Microsoft

Technical Summit Workshop
Become a Data Ninja

Olivia Klose, Software Development Engineer, Microsoft

Agenda

09:00 – 10:30 Cortana Intelligence Suite + Azure Env.

10:30 – 11:00 Kaffeepause

11:00 – 12:30 Stream Processing

12:30 – 13:30 Mittagspause

13:30 – 15:00 Machine Learning

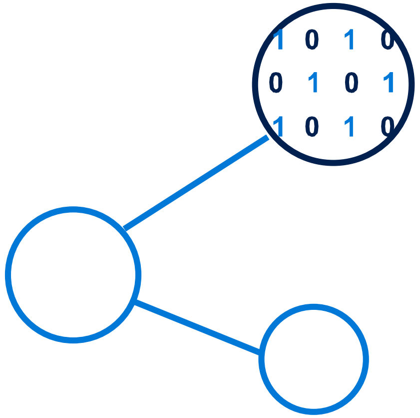
15:00 – 15:30 Kaffeepause

15:30 – 17:00 Pre-Processing and Orchestration

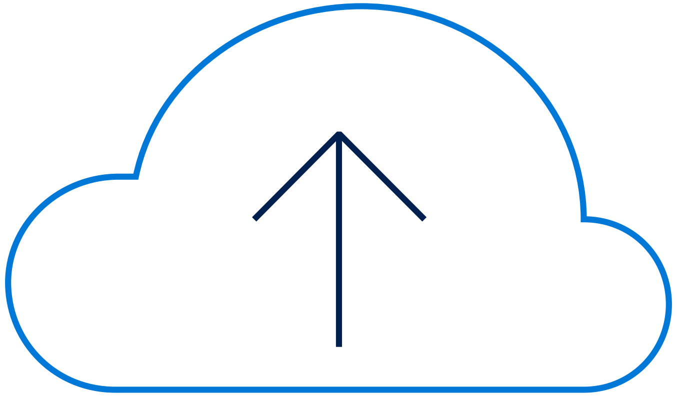
Installing...

1. Visual Studio Code with MSSQL Extension
   (http://code.visualstudio.com)
   OR
   SQL Server Management Studio
   OR
   another similar tool to access a SQL server DB
2. Microsoft Azure Storage Explorer
   (http://storageexplorer.com)

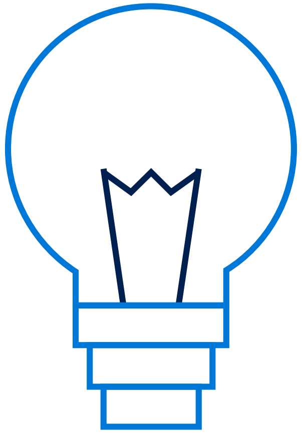
Business is being transformed by three trends



