Anomaly Detection in Wind Turbine using Data Science and Big Data

Anish Mahapatra | 19 October, 2021

Agenda

upGrad

Introduction

SENIOR DATA SCIENTIST

ANISH MAHAPATRA

Google Me.

anishmahapatra01@gmail.com

Senior Data Scientist 🥷

LinkedIn: www.linkedin.com/in/**anishmahapatra**

Medium: anishmahapatra.medium.com

Website: www.**anishmahapatra**.com

GitHub: www.github.com/**anishmahapatra**

Q. How to get into Data Science?

A. Simple, ask a real-world Data Scientist.

The Data Lifecycle

01
Understand the Data
Source the data,
Create Data Dictionary,
View Data Summary,
View Data Distribution

02
Data Exploration
Exploratory Data Analysis,
Make ER Diagram,
Main Analytical Data frame

03
Data Roll-up
Aggregation and roll-up of data,
Merging of Data on relevant keys,
Unit testing,
Building pipelines to experiment,

06
Model Evaluation
Model Evaluation,
Monitoring Model Performance and Drift,
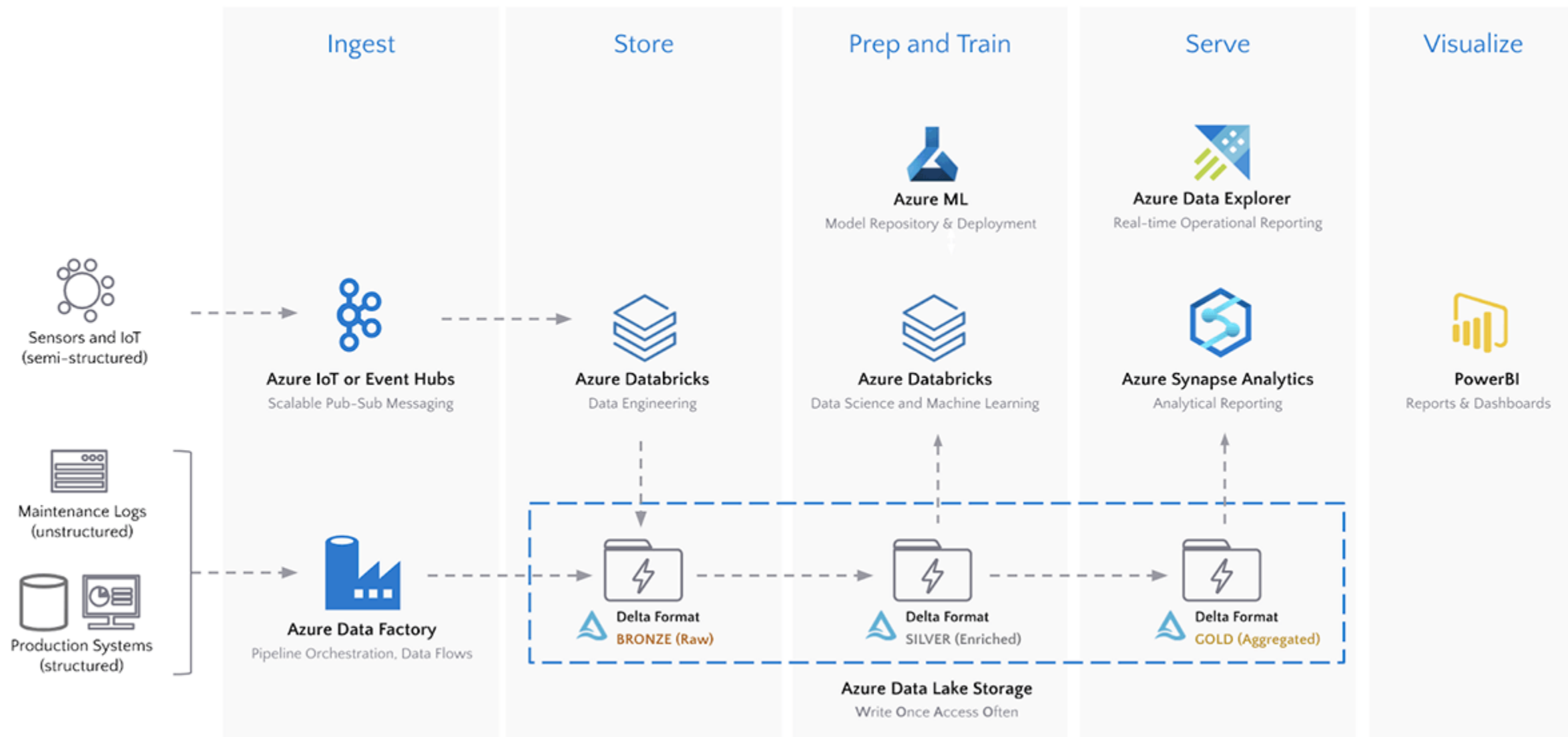Data Governance Policies

