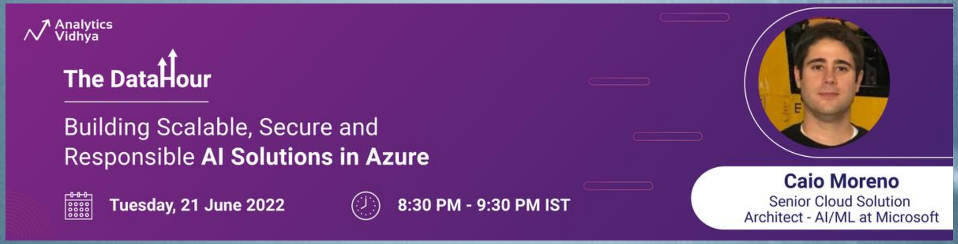
Microsoft

Building Scalable, Secure and Responsible AI Solutions in Azure

Caio Moreno
Senior Cloud Solution Architect



Agenda

- Introduction to Machine Learning
- How to become an Azure Data Scientist?
- Building a Data Science Team
- Data and AI Strategy & The AI/ML Journey
- Building Scalable, Secure and Responsible AI Solutions in Azure
  - ML Ops (Machine Learning Ops)
  - Responsible AI (Artificial Intelligence)

About me

Caio Moreno is a **Senior Cloud Solution Architect at Microsoft**, responsible for helping Microsoft empower every person and organization on the planet to achieve more using Data and AI. He has experience in artificial intelligence, machine learning, big data, IoT, distributed systems, analytics, streaming, business intelligence, data integration and visualization. He is also a PhD. Candidate at the Complutense University of Madrid. He enjoys travelling and all kinds of sports. He lives in London with his wife and 3 daughters.

https://www.linkedin.com/in/caiomsouza/

https://linktr.ee/caiomoreno

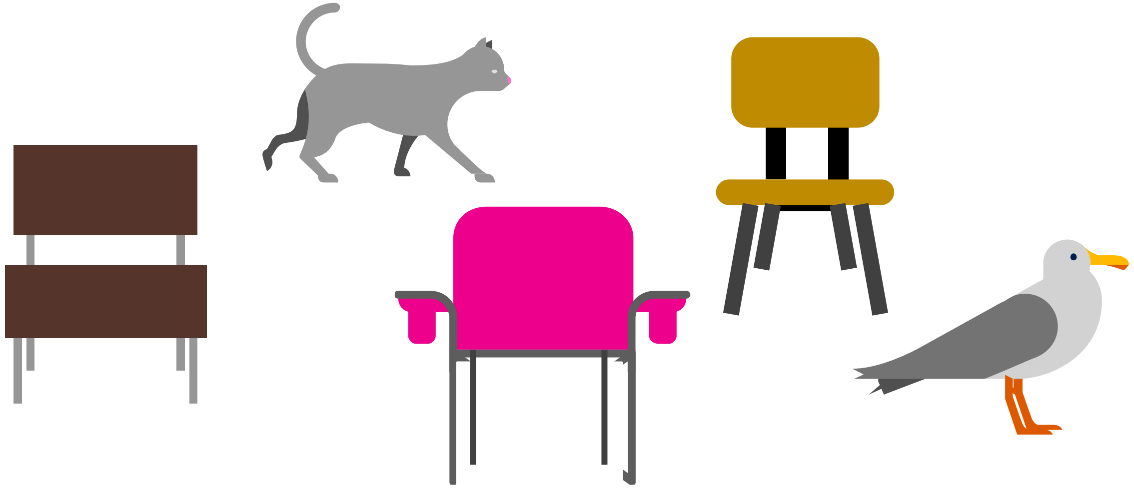
AI isn’t just another piece of technology. It could be one of the **most fundamental pieces** of technology the human race has ever created.

Satya Nadella

Introduction to Machine Learning

Machine Learning

Grew out of artificial intelligence within computer science. Aims to learn from and make predictions on data.



Is this a chair?

Preliminary Concepts

Input vs output variables

- Input synonyms: feature, attribute, predictor variable
- Output synonyms: target, class, label

Type of variables

- Categorical / qualitative / nominal

  Distinct category of values that is not meaningful to rank / measure difference, e.g. gender, color
- Numeric / quantitative

  Values that can be ranked and compared in a meaningful way, e.g. weight, price

Supervised vs. Unsupervised

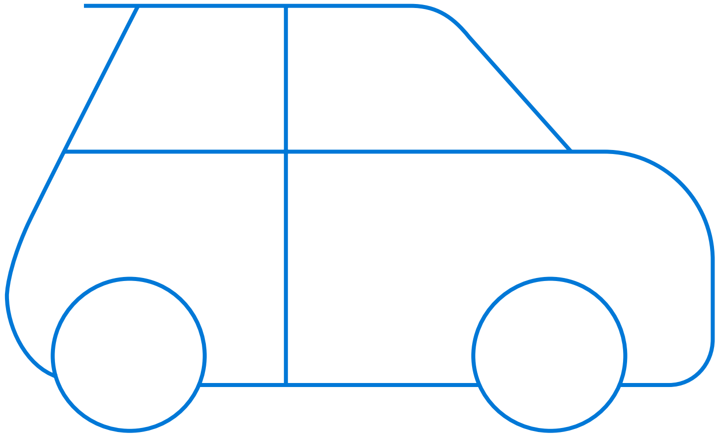
- Supervised: Learn from data with ground truth (the true answer, label)
  - Classification problems such as spam email detection
  - Regression problems such as income prediction
- Unsupervised: Learn without ground truth
  - Clustering problems such as user clustering
- Semi-supervised: Learn with some labelled and some unlabeled data

And More…

- Reinforcement Learning: Learn to take actions to achieve a goal and maximize (cumulative) reward
- Recommender Systems: A system that predicts how much a user will like an item
  - Simplified: Choose top n that user would like the most

