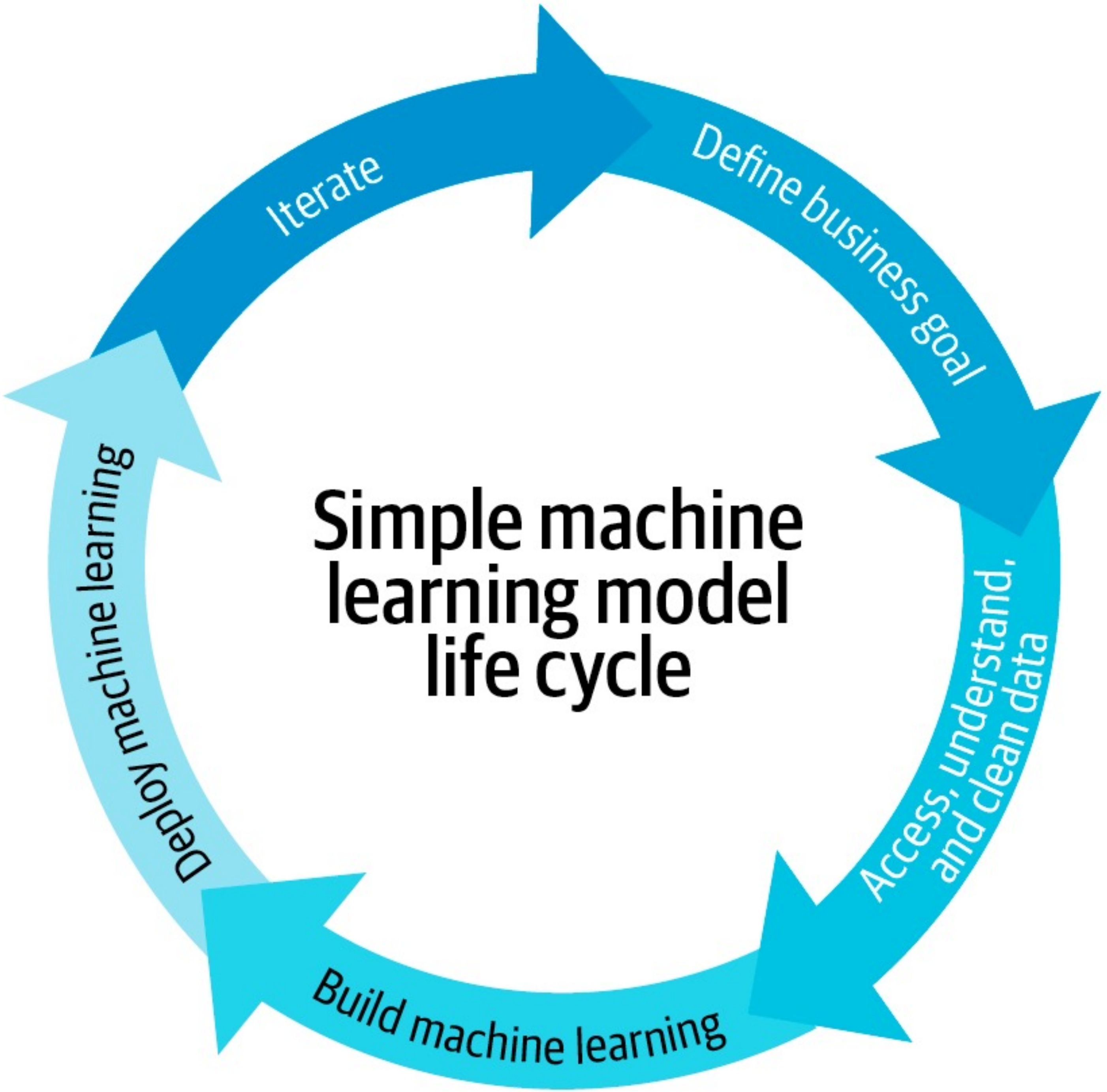
Requirements to Achieve MLOps



- ❑ Reproducible — Must be able to re-train a 9-month-old model to within few %
- ❑ Accountable — Must be able to track back from model in Production to its provenance
- ❑ Collaborative — Must be able to do asynchronous collaboration
- ❑ Continuous — Must be able to deploy automatically & monitor statistically