05
Model Interpretability
Model Interpretability using SHAP,
Machine Learning interpretation,
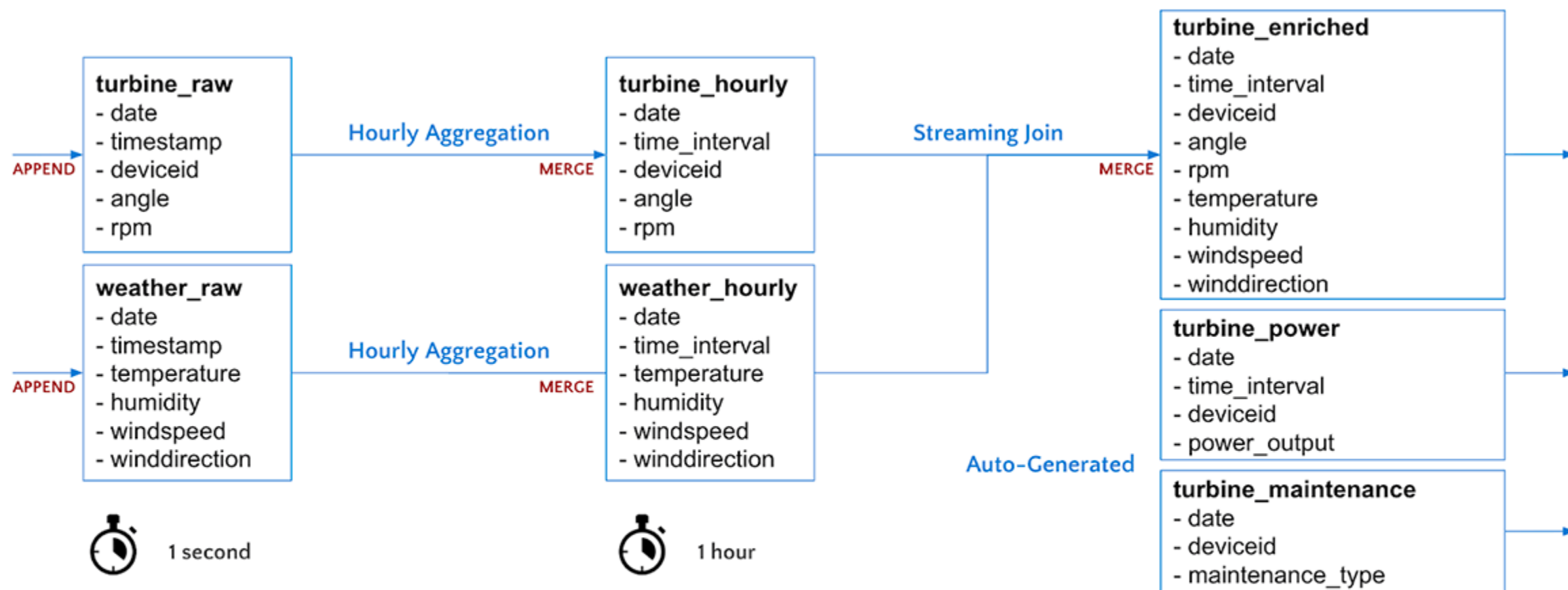Business Explainability

