Microsoft

Analytics on Azure

What to Use When

Christina E. Leo

Cloud Solution Architect | Data & AI | One Commercial Partner | United Kingdom
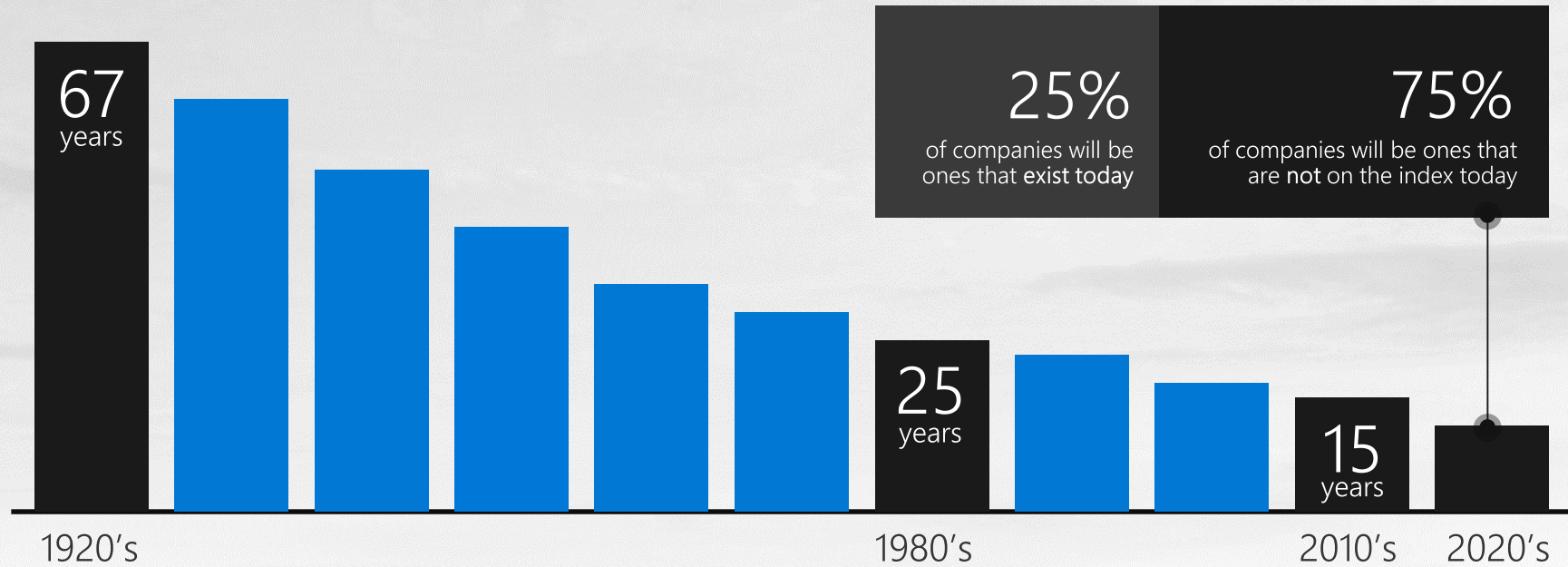
christina.leo@microsoft.com

@christinaleo

A World of Unprecedented Change
Average lifespan of a company on S&P 500

67 years

25 years

15 years

25% of companies will be ones that **exist today**

75% of companies will be ones that are **not** on the index today

1920's

1980's

2010's   2020's

Source: Yale Professor Richard Forster

The Azure Data Landscape

Data Sources

Apps

Sensors and devices

Orchestration

- Data Factory
- Integrations Services
- Data Catalog
- Storage Blob
- Table Storage
- Event Hubs

IoT
IoT Hub

Data Storage

- Data Lake Store
- CosmosDB
- SQL Azure
- SQL Managed Instance
- SQL Data Warehouse
- SQL Server 2017

Machine Learning

- Studio
- ML Development Environments
- Azure ML Services
- Server
- Data Science & Deep Learning VM

Big Data & Analytics

- Stream Analytics
- Analysis Services
- HDInsight
- DataBricks

Prebuild Models

- Cognitive Services

Operationalization

- Functions & Serverless
- Bot Framework
- Azure Search

Dashboards & Visualizations

- Power BI

People

Apps
- Web
- Mobile
- Bots

Automated Systems

DATA → STORAGE → INTELLIGENCE → ACTION

Solution Scenarios
Big Data & Advanced Analytics

Modern Data Warehousing

"We want to integrate all our data including 'big data' with our data warehouse"

Real-Time Analytics

"We are trying to get insights from our devices in real-time, etc."