https://mlops.community

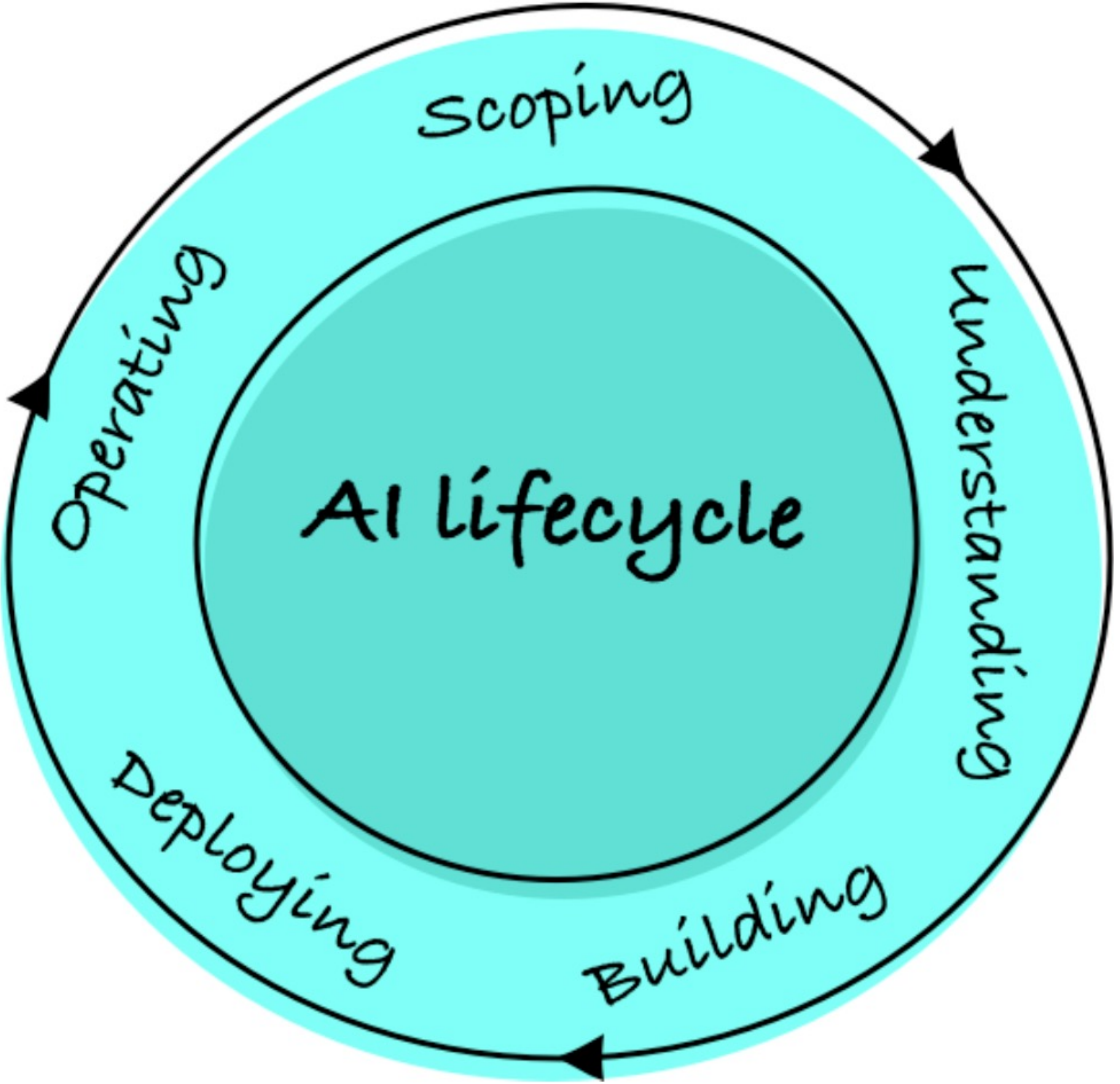
ML Lifecycle



Re-cap