04
Feature Engineering
Building new features,
Combining external data,
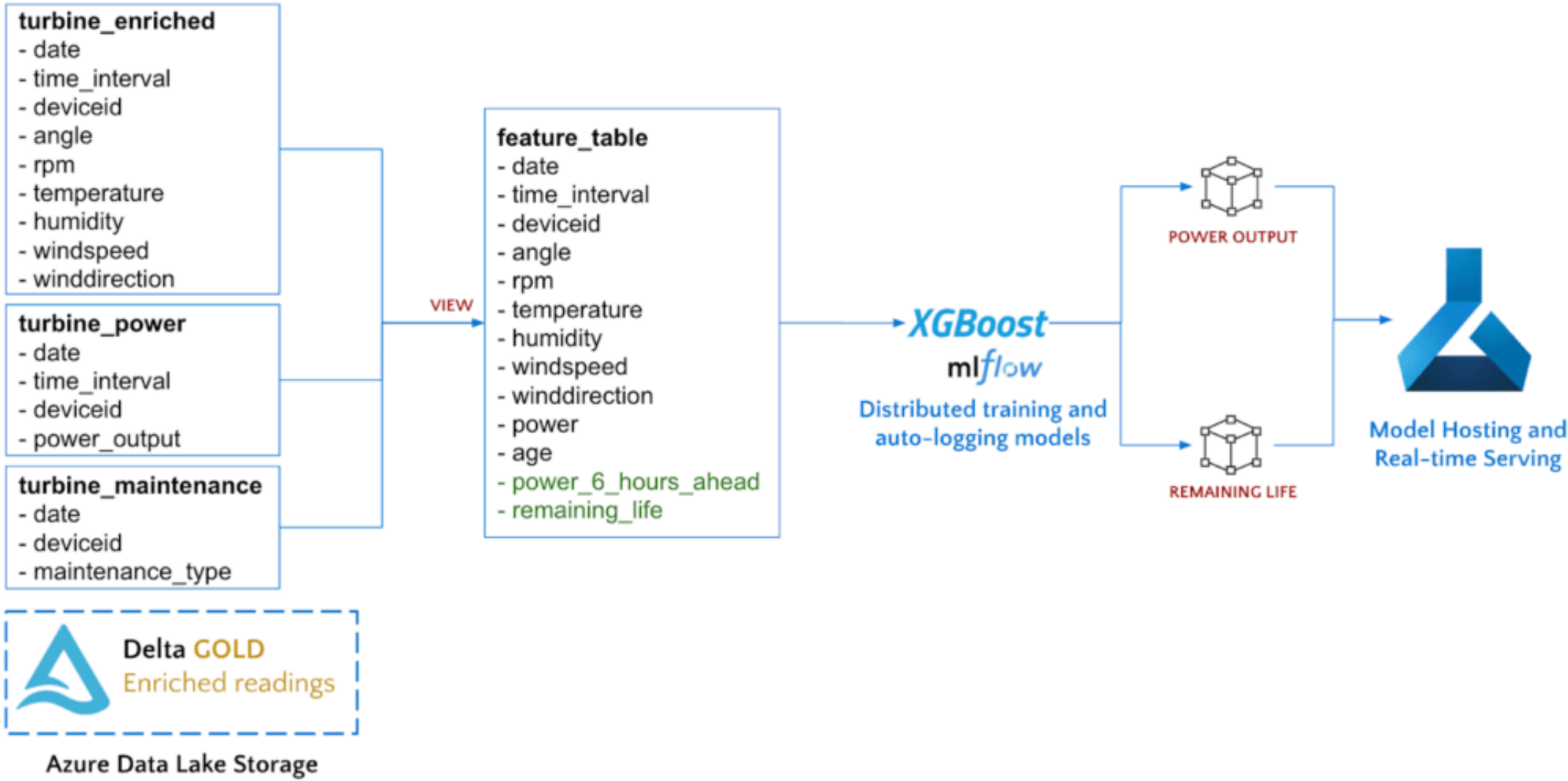Building a feature store