Big Data

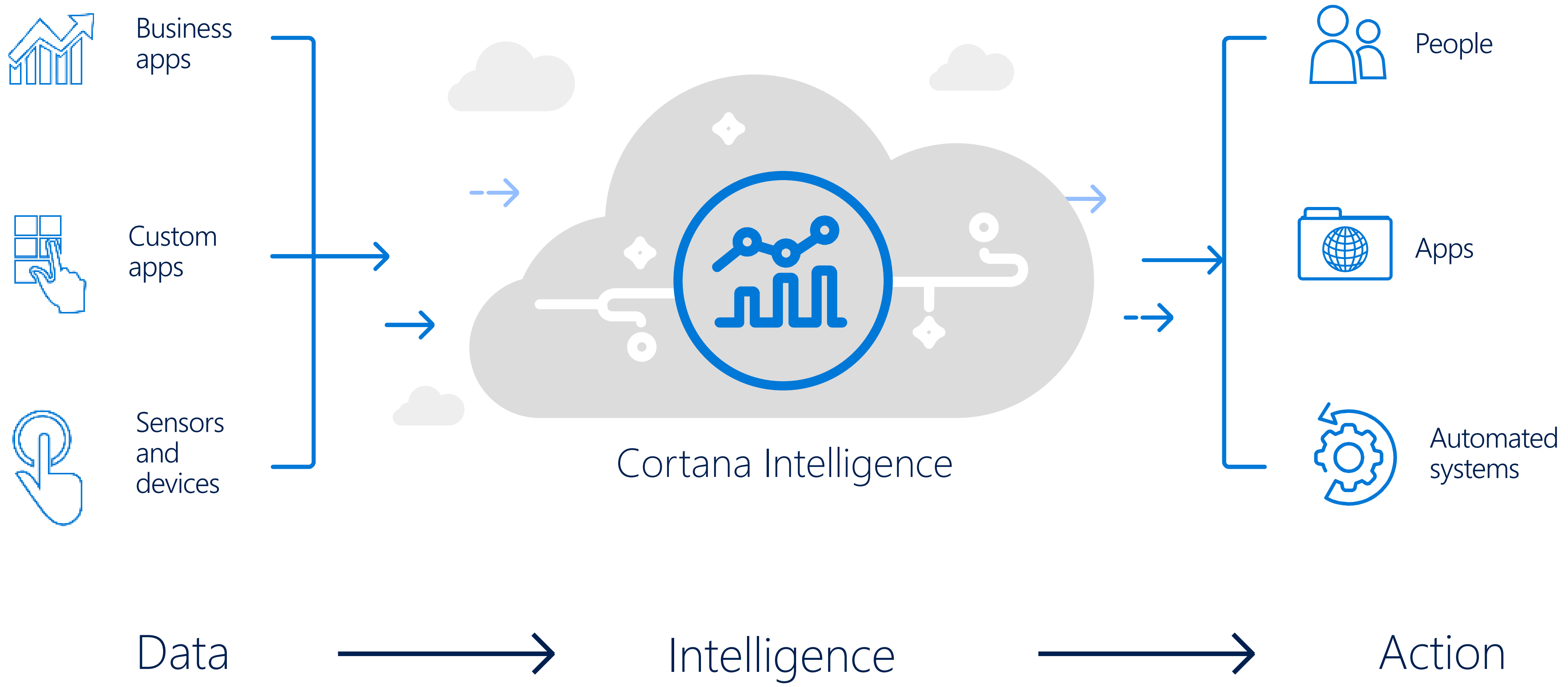


Cloud



Intelligence

Stay ahead of the curve with Cortana Intelligence Suite



Examples of Analytics Use Cases

Improving visibility and making accurate predictions with **remote monitoring**

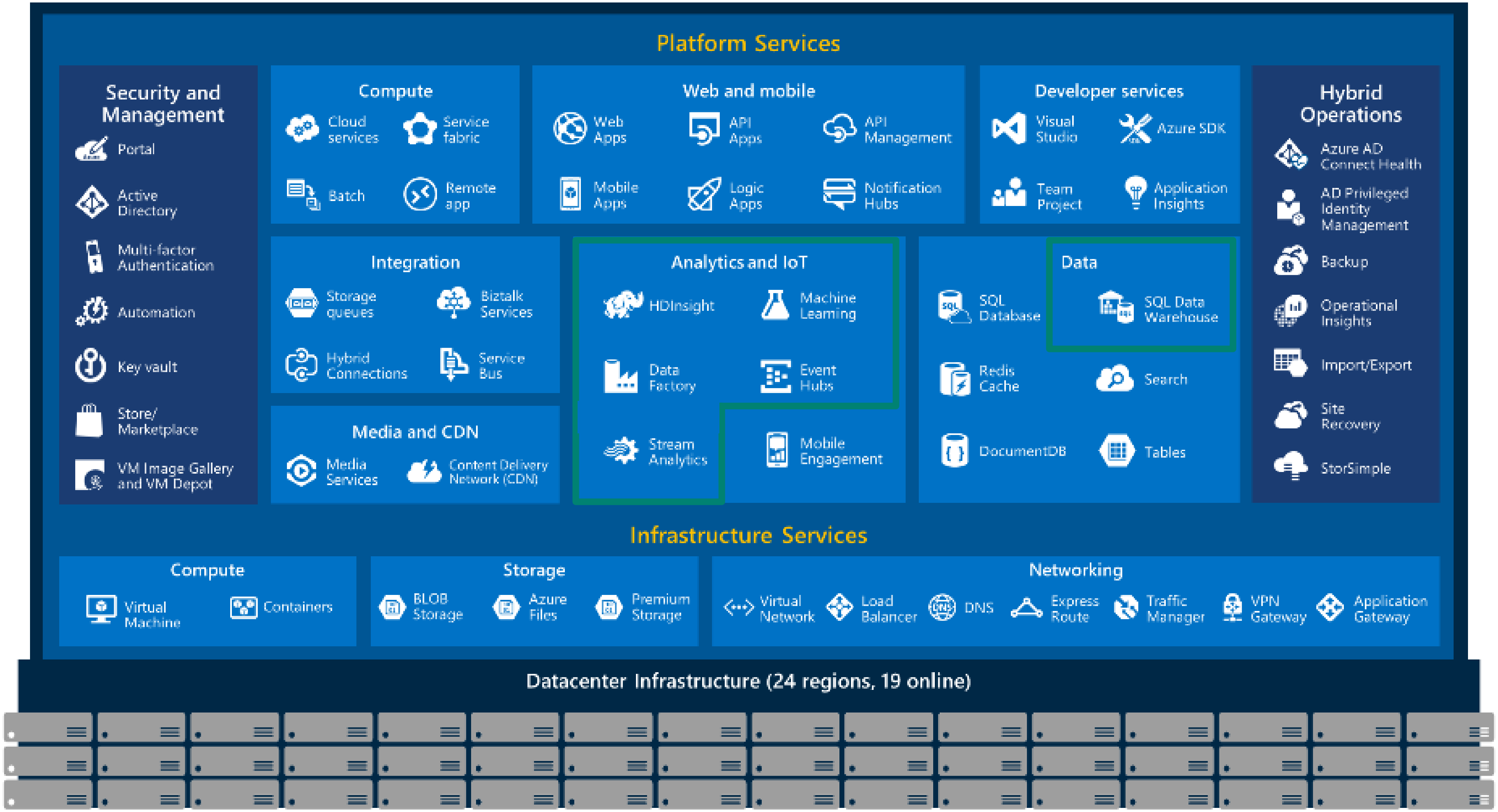
Getting the right products to the right places with **inventory management**