Building your own AI models

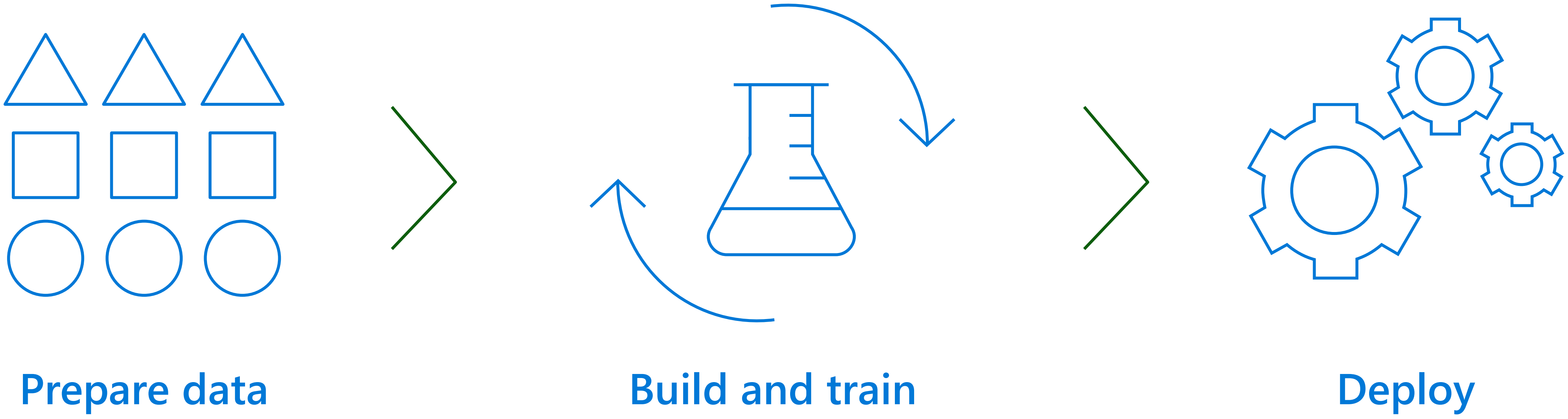
Transforming Data into Intelligence



Q: How much is this car worth?