Production Machine Learning Architecture
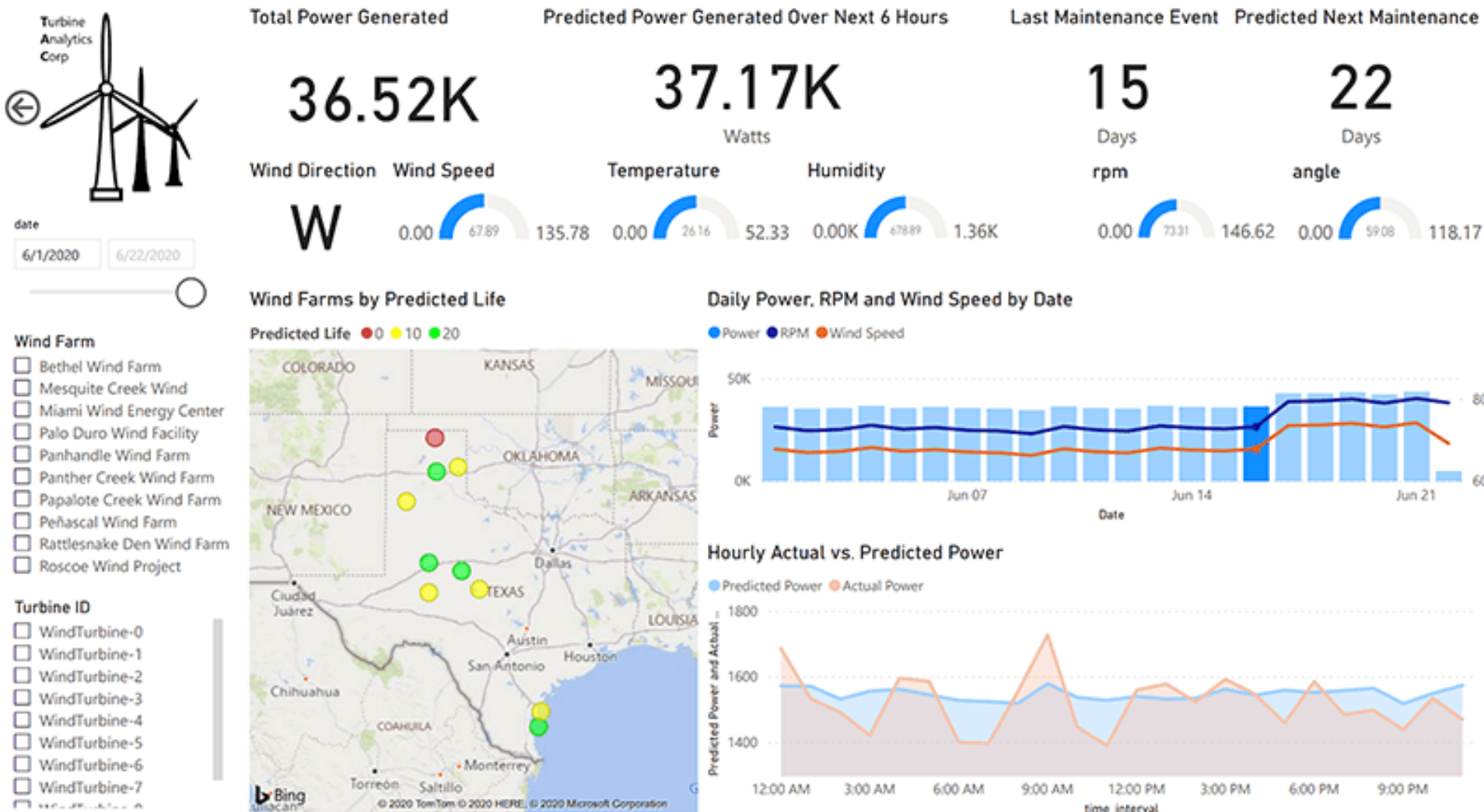
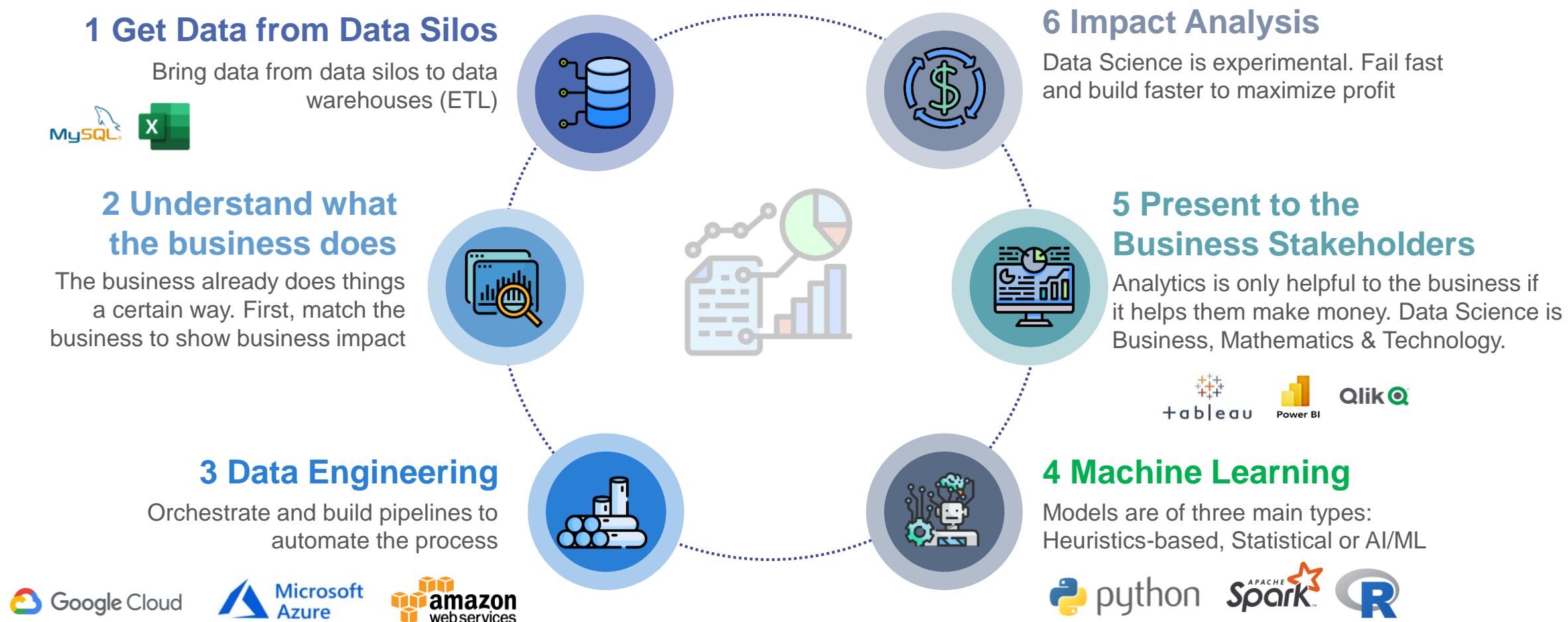# Implementation in Data Lake

Implementation in Data Lake

Predictive Maintenance Dashboard

Predictive Maintenance Dashboard

The Data Science Process Flow

1 Get Data from Data Silos

Bring data from data silos to data warehouses (ETL)

2 Understand what the business does

The business already does things a certain way. First, match the business to show business impact

3 Data Engineering

Orchestrate and build pipelines to automate the process

6 Impact Analysis

Data Science is experimental. Fail fast and build faster to maximize profit

5 Present to the Business Stakeholders

Analytics is only helpful to the business if it helps them make money. Data Science is Business, Mathematics & Technology.

4 Machine Learning

Models are of three main types: Heuristics-based, Statistical or AI/ML

Anomaly Detection in Wind Turbine using Data Science and Big Data

Anish Mahapatra

upGrad

Q&A Session

LinkedIn: www.linkedin.com/in/**anishmahapatra**

Email: anishmahapatra01@gmail.com

Anomaly Detection in Wind Turbine using Data Science and Big Data

Anish Mahapatra

upGrad

Q. *How to get into Data Science?*

A. Simple, ask a real-world Data Scientist.

Thank you.

Email: anishmahapatra01@gmail.com