Offering customers exactly what they want, when they want it, with **personalization**

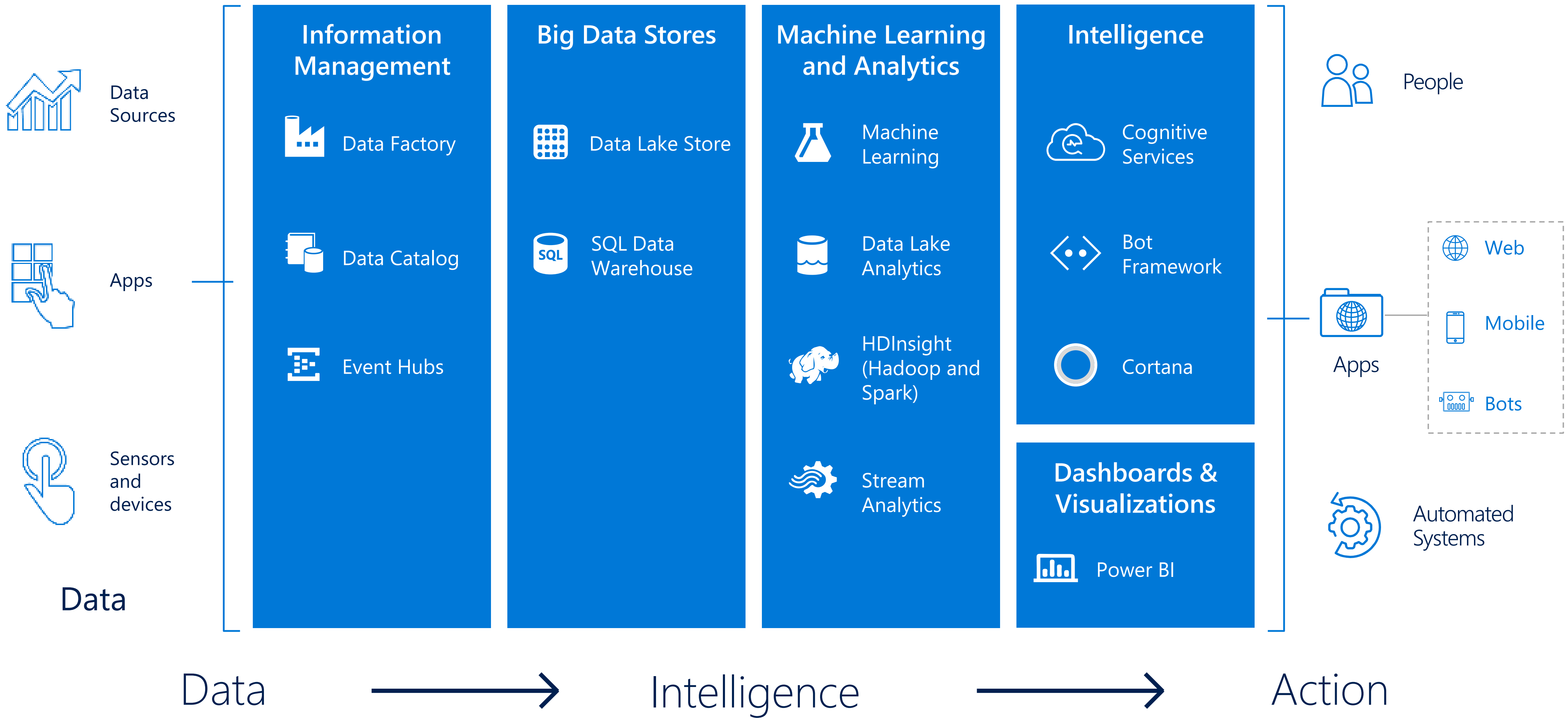
Fixing problems proactively before they start with **predictive maintenance**