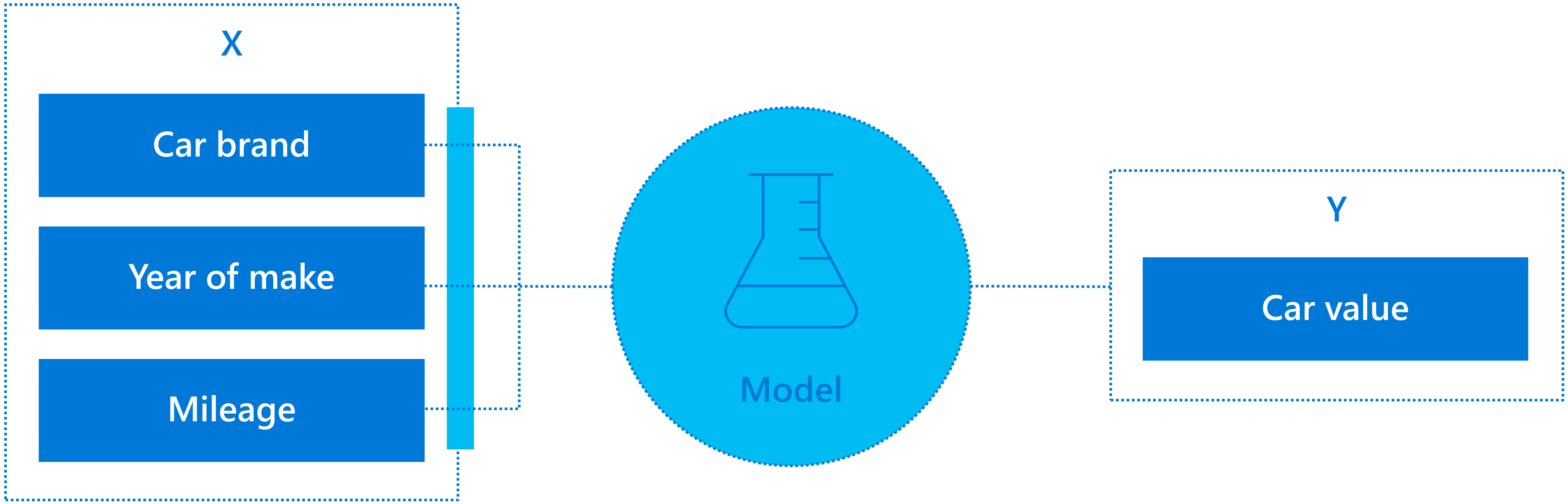
Building your own AI models

Transforming data into intelligence



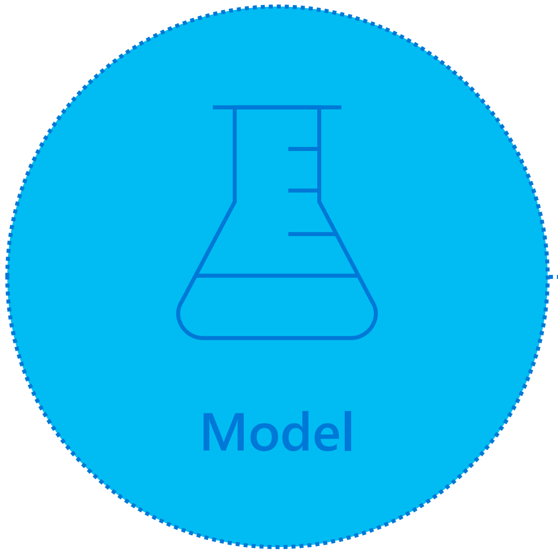
Building your own AI models

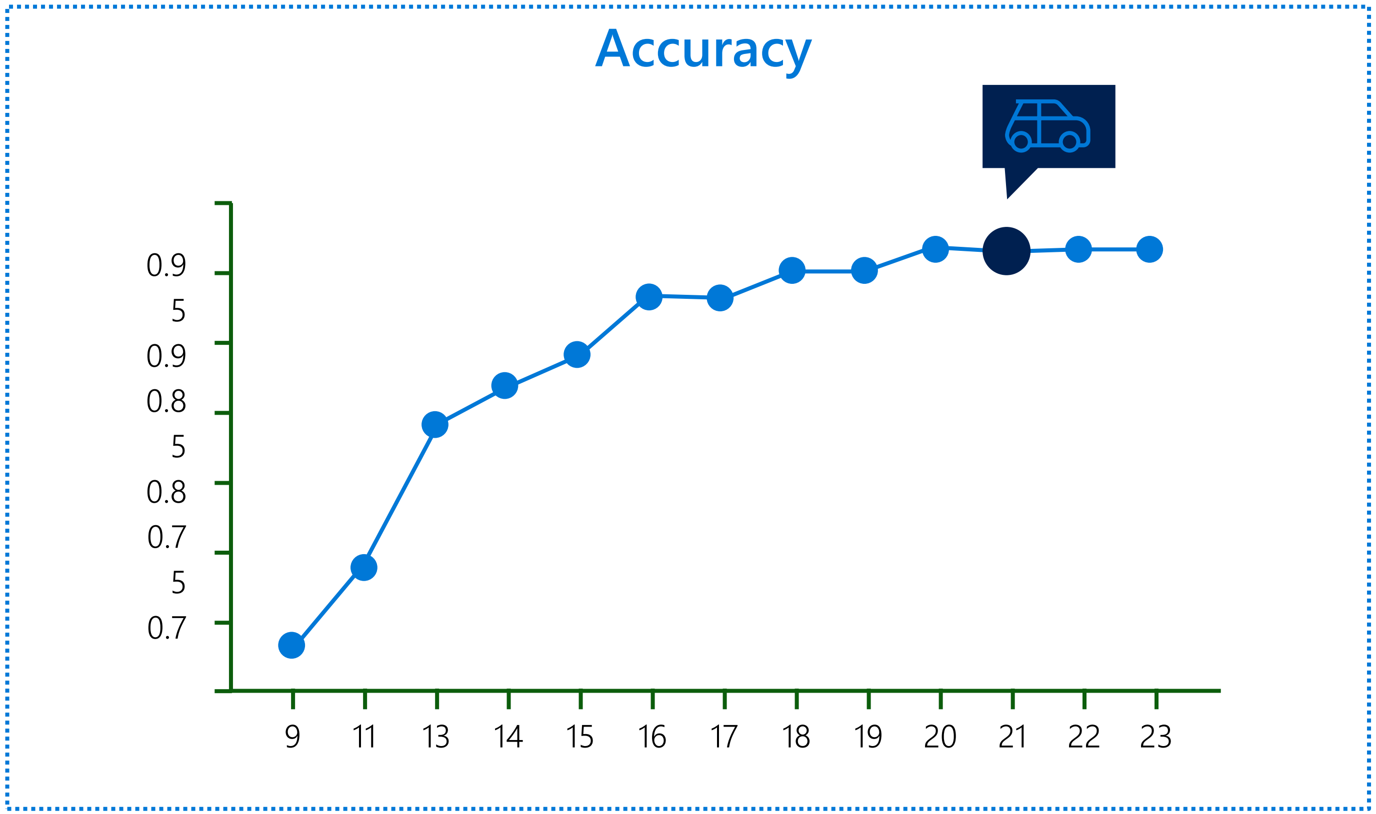
Step: Build and Train



Building your own AI models

Step: Build and train

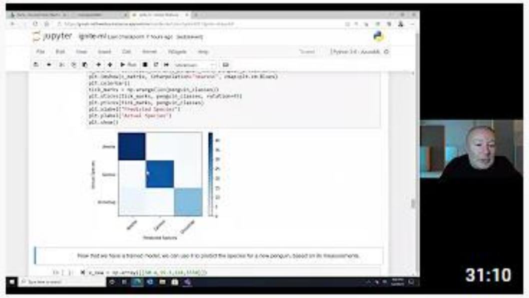




How to become an Azure Data Scientist?

How to become an Azure Data Scientist?

How do I start?



How to be a Data Scientist on Azure | LRN236
https://www.youtube.com/watch?v=n6OOowy7wAc

How to become an Azure Data Scientist
https://www.youtube.com/watch?v=iYBKTkDBB8g

Building a Data Science team

Data Science Team

Isabel
Team Lead

Own the AI Data Product(s). Manage the team.

Caio
AI Solution Architect

Design the AI technical solution that meets business requirements, and operations architecture for those solutions.

Chen
Data Scientist

Frame business problem, exploratory data analysis, develop machine learning models, communicate outcomes.

Roger
Data Engineer

Operate and maintain the data stack, pull data from different sources, data integration and prep, set up data pipelines.

Ana
Data Visualization Expert

Create data visualizations to help leaders make business decisions.

Helena
Data Architect

Design databases for mission-critical line-of-business apps. Designing and implementing data security.

Matt
ML Ops Engineer

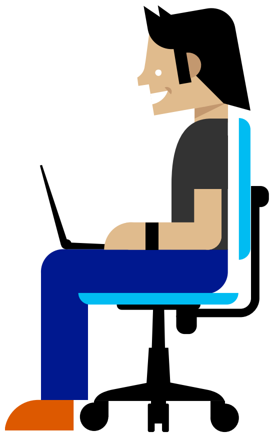
Design and build end to end ML Ops pipelines to deploy AI Solutions to production.

David
App Developer

Build apps that use data and analytic models.

https://docs.microsoft.com/en-us/azure/architecture/data-science-process/overview

General Roles in a Data Science Project

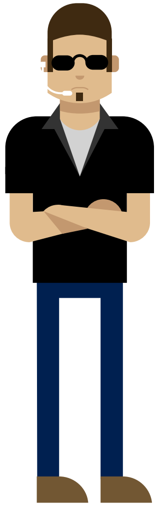


3. Data Engineer

Acquire, transform and stage/store data sources

Orchestrate data acquisition and flow through to organizational data stores

Create ,schedule and provision foundational orchestration services

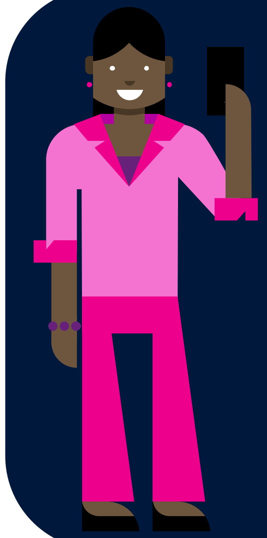


1. Data Architect

Architect the whole data platform

Design the data governance strategy

Ensure data security and privacy



4. Machine Learning Engineer

Implement the MLOps pipeline

Deploy machine learning solution into production

Optimize solutions for performance and scalability

2. Data Scientist

Understand and help refine/shape the analytical question to answer

Craft approaches for new analytic technical foundations (i.e., Statistical Testing, Machine Learning, Data Lakes, etc.)

Conduct analytical modeling, testing, evaluation and interpretation

Assist in monitoring and adjusting for model performance over time

Data & AI Strategy

Data and AI Strategy

It is proposed a new formal framework for companies that want to adapt to a Big Data MO strategy.

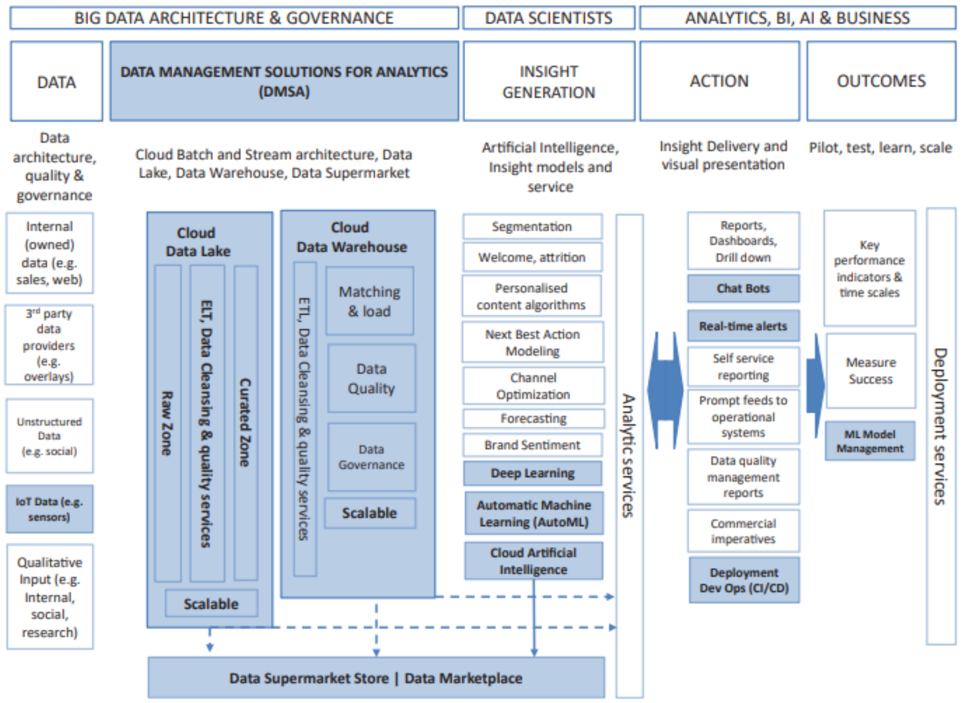


Fig. 3. A conceptual Big Data MO strategy architecture.