Advanced Analytics

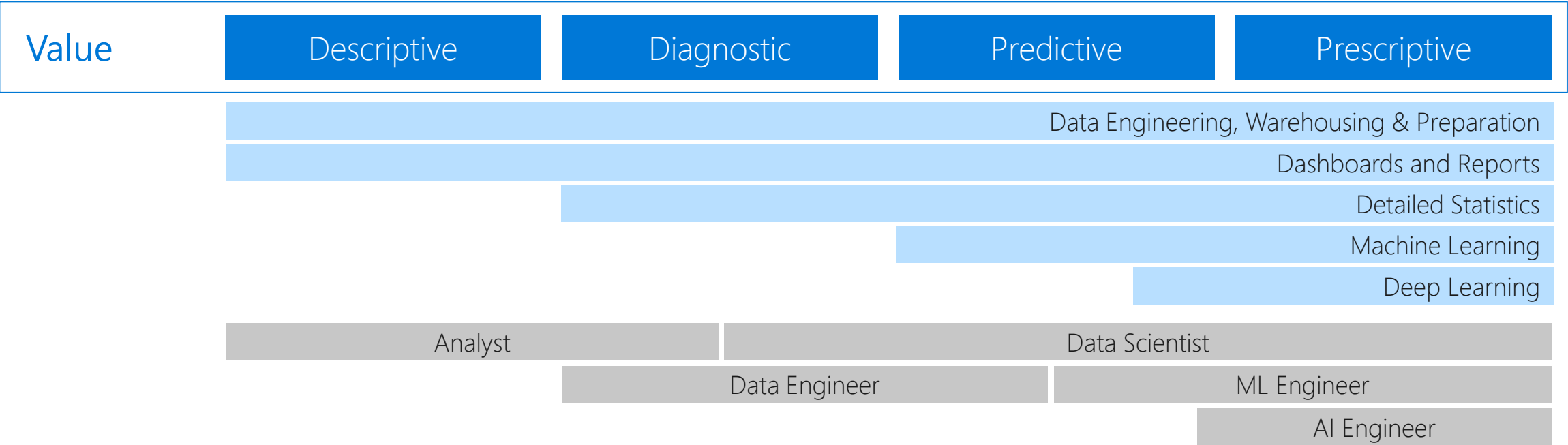"We are trying to predict when our customers churn."

Uses of Data

	BI	AA		AI
Value	Descriptive	Diagnostic	Predictive	Prescriptive

Data Engineering, Warehousing & Preparation

Dashboards and Reports

Detailed Statistics

Machine Learning

Deep Learning

Analyst

Data Scientist

Data Engineer

ML Engineer

AI Engineer

Modern Data Warehousing

The Modern Data Warehouse extends the scope of the data warehouse to serve "big data" that is prepared with techniques beyond relational ETL.

Modern Data Warehousing

"We want to integrate all our data including 'big data" with our data warehouse"

Real-Time Analytics

"We are trying to get insights from our devices in real-time, etc."

Advanced Analytics

"We are trying to predict when our customers churn."