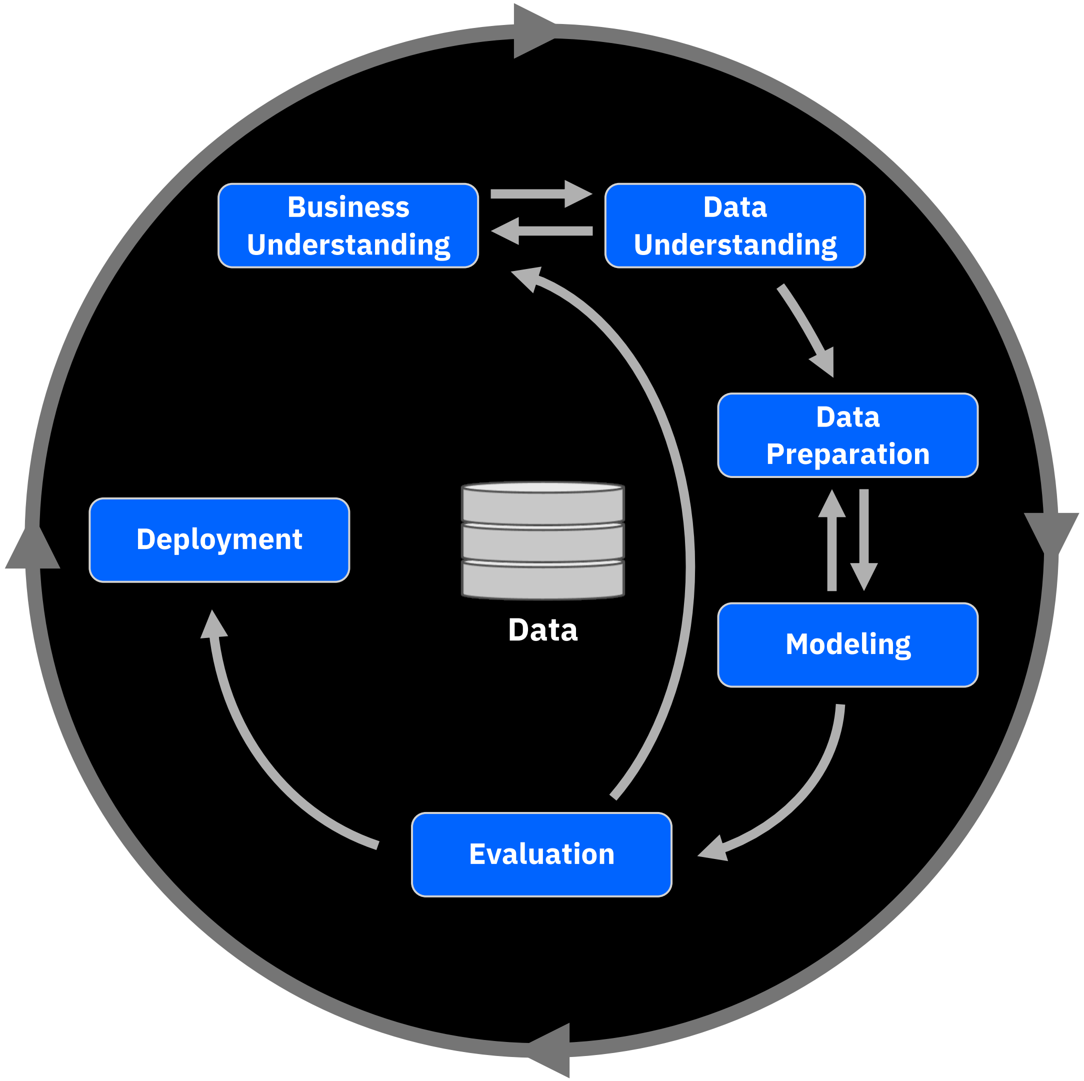
1. Data & AI Project Lifecycles
2. Data & AI Project Roles