Special Issue on Use Cases of Artificial Intelligence, Digital Marketing and Neuroscience

Data and Artificial Intelligence Strategy: A Conceptual Enterprise Big Data Cloud Architecture to Enable Market-Oriented Organisations

Caio Moreno[1], Ramón Alberto Carrasco[1], Enrique Herrera-Viedma[2]*

[1] Department of Management and Marketing, Complutense University, Madrid 28223 (Spain)
[2] Andalusian Research Institute in Data Science and Computational Intelligence, University of Granada, Granada 18071 (Spain)

unir LA UNIVERSIDAD EN INTERNET

Received 11 April 2019 | Accepted 25 April 2019 | Published 22 June 2019

Abstract

Market-Oriented companies are committed to understanding both the needs of their customers, and the capabilities and plans of their competitors through the processes of acquiring and evaluating market information in a systematic and anticipatory manner. On the other hand, most companies in the last years have defined that one of their main strategic objectives for the next years is to become a truly data-driven organisation in the current Big Data context. They are willing to invest heavily in Data and Artificial Intelligence Strategy and build enterprise data platforms that will enable this Market-Oriented vision. In this paper, it is presented an Artificial Intelligence Cloud Architecture capable to help global companies to move from the use of data from descriptive to prescriptive and leveraging existing cloud services to deliver true Market-Oriented in a much shorter time (compared with traditional approaches).

Keywords

Market-oriented Organisations, Big Data, Cloud Architecture, Artificial Intelligence Strategy, Data Supermarket.

DOI: 10.9781/ijimai.2019.06.003

https://doi.org/10.9781/ijimai.2019.06.003

Moreno, C., González, R. A. C., & Viedma, E. H. (2019). Data and artificial intelligence strategy: A conceptual enterprise big data cloud architecture to enable market-oriented organisations. IJIMAI, 5(6), 7-14.

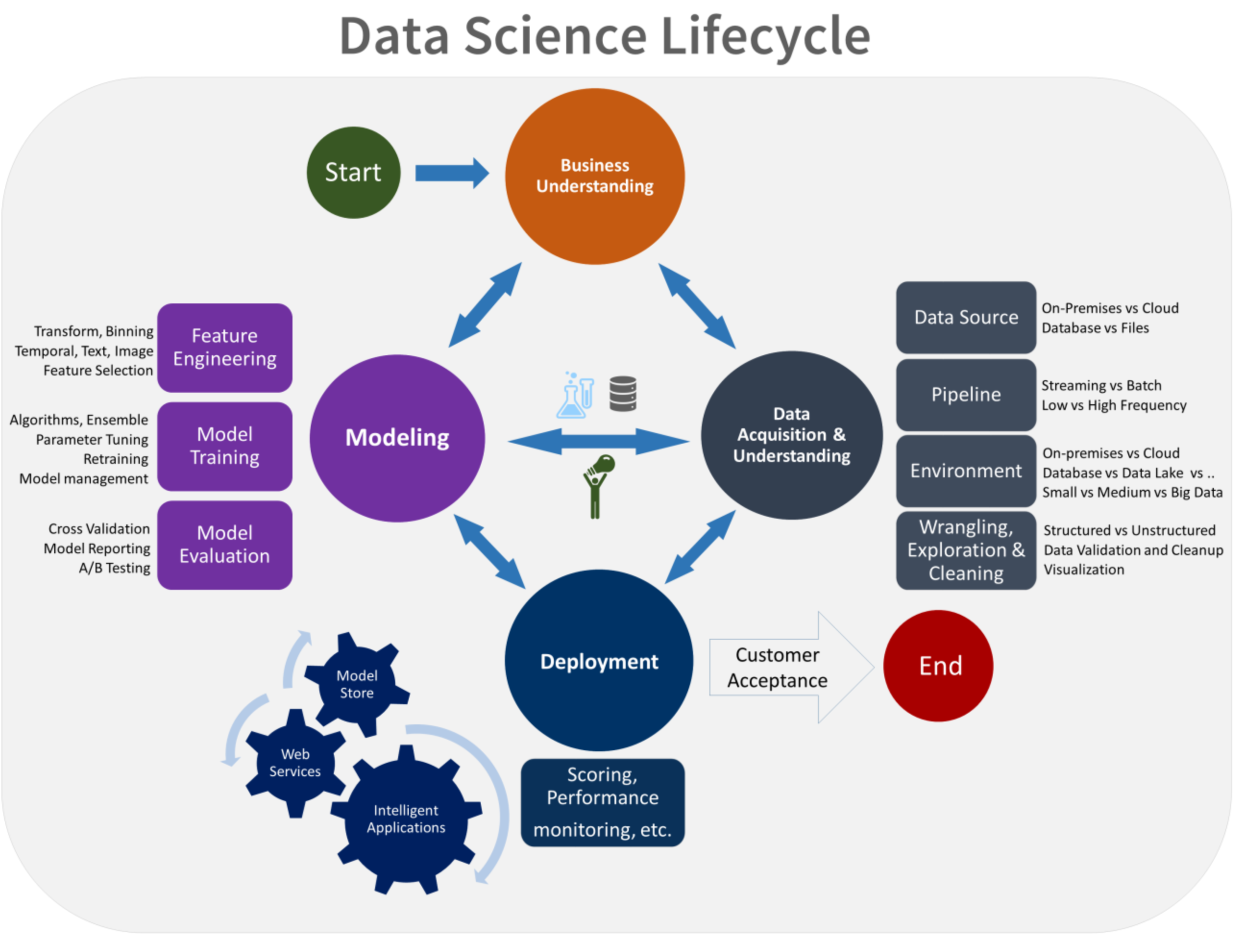
The AI/ML Journey

Data Science Lifecycle

The lifecycle outlines the major stages that projects typically execute, often iteratively:

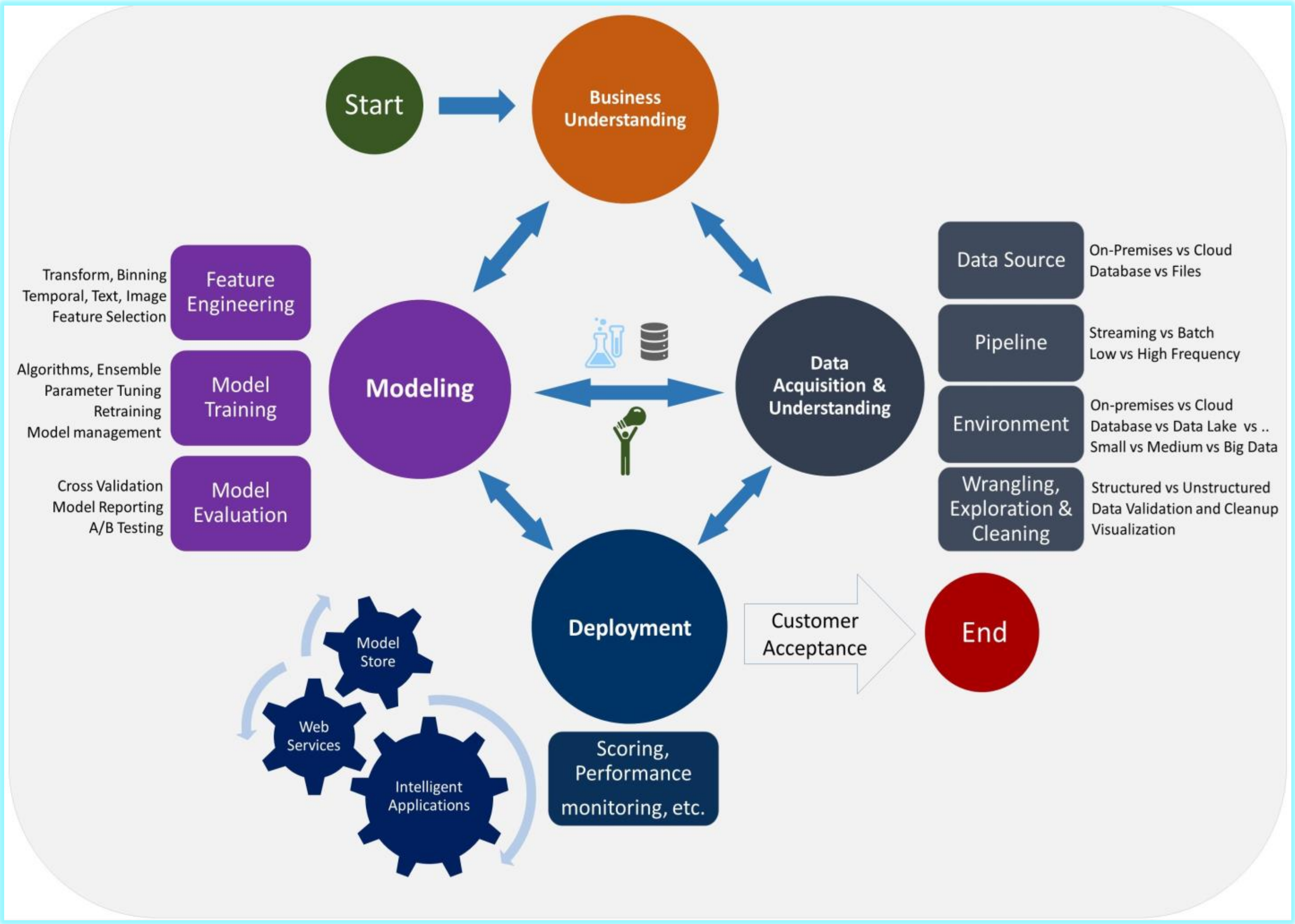
- **Business Understanding**
- **Data Acquisition and Understanding**
- **Modeling**
- **Deployment**

Here is a visual representation of the **Team Data Science Process lifecycle**.



https://docs.microsoft.com/en-us/azure/architecture/data-science-process/overview

Data Science Lifecycle



- Business Understanding
  - Identify business problem
  - Find relevant data sources
  - Establish a business strategy
- Data Acquisition and Understanding
  - Data discovery
  - Data cleansing and prep
  - Data pipeline
- Modeling
  - Optimal data features
  - Determine ML model
  - Implement data pipeline
- Deployment
  - Deploy
  - Monitor
  - Retrain
  - Visualize

https://docs.microsoft.com/en-us/azure/architecture/data-science-process/overview

Lifecycle stages can be integrated with specific deliverables & checkpoints

Business Understanding

- Project Objective
- Data, Target & Feature Definition
- Data Dictionary

Data acquisition and understanding

- Data Quality Report
- Architecture Diagram (initial draft)

Modeling

- Featurization
- Modeling Report

Deployment

- ML Model
- Final model reporting
- Deployment platform & architecture

Build Business Understanding

Begin with the exploration of a problem

"We spend a lot of money to acquire and retain customers. Turning over a customer, or reacquiring them is also expensive."

Arrive at an analytical question that addresses the problem

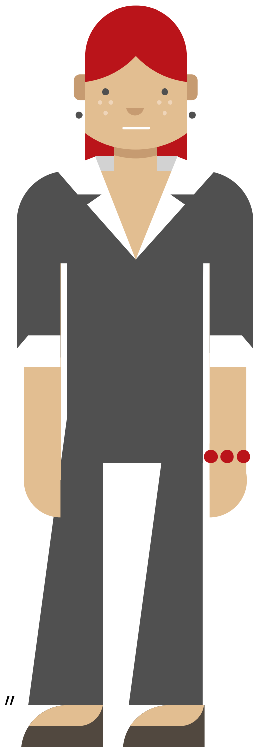
"It would be great if we could predict whether a customer is likely to leave us next month and target them for retention."

Understand the data required to address the problem

"We have historical data representing customer acquisition and turnover for the past n years, including all customer interactions."

Define the consumption experience

How will people or systems use these conclusions, on which devices. How will they want to consume the analytic output and what will they do with it?



Define business goals with "sharp" questions that can be answered by Data Science:

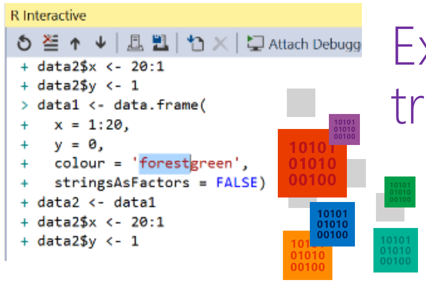
- How much or how many? (regression)
- Which category? (classification)
- Which group? (clustering)
- Is this weird? (anomaly detection)
- Which option should be taken? (recommendation)

Data Prep and Modelling

1110
1110 1110
1010 1010
1010 1010 1010

Acquire data sets

Plan and locate potential sets of data. Gather data from internal and external sources



Explore and transform the data

Using data transformation tools and engineering techniques, mine, extract, clean, mark up, label, transform, enrich and stage data

Begin the modelling experimentation

Feature engineering, model fitting, model evaluation on “unseen” data

10101
01010
00100
11011
0101
010
00

Experimentation may cause re-visitation of data understanding or even the business problem