Exploring new business opportunities with **data-driven services**

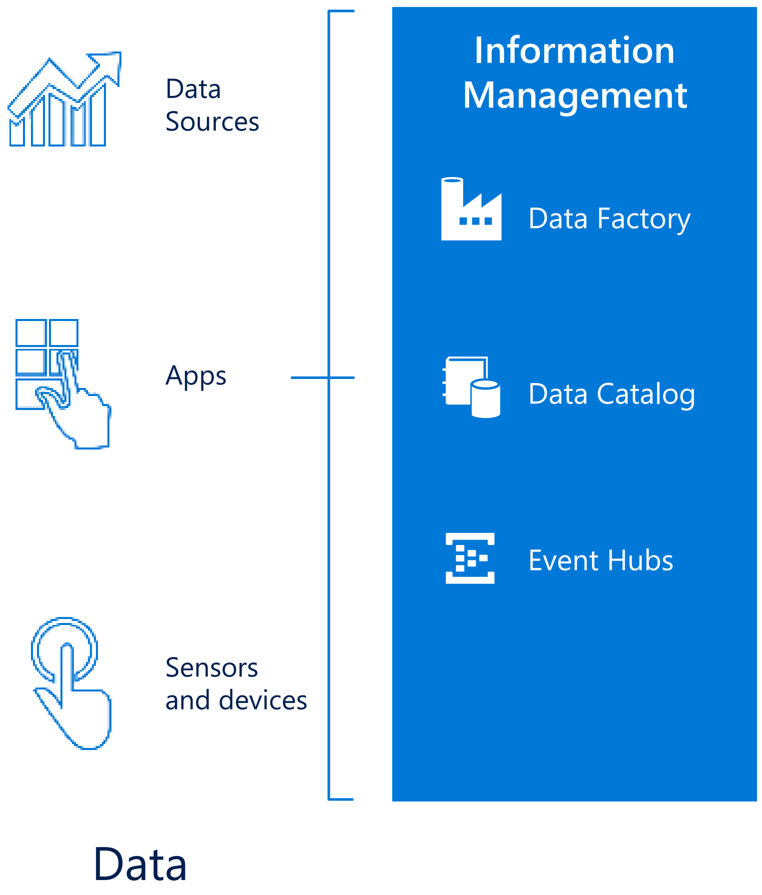
Cortana Intelligence combines the services you already know