Resources


AI lifecycle
Scoping
Understanding
Building
Deploying
Operating

Solution Development Method Approach

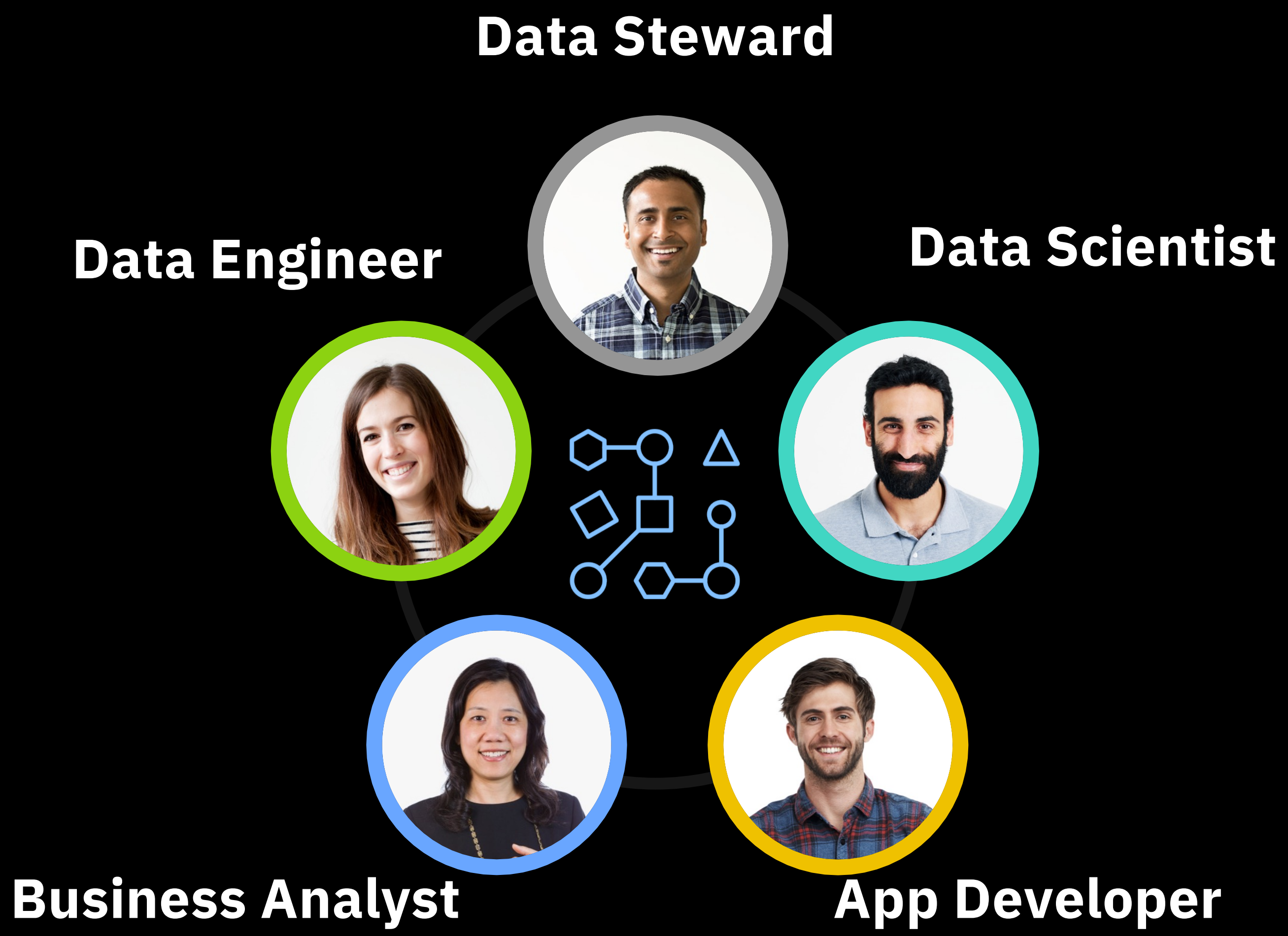
CRossIndustry Standard Process for Data Mining (CRISP-DM)



Seven steps to successful Data Mining/Predictive Analytics

1. **Define the business challenge in a precise statement**
2. **Define the data model and data requirements**
3. **Source data from all available repositories**
4. **Evaluate the data quality**
5. **Select the machine learning algorithm**
6. **Interpret the results and iterate to improve model**
7. **Deploy the model into your business**

Data & AI Project Roles


Data Steward
Data Engineer
Data Scientist
Business Analyst
App Developer

Drives governance policy effectiveness while tracking how data is used and its value to the company

Data Steward

Builds data pipelines that power dashboards and data platforms while ensuring high quality

Data Engineer

Prepares data to tease out the insights they’re looking for, without IT involvement