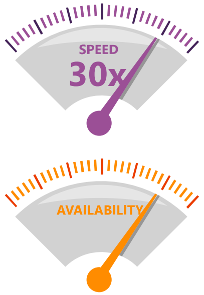
Data prep tasks can be repeated multiple times

Rigorous model evaluation

Deploy and Monitor Performance

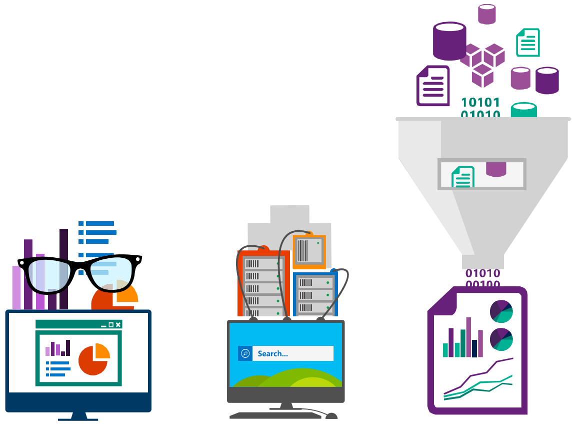
Model deployment/ operationalization

Systematic integration into dashboards, systems, applications including continued data ingestion and training



Monitor performance

Continuous monitoring of model performance, refine model based on field feedback, retrain model on new data

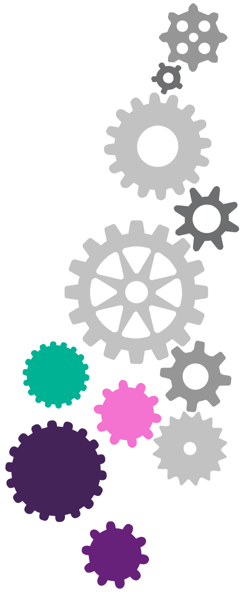


Move beyond "experimentation" into production

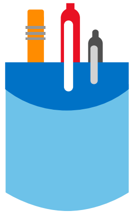
Integrate into dashboards, apps and acquire new data sets

Monitor performance accuracy

Revisit the modelling stage to refine the model as needed



Success Ingredients for a Data Science Project

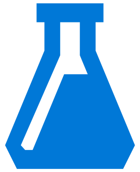


Know your problem

Is there a clear business need?
Is the outcome measurable?

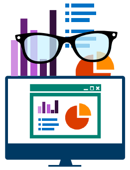
Know your data

Is there data to support the business scenario? What are the caveats? Data exploration and validation are essential.



Build relevant models

Which types of models can answer the business question? Is it a regression, classification, anomaly detection, clustering, recommendation, or other problem?



Integrate with business

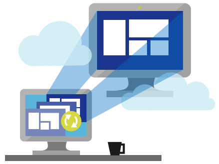
Trust building through testing and validation, then deploy

Building Scalable, Secure and Responsible AI Solutions in Azure

Azure Machine Learning service

Bring AI to everyone with an end-to-end, scalable, trusted platform

Boost your data science productivity



Increase your rate of experimentation

Deploy and manage your models everywhere

Built with your needs in mind

- Automated machine learning
- Managed compute
- Simple deployment
- DevOps for machine learning
- Support for open source frameworks
- Tool agnostic Python SDK

Seamlessly integrated with the Azure Portfolio

Machine Learning on Azure

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 advanced 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

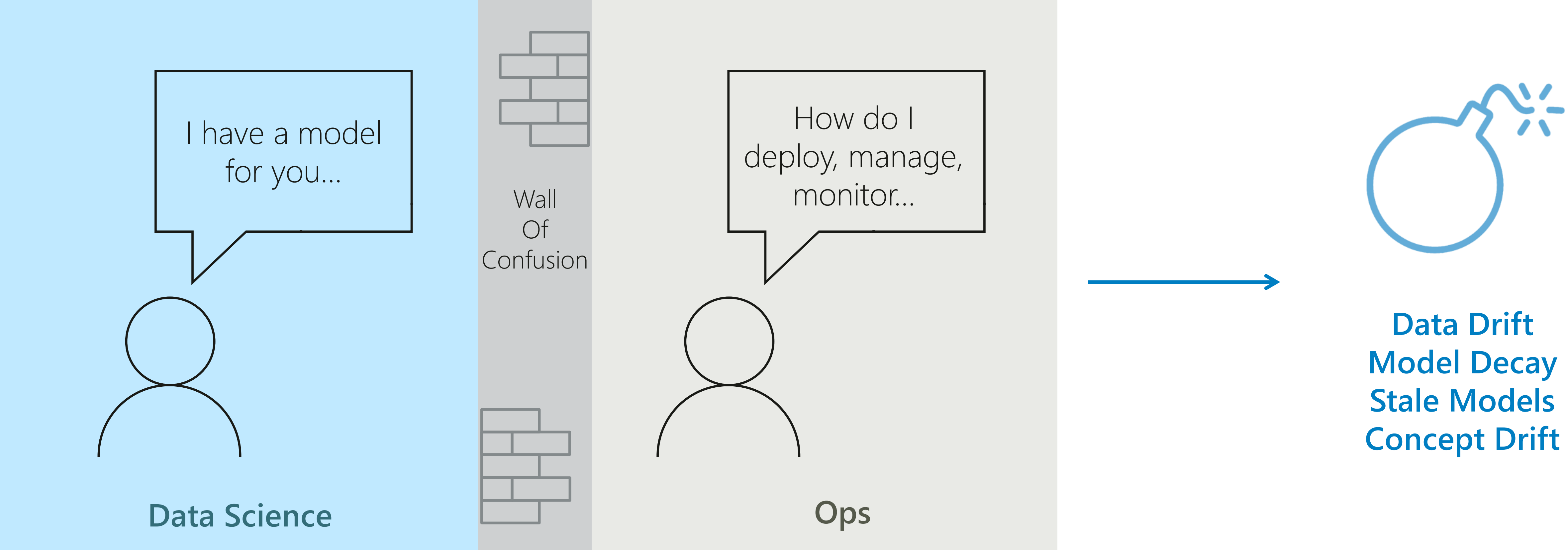
Powerful infrastructure
To accelerate deep learning

- CPU
- GPU
- FPGA

From the Intelligent Cloud to the Intelligent Edge

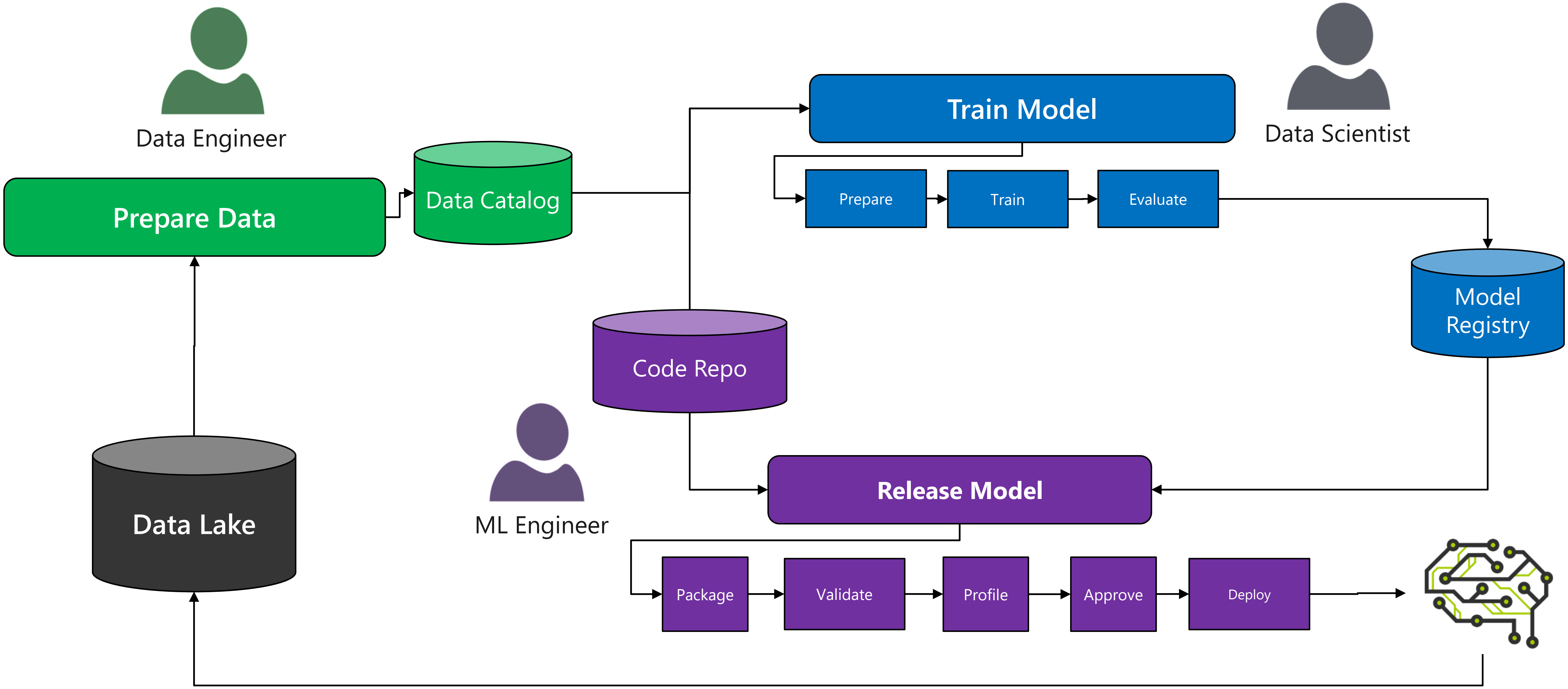
ML Ops (Machine Learning Ops)

Traditional DS Delivery

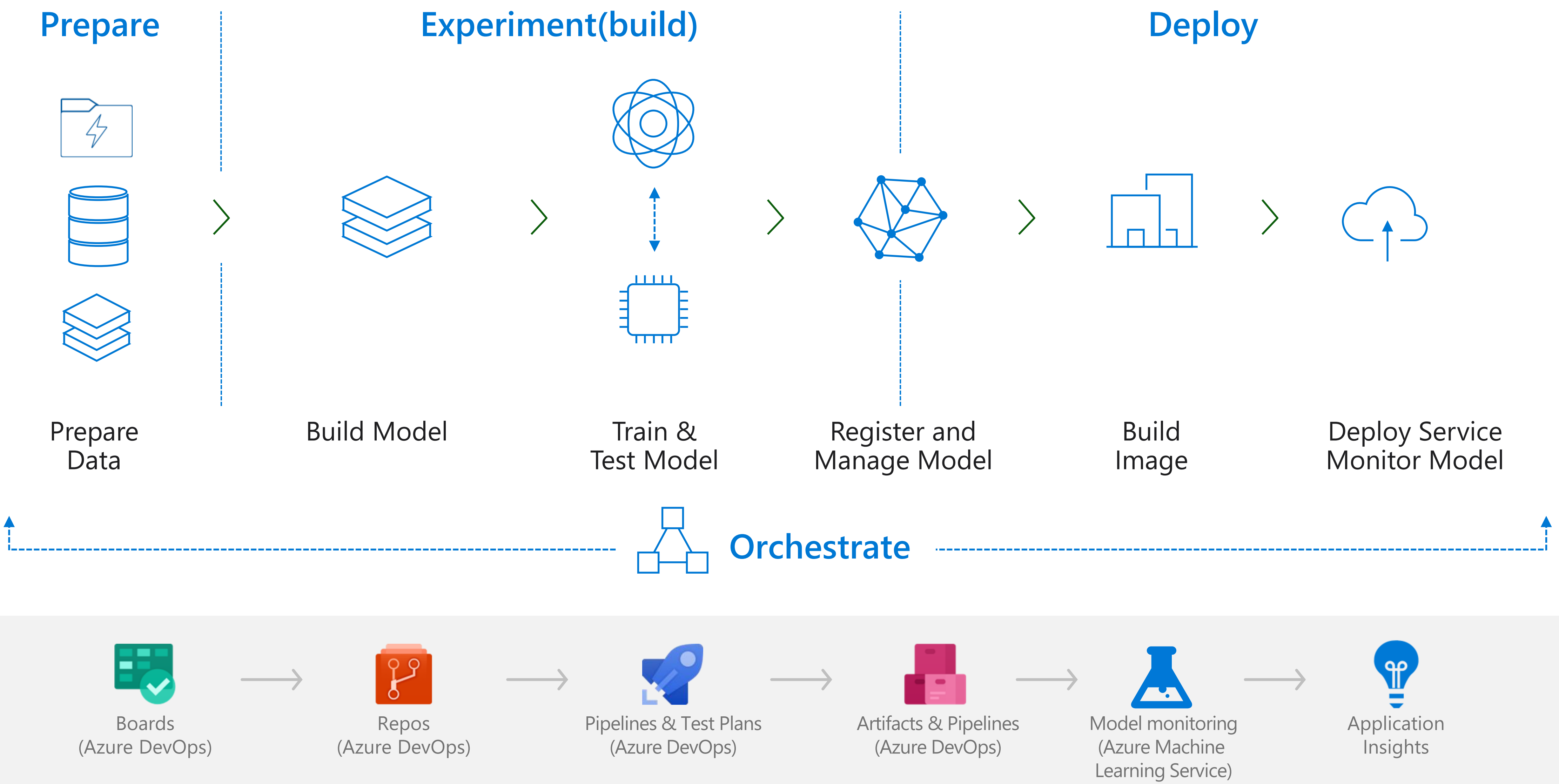


MLOps Process

Enterprise ready machine learning development



MLOps Workflow



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

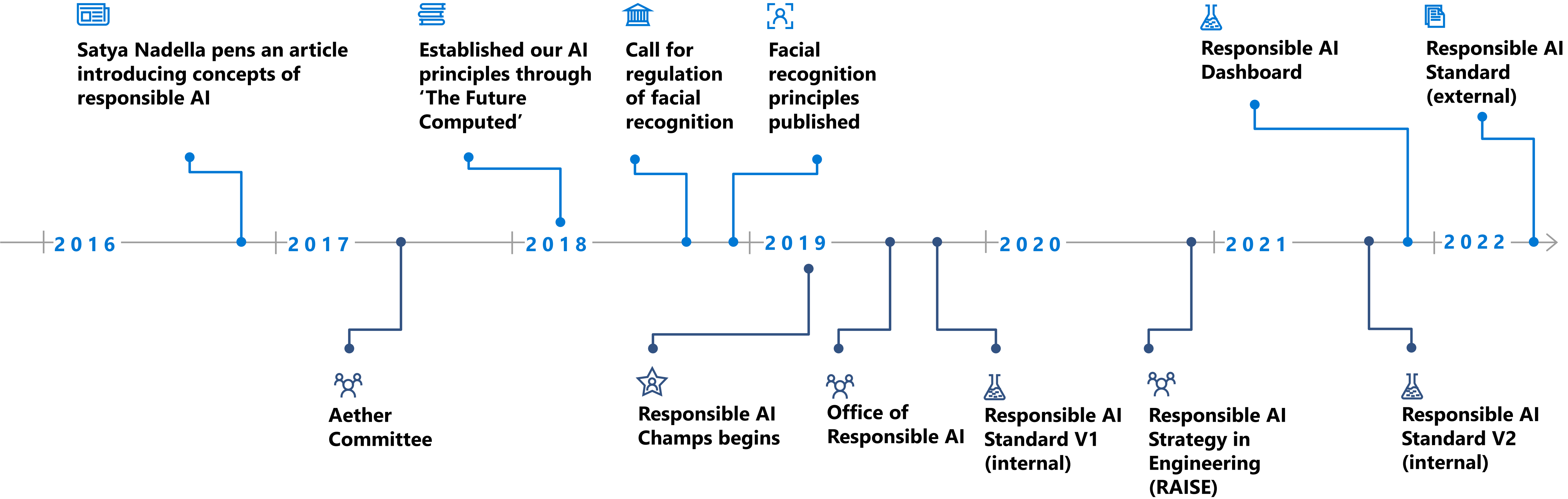
Responsible AI (Artificial Intelligence)

Why responsible AI?

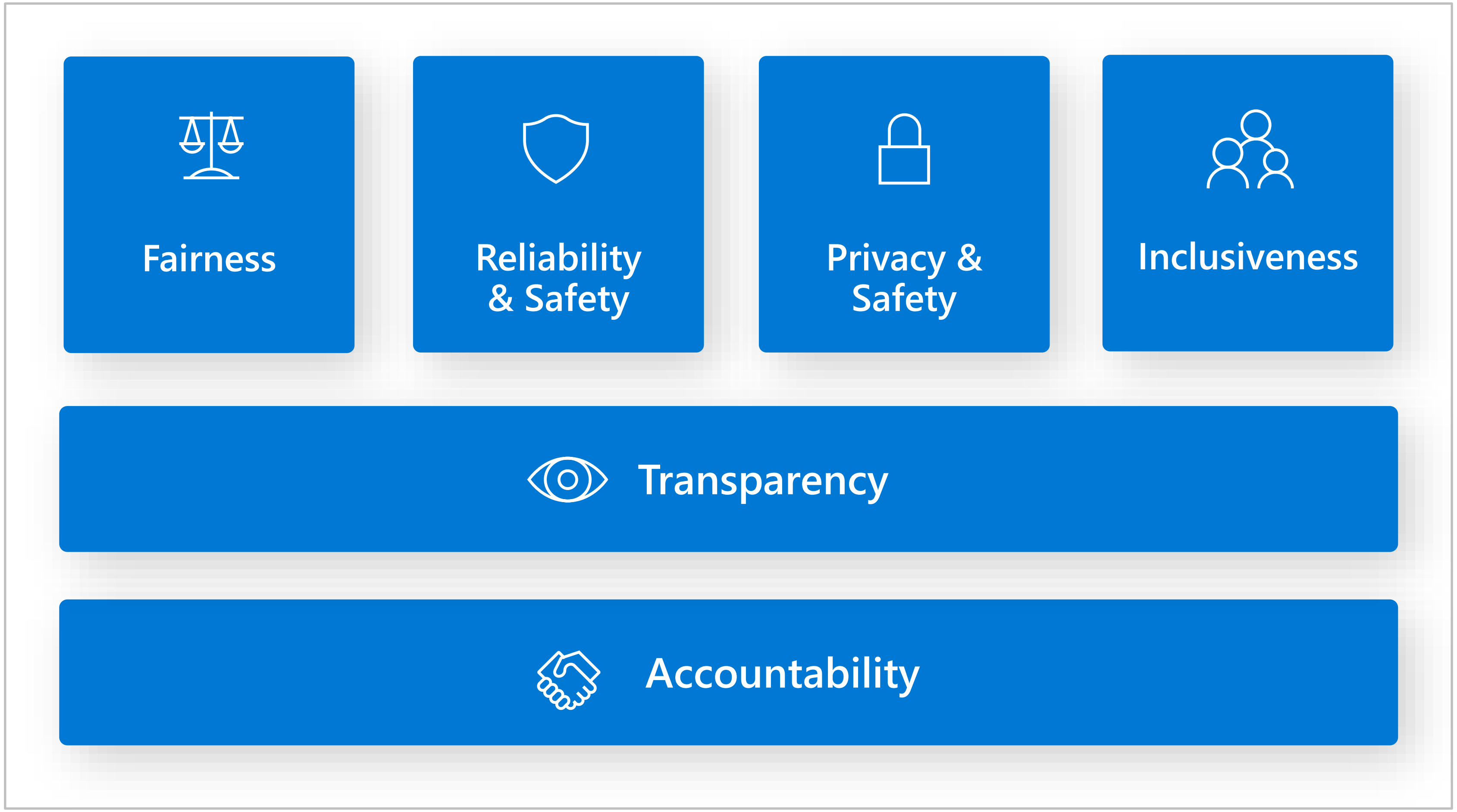
"When you create technology that changes the world, you have to assume a responsibility for the world that you've helped to create."

Brad Smith
President and Chief Legal Officer, Microsoft

Our Responsible AI journey



Learn Microsoft's AI principles



Customer Story: NHS

Northumbria Healthcare NHS Foundation adopts Responsible AI philosophy with Azure Machine Learning: https://www.youtube.com/watch?v=LRZHcipcweY

Thank you
Gracias
Obrigado