MODERN DATA WAREHOUSING
CANONICAL OPERATIONS
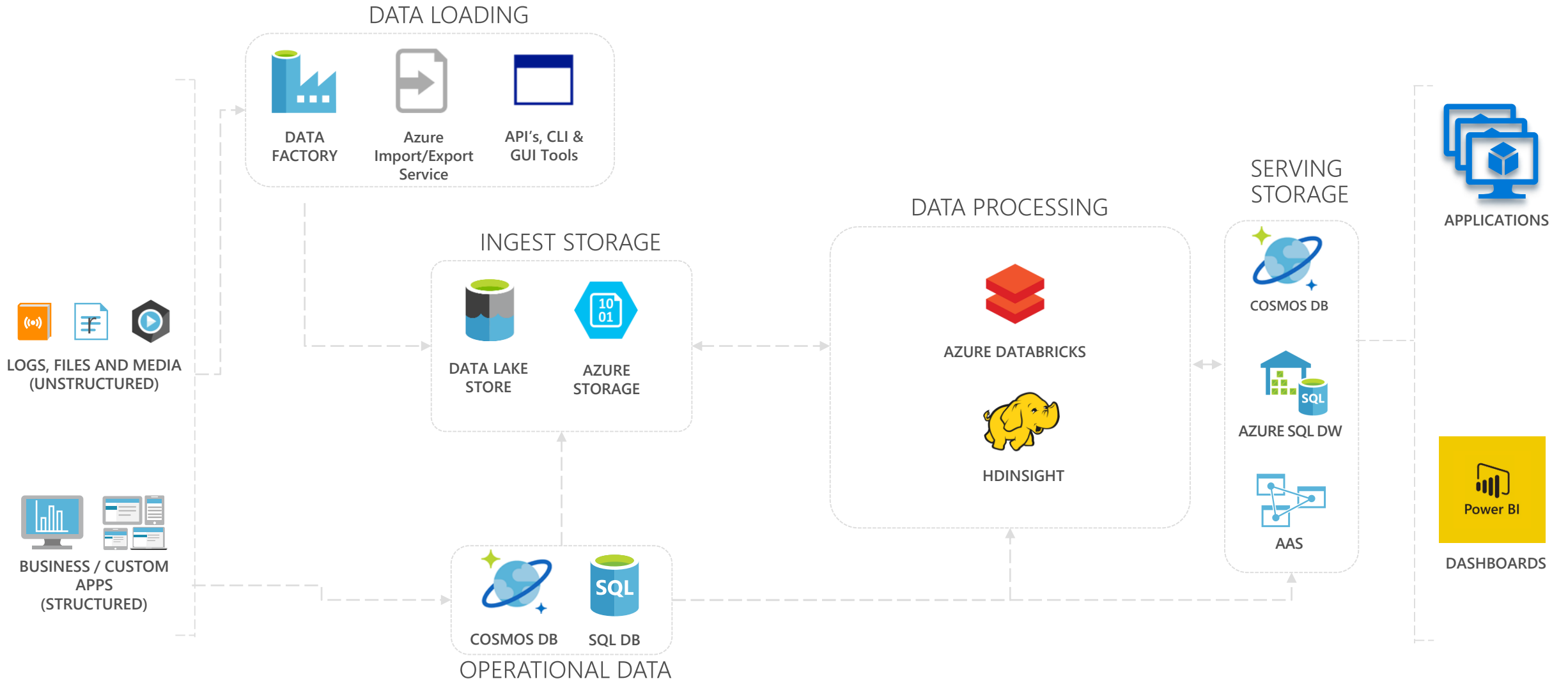
Load & Ingest

TRANSFER, STORE

Process

PROCESS, CLEAN

Serve

SERVE, ANALYZE

DATA WAREHOUSING PATTERN IN AZURE

Loading and preparing data for analysis with a data warehouse

DATA LOADING

DATA FACTORY

Azure Import/Export Service

API's, CLI & GUI Tools

LOGS, FILES AND MEDIA (UNSTRUCTURED)

INGEST STORAGE

DATA LAKE STORE

AZURE STORAGE

BUSINESS / CUSTOM APPS (STRUCTURED)

DATA PROCESSING

AZURE DATABRICKS

HDINSIGHT

OPERATIONAL DATA

COSMOS DB

SQL DB

SERVING STORAGE

COSMOS DB

AZURE SQL DW

AAS

APPLICATIONS

Power BI

DASHBOARDS

Real-Time Analytics

Real-time Analytics (aka Stream Analytics) is the phenomenon of processing data as soon as it is generated, to derive very quick analysis/insight for timely action.

Modern Data Warehousing

"We want to integrate all our data including 'big data' with our data warehouse"

Real-Time Analytics

"We are trying to get insights from our devices in real-time, etc."

Advanced Analytics

"We are trying to predict when our customers churn."