Transform data into intelligent action



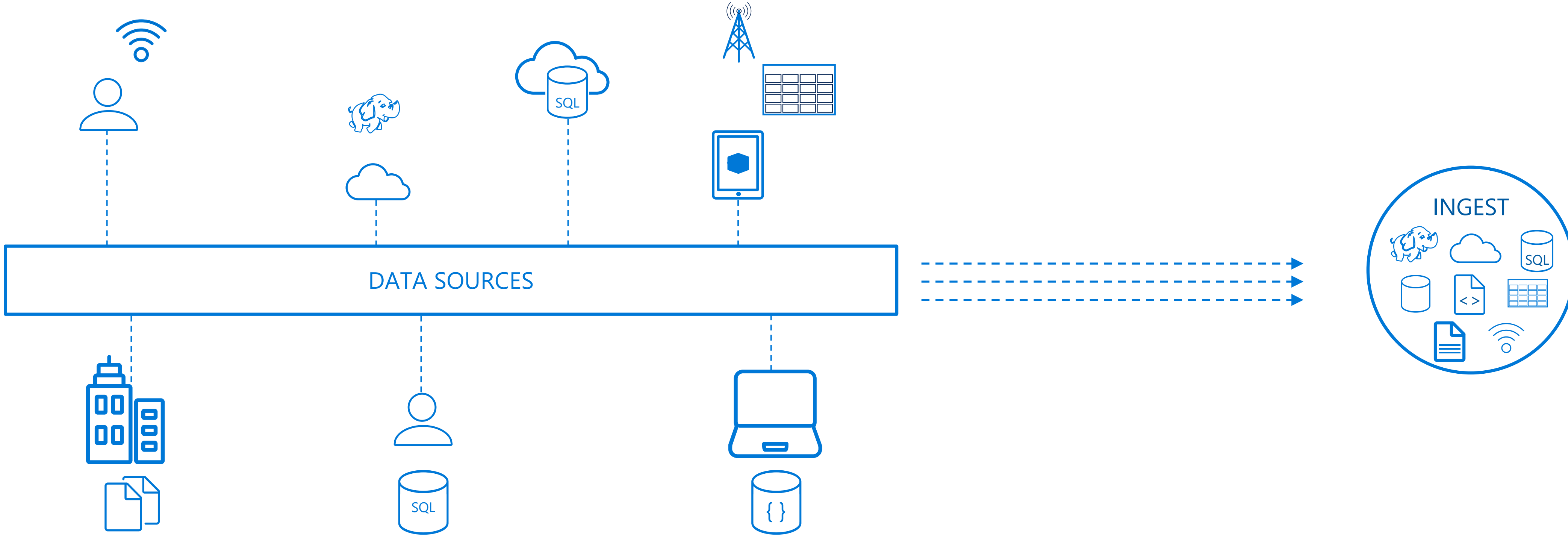
Information Management



Compose and orchestrate data services at scale

Information Management

- Data Factory
- Data Catalog
- Event Hubs



- Create, schedule, orchestrate, and manage data pipelines
- Visualize data lineage
- Connect to on-premises and cloud data sources
- Monitor data pipeline health
- Automate cloud resource management
- Move relational data for Hadoop processing
- Transform with Hive, Pig, or custom code

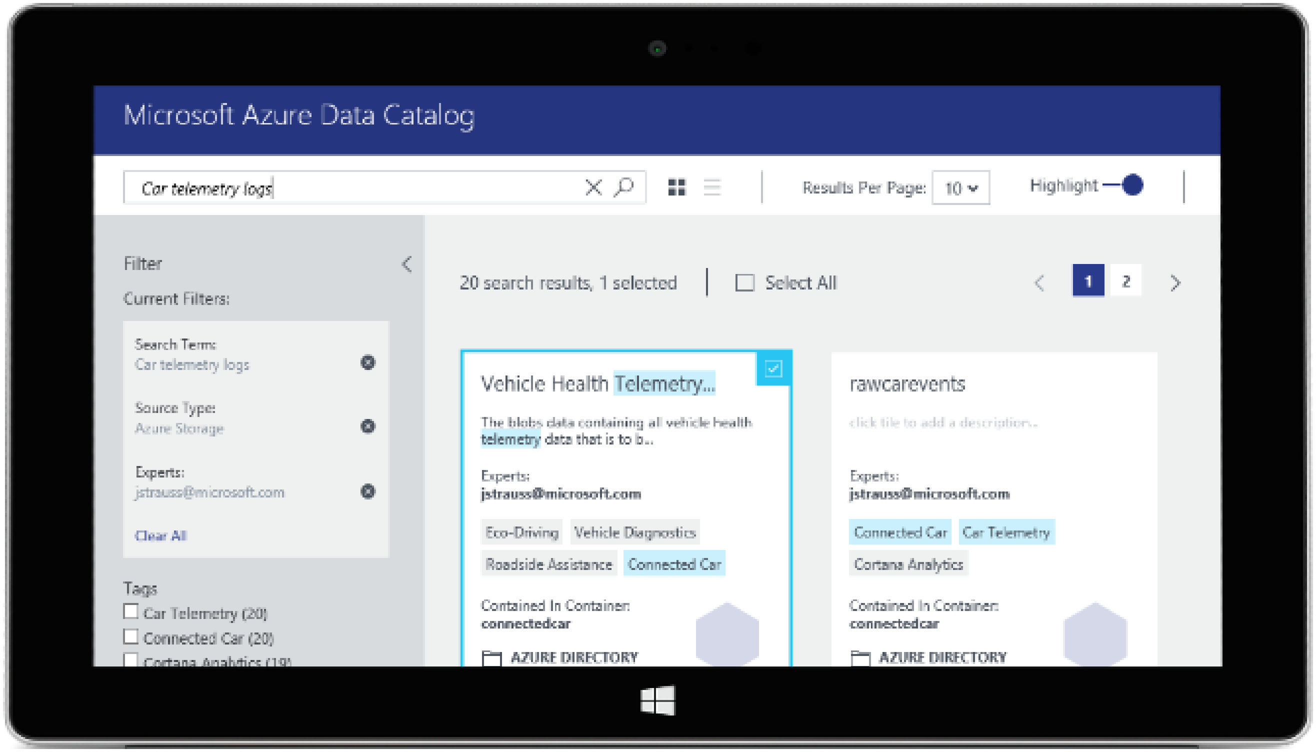
Get more value from your enterprise data assets