Data Scientist

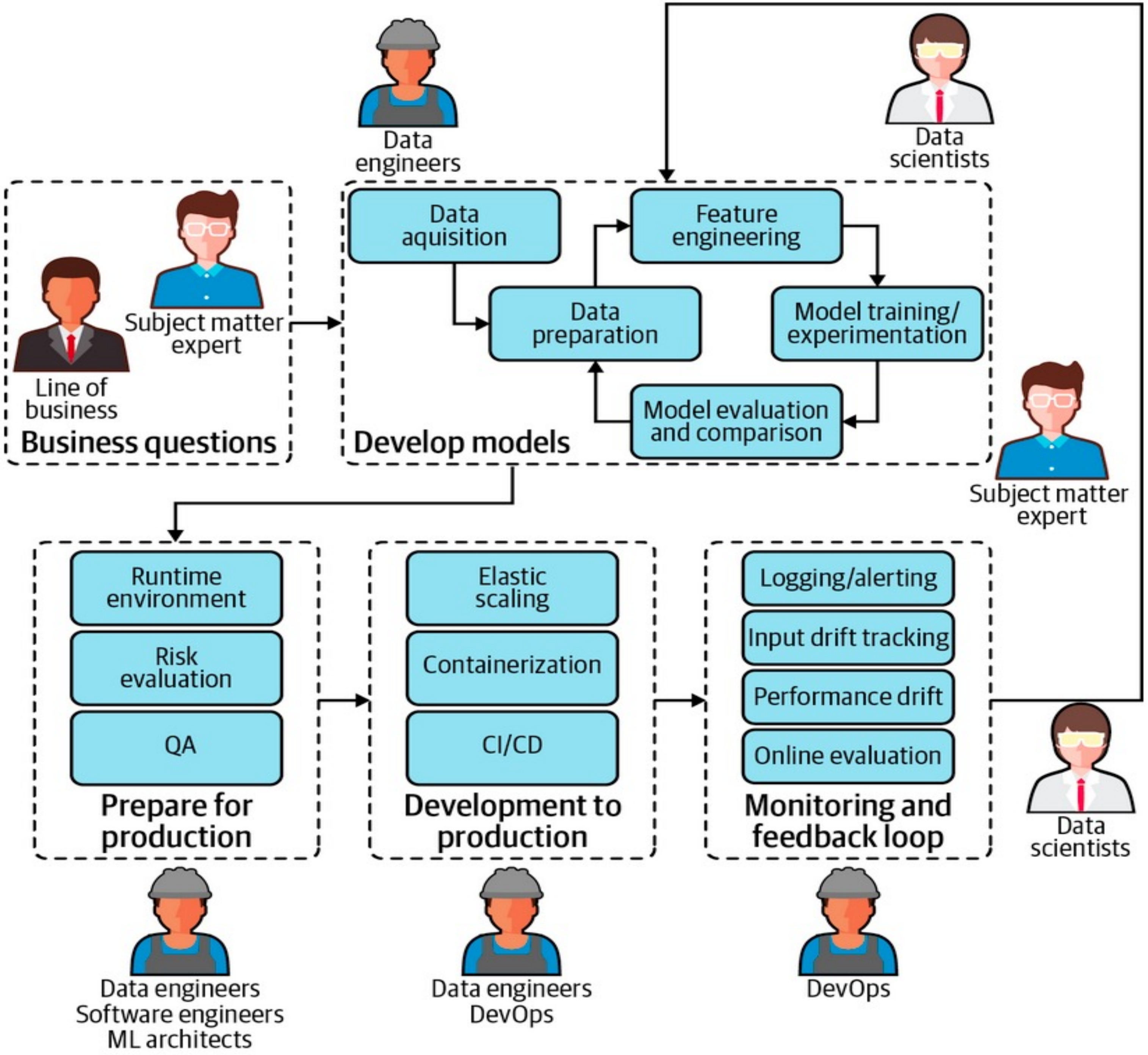
Business Analyst

Works with data to apply insights to the business strategy

App Developer

Makes insights immediately actionable and adds intelligence to apps in straightforward manner

Enterprise MLOps Lifecycle



AI Factory

1. Case study
2. Operationalizing AI stages
3. Best practices

CASE STUDY

IBM helps Walmart set up an enterprise-wide framework for **Trusted AI**

CHALLENGE

Help Walmart ensure fairness and trust is established for their AI/ML models, to accelerate efforts around digital citizenship and social responsibility.

SOLUTION

Platform:
OpenScale (Cloud Pak for Data) on Azure, GCP

Process:
MLOps-Trust (part of AI Factory)

People:
DSE & Expert Labs teams

Quickly and *consistently* **detect and counteract bias in Walmart's AI/ML models** and support business decisions.

IBM's Watson OpenScale on Cloud Pak for Data monitors health and fairness in Walmart's hiring system. This enables transparency in the decision making, both for **uniformity of decisions** and to **"explain" any specific decision.** The monitoring and explanations allow Walmart to inspect, understand, and correct decisions, maintaining its commitments as a fair employer, and fairness for each of its candidates and new employees.

Walmart is now expanding beyond the hiring use case to an enterprise-wide **Trusted AI** paradigm for all AI/ML models

Operationalizing AI involves challenges across personas

C-Level Exec

"We want to get down from 6-9 months for building & deploying AI to 6-8 weeks. Being able to get to market quickly is very important"

Operational Challenges

Data Science Lead

"Our teams use a variety of ML tools; we need consistent model management"

Ops Team Lead

"Need to integrate with our Continuous Integration and Continuous Delivery (CI/CD) process"

Risk Team Lead

"How do we consistently trust what the models recommend?"