STREAMING - CANONICAL OPERATIONS
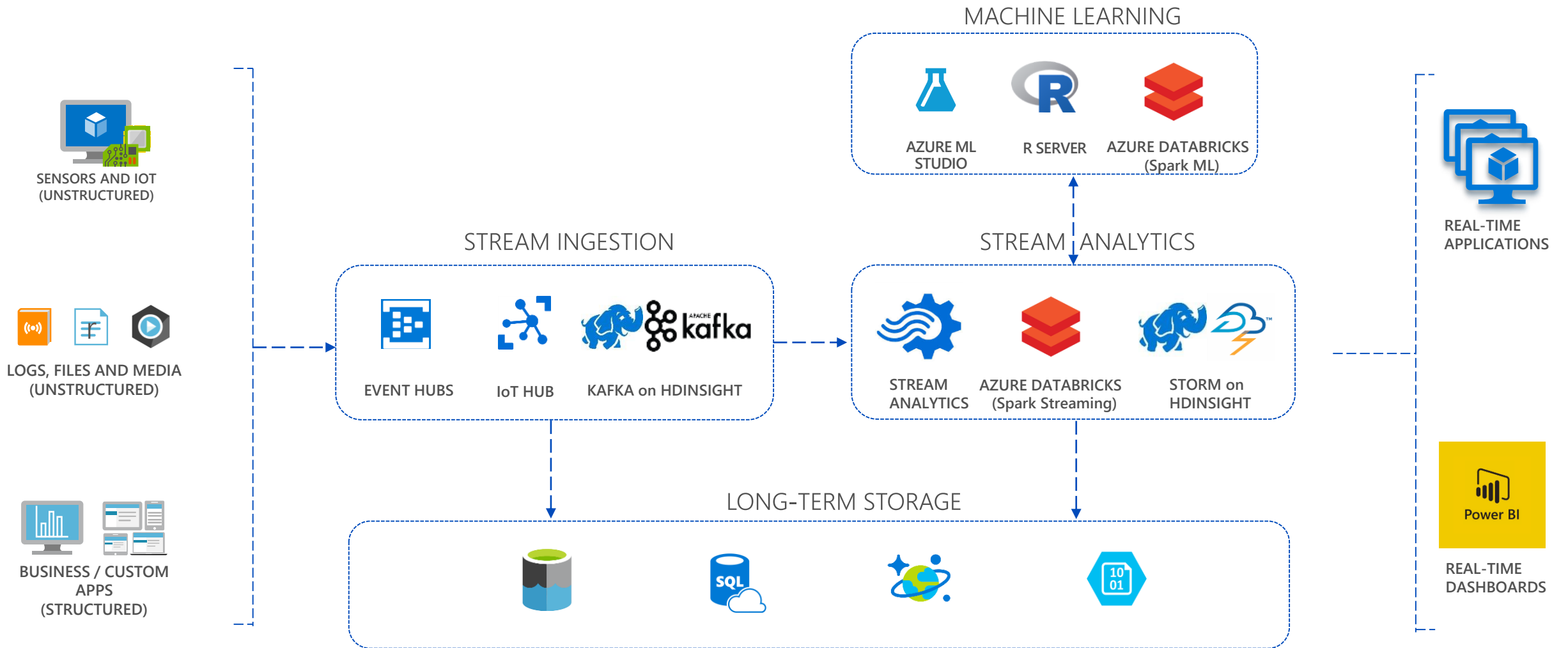
Ingest
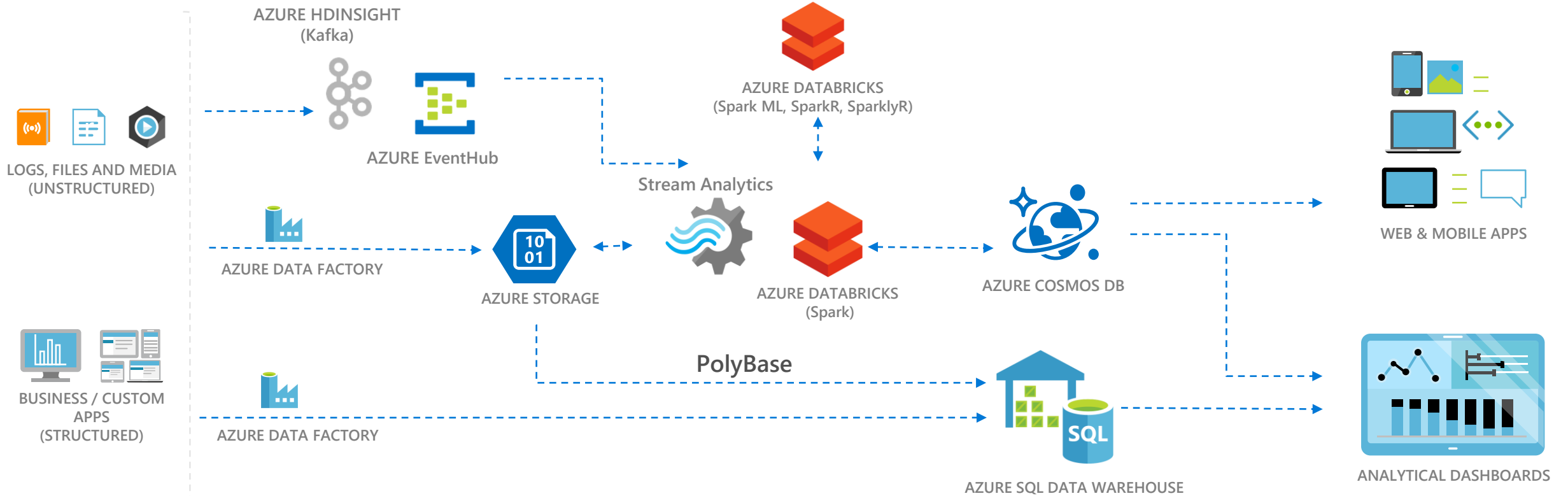**CONNECT, COLLECT, STORE**