Information Management

Data Factory

Data Catalog

Event Hubs

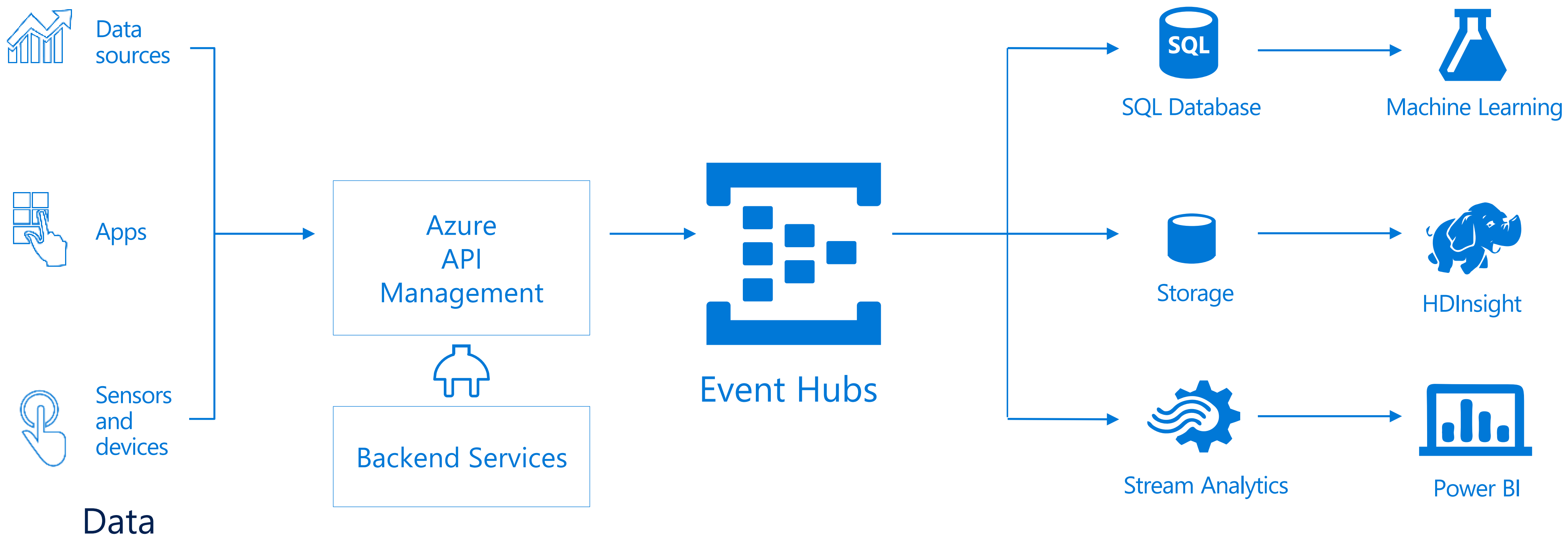


- Spend less time looking for data, and more time getting value from it
- Register enterprise data sources, discover data assets and unlock their potential, and capture tribal knowledge to make data understandable
- Bridge the gap between IT and the business, allowing everyone to contribute their insights, tags, and descriptions
- Intuitive search and filtering to understand the data sources and their purpose
- Let your data live where you want; connect using tools you choose
- Integrate into existing tools and processes with open REST APIs

Ingest events from websites, apps and devices at cloud scale

Information Management

- Data Factory
- Data Catalog
- Event Hubs



- Log millions of events per second in near real time
- Connect devices using flexible authorization and throttling
- Use time-based event buffering
- Get a managed service with elastic scale

- Get a managed service with elastic scale
- Reach a broad set of platforms using native client libraries
- Pluggable adapters for other cloud services