Business View

Business Executive

"How do we apply AI to transform our business?"

"How do we go from ideas to market quickly?"

Approach

C-Level Exec

"We need a strategy and repeatable process for AI at scale and at speed"

Creating an AI/ML model for a proof-of-concept looks like this...

Build AI assets

Connect data

Analyze data

Prepare data

Build models

Train models

Visualize & evaluate models

Test for Bias, Fairness

Data engineer

Data scientist

But this is not the complete AI lifecycle!

Stages in Operationalizing AI

Continuous integration / continuous delivery pipeline

Scope AI project(s)

Explore, prioritize, select use cases – feasibility vs impact

Detail selected use case(s) – KPIs, data, workflow, success criteria

Technical design: tools, infrastructure, data details, approach

Business
Data Scientist
IT/ Ops
Data provider

DataOps

All: Define sources & needs

Provider: Set up catalog for discovery, lineage, curation, access rules

Steward: Create data policies & access

Consumer: Find, understand, add, explore, review, share

Data provider
Data steward

Build AI assets

Connect data

Analyze data

Prepare data

Build models

Train models

Visualize & evaluate models

Test for Bias, Fairness

Data engineer
Data scientist

ML Ops – Deploy

Review, 3rd party oversight, unit tests, validation, approve deployable version

Deploy model to ML runtime engine

Monitor & evaluate model execution

Manage against thresholds

Validation reports

Software engineer
IT/ Ops
Data scientist

Manage and Trust

Configure monitoring
Configure integration with other systems

Monitor and manage:
1. Quality
2. Performance
3. Custom metrics
4. Bias/fairness
5. Accuracy & Data Drift

Explain on demand

Risk & Compliance
Business

How can we trust an AI-infused recommendation?

Trustworthiness presents new challenges to operationalizing AI

Bias

Training data and AI models may be biased.

Are privileged groups at a systematic advantage compared to unprivileged groups?

Quality

Models need to perform well across the AI/ML lifecycle.

Are relevant performance metrics being monitored over time?

Drift

Changes in input data cause model to make inaccurate decisions.