Analytics
PROCESS, ANALYZE

Actions
REPORT, VISUALIZE, ACT

BIG DATA STREAMING PATTERN WITH AZURE

MACHINE LEARNING

AZURE ML STUDIO R SERVER AZURE DATABRICKS (Spark ML)

SENSORS AND IOT (UNSTRUCTURED)

LOGS, FILES AND MEDIA (UNSTRUCTURED)

BUSINESS / CUSTOM APPS (STRUCTURED)

STREAM INGESTION

EVENT HUBS IoT HUB KAFKA on HDINSIGHT

STREAM ANALYTICS

STREAM ANALYTICS AZURE DATABRICKS (Spark Streaming) STORM on HDINSIGHT

LONG-TERM STORAGE

REAL-TIME APPLICATIONS

Power BI

REAL-TIME DASHBOARDS

DATA WAREHOUSING PATTERN IN AZURE

Big Data Lambda Architecture

AZURE HDINSIGHT
(Kafka)

AZURE EventHub

LOGS, FILES AND MEDIA
(UNSTRUCTURED)

AZURE DATA FACTORY

AZURE STORAGE

Stream Analytics

AZURE DATABRICKS
(Spark ML, SparkR, SparklyR)

AZURE DATABRICKS
(Spark)

AZURE COSMOS DB

WEB & MOBILE APPS

BUSINESS / CUSTOM
APPS
(STRUCTURED)

AZURE DATA FACTORY

PolyBase

AZURE SQL DATA WAREHOUSE

ANALYTICAL DASHBOARDS

Advanced Analytics

Advanced Analytics is the process of applying machine learning and/or deep learning techniques to data for the purpose of creating predictive/prescriptive insights.

Modern Data Warehousing

"We want to integrate all our data including 'big data" with our data warehouse"

Real-Time Analytics

"We are trying to get insights from our devices in real-time, etc."

Advanced Analytics

"We are trying to predict when our customers churn."