Big Data Stores

Data Sources

Apps

Sensors and devices

Data

Information Management

- Data Factory
- Data Catalog
- Event Hubs

Big Data Stores

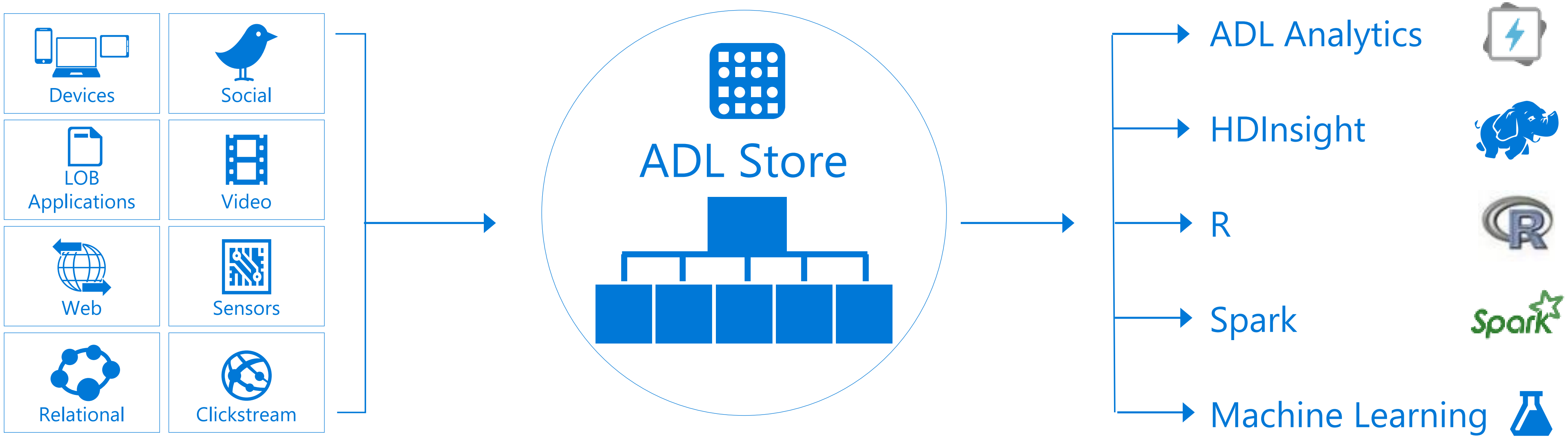
- Data Lake Store
- SQL Data Warehouse

A hyper-scale repository for big data analytics workloads

Big Data Stores

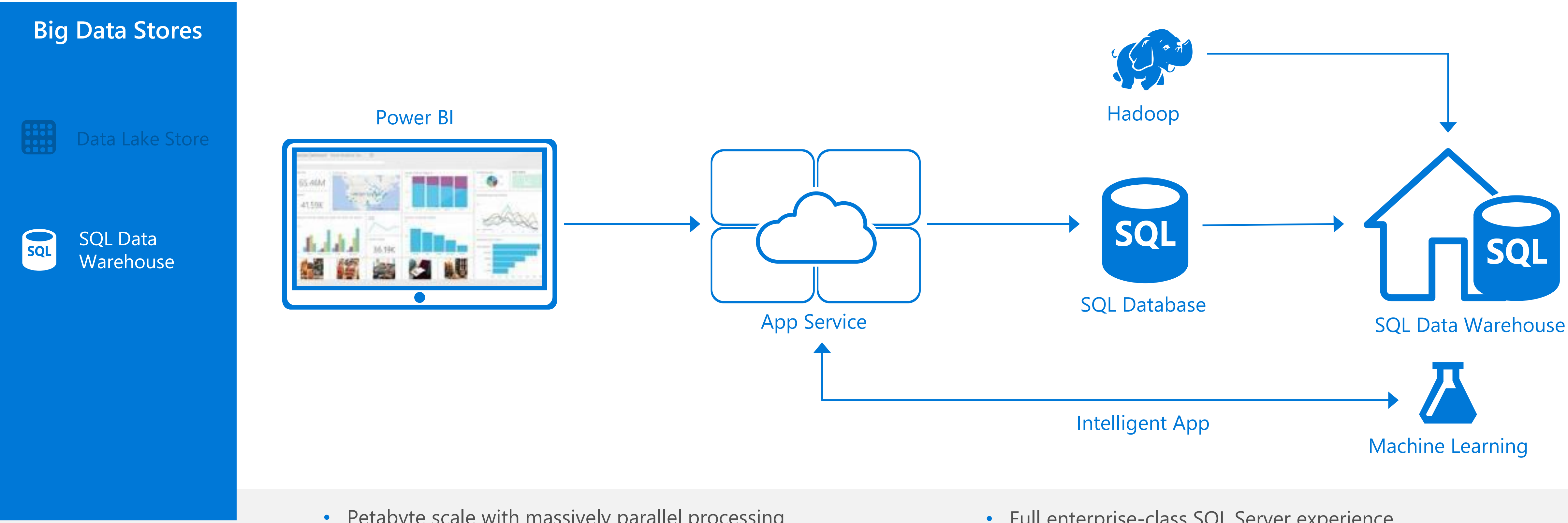
Data Lake Store

SQL Data Warehouse



- A Hadoop Distributed File System for the cloud
- No fixed limits on file size
- No fixed limits on account size
- Unstructured and structured data in their native format
- Massive throughput to increase analytic performance
- High durability, availability, and reliability
- Azure Active Directory access control

Elastic data warehouse as a service with enterprise-class features



- Petabyte scale with massively parallel processing
- Independent scaling of compute and storage—in seconds
- Transact-SQL queries across relational and non-relational data
- Full enterprise-class SQL Server experience
- Works seamlessly with Power BI, Machine Learning, HDInsight, and Data Factory

Machine Learning and Analytics

Data Sources

Apps

Sensors and devices

Information Management

- Data Factory
- Data Catalog
- Event Hubs

Big Data Stores

- Data Lake Store
- SQL Data Warehouse

Machine Learning and Analytics

- Machine Learning
- Data Lake Analytics
- HDInsight (Hadoop and Spark)
- Stream Analytics

a

Intelligence

Easily build, deploy, and share predictive analytics solutions

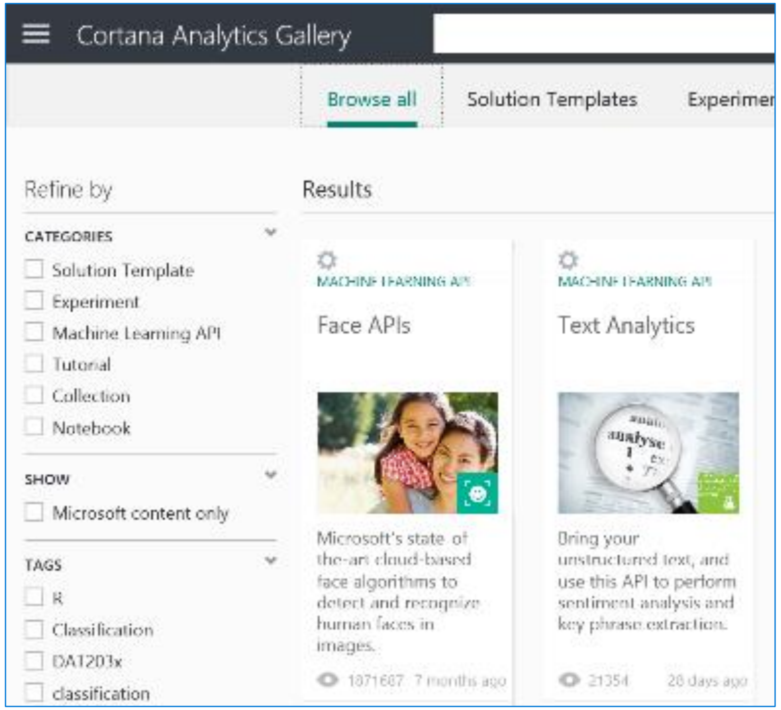
Machine Learning and Analytics

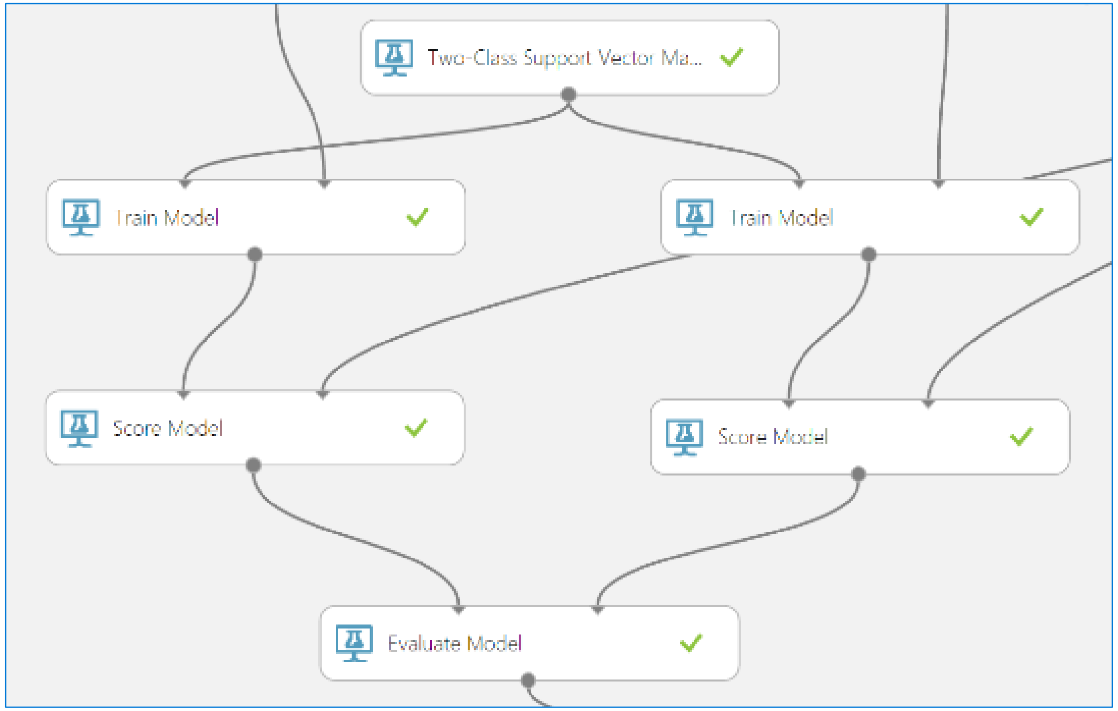
Machine Learning

Data Lake Analytics

HDInsight (Hadoop and Spark)

Stream Analytics