Training data may not include data ranges or combinations seen in real life

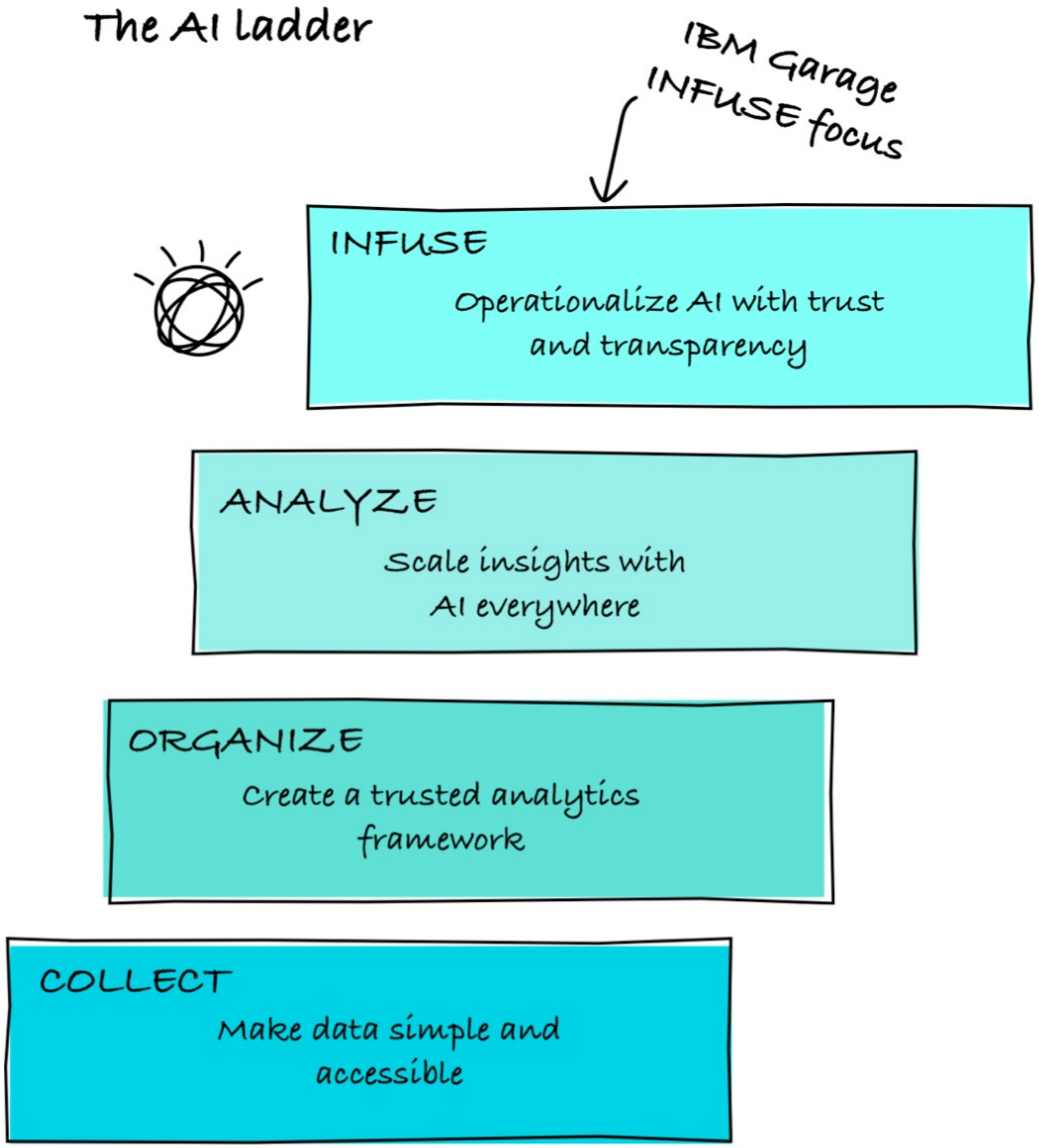
Explainability

Traditional statistical models are simpler to interpret and explain.

At what point would the outcome have been different?

Best practices

Resources


The AI ladder
IBM Garage INFUSE focus
INFUSE
Operationalize AI with trust and transparency
ANALYZE
Scale insights with AI everywhere
ORGANIZE
Create a trusted analytics framework
COLLECT
Make data simple and accessible

Process: Best Practices Throughout the Life Cycle

Continuous Improvement

Scope

- Strategic planning provides the foundation
- Structured design thinking defines business need
- Structured exercises translate business need into specific technical task and shape understanding steps

Understand

- Flesh out details of contributing data sources according to model
- Flesh out details of technical task and approach based on experiments with clear validation criteria

Build

- Develop in agile sprints
- Continuously validate and adapt approach, data, and integration plan
- Develop for data science outcome *and* for integration into broader systems

Deploy

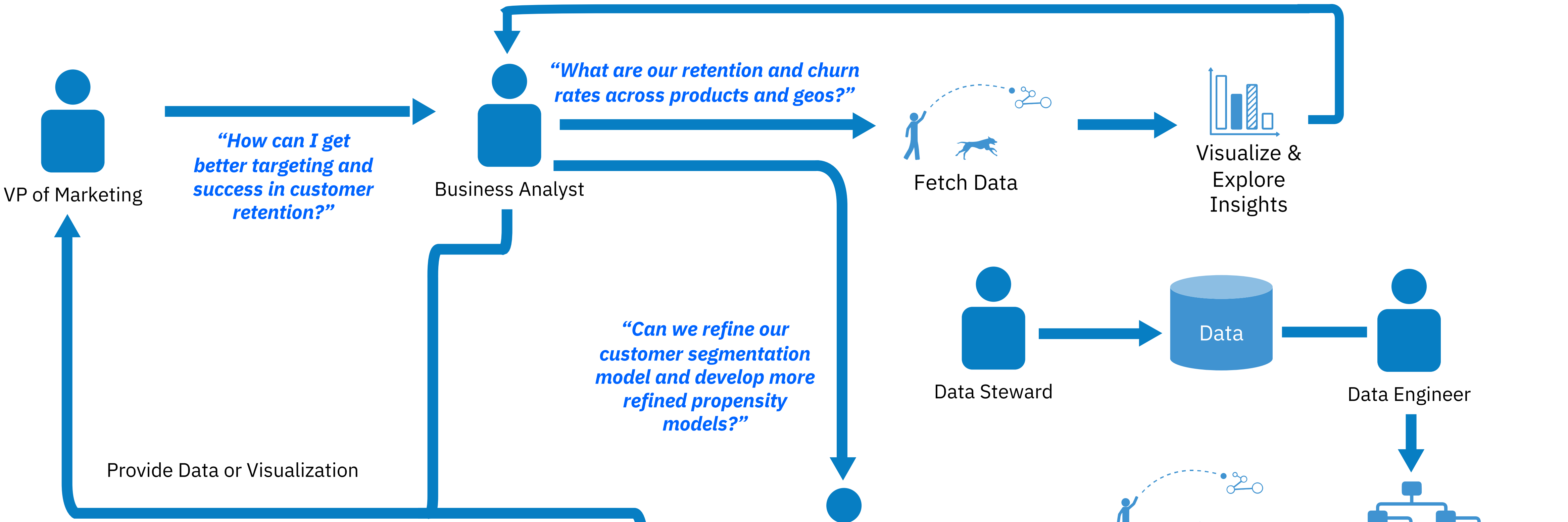
- Integrate into software applications and business workflow
- Automate, integrate with CI/CD systems
- Deploy based on QA of both data science outcomes and full system