ADVANCED ANALYTICS-CANONICAL OPERATIONS

Data Acquisition & Understanding

Modeling

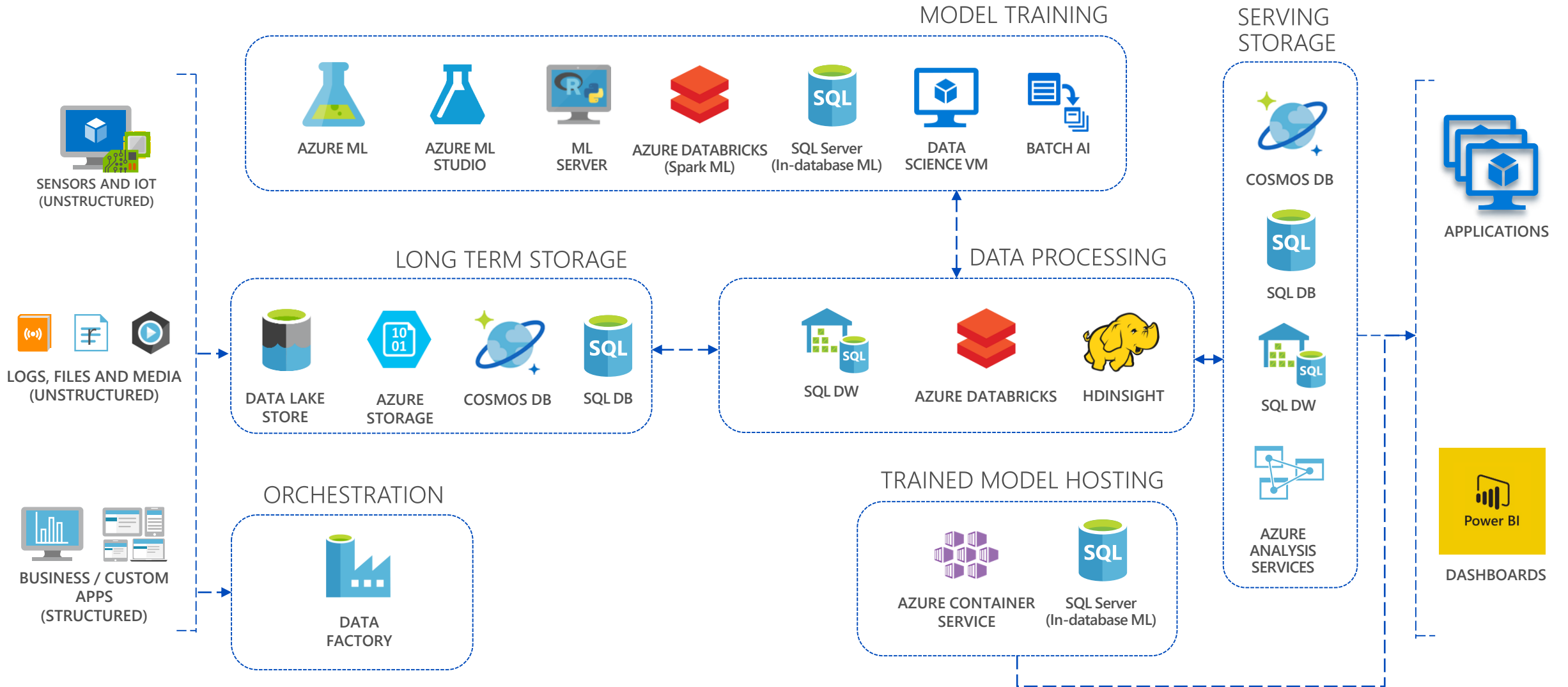**Deployment**

ACQUIRE, UNDERSTAND

TRAINING, VALIDATION

DEPLOY, INTEGRATE

ADVANCED ANALYTICS PATTERN IN AZURE

Performing data collection/understanding, modeling and deployment

MODEL TRAINING

AZURE ML

AZURE ML STUDIO

ML SERVER

AZURE DATABRICKS (Spark ML)

SQL Server (In-database ML)

DATA SCIENCE VM

BATCH AI

SENSORS AND IOT (UNSTRUCTURED)

LOGS, FILES AND MEDIA (UNSTRUCTURED)

BUSINESS / CUSTOM APPS (STRUCTURED)

LONG TERM STORAGE

DATA LAKE STORE

AZURE STORAGE

COSMOS DB

SQL DB

DATA PROCESSING

SQL DW

AZURE DATABRICKS

HDINSIGHT

SERVING STORAGE

COSMOS DB

SQL DB

SQL DW

AZURE ANALYSIS SERVICES

ORCHESTRATION

DATA FACTORY

TRAINED MODEL HOSTING

AZURE CONTAINER SERVICE

SQL Server (In-database ML)

APPLICATIONS

Power BI

DASHBOARDS

DEMO

THINGS TO NOTE

› There are no right or wrong solutions, only optimal solutions

› We lead with certain solutions and customise based on customer scenarios

› Customer voice and product and service maturity govern lead solutions

› Consider price and performance, ease of use, and ecosystem acceptance as factors

› Everything is fluid - a lead solution today might be non-optimal tomorrow, based on the factors above and new releases

Decision Points

	IaaS	PaaS	SaaS

Customer:
- ☐ Regionality
- ☐ VNET requirements
- ☐ Speed of business

- ☐ Data Volumes
- ☐ Data Types
- ☐ Governance
- ☐ Team
- ☐ Skills

AI Decision Tree

FEEDBACK FORMS

PLEASE FILL OUT AND PASS TO YOUR ROOM
HELPER BEFORE YOU LEAVE THE SESSION