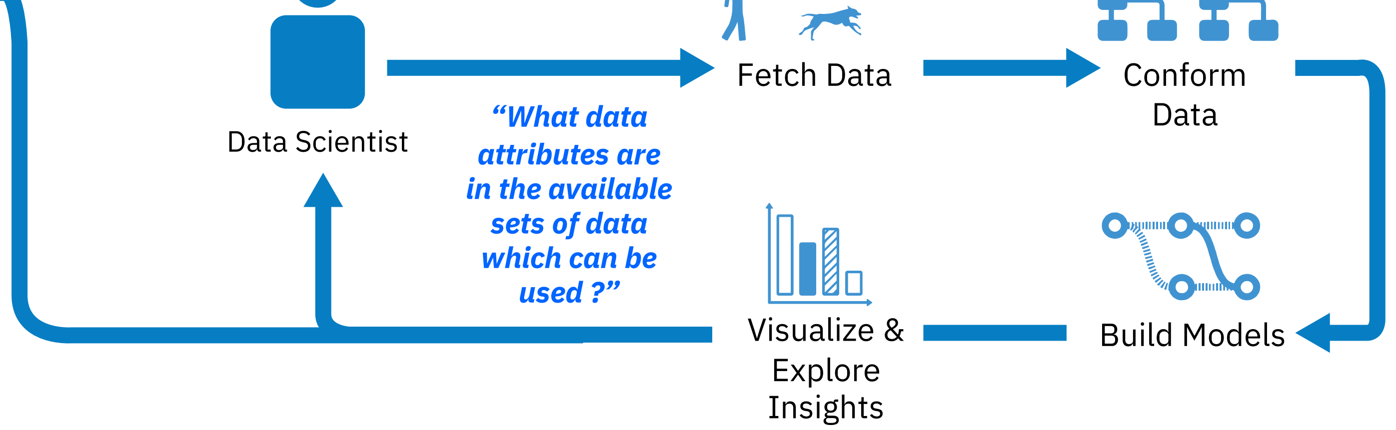
Trust

- Monitor for accuracy and drift to identify the need to update the model(s)
- Demystify AI models by explaining key results
- Identify and correct for unfair outcomes

Demos - Lab

Sample business problem: Churn Analysis





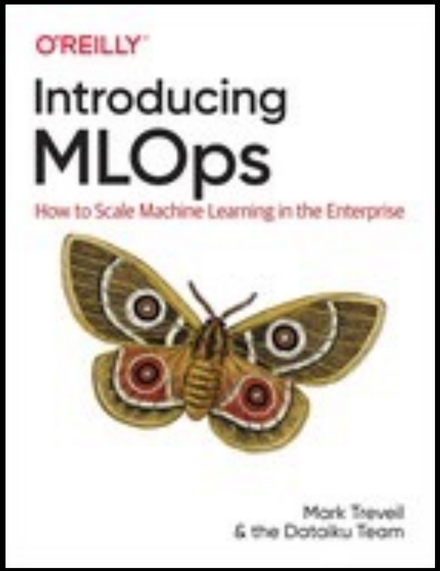
- Rapid data evaluation and extraction of necessary attributes to support analysis
- Discovery through shop for data
- Rapid hypothesis testing
- Iterative analysis and model building through discovery and collaboration
- Deploy scoring models into existing campaign management processes

Q&A

Ivan Portilla

ivanp@us.ibm.com

@iportilla



https://medium.com/inside-machine-learning/ai-ops-managing-the-end-to-end-lifecycle-of-ai-3606a59591b0

Resources

https://www.ibm.com/artificial-intelligence/ethics
https://developer.ibm.com/technologies/artificial-intelligence/
https://developer.ibm.com/technologies/artificial-intelligence/learningpaths/

https://github.com/jiportilla/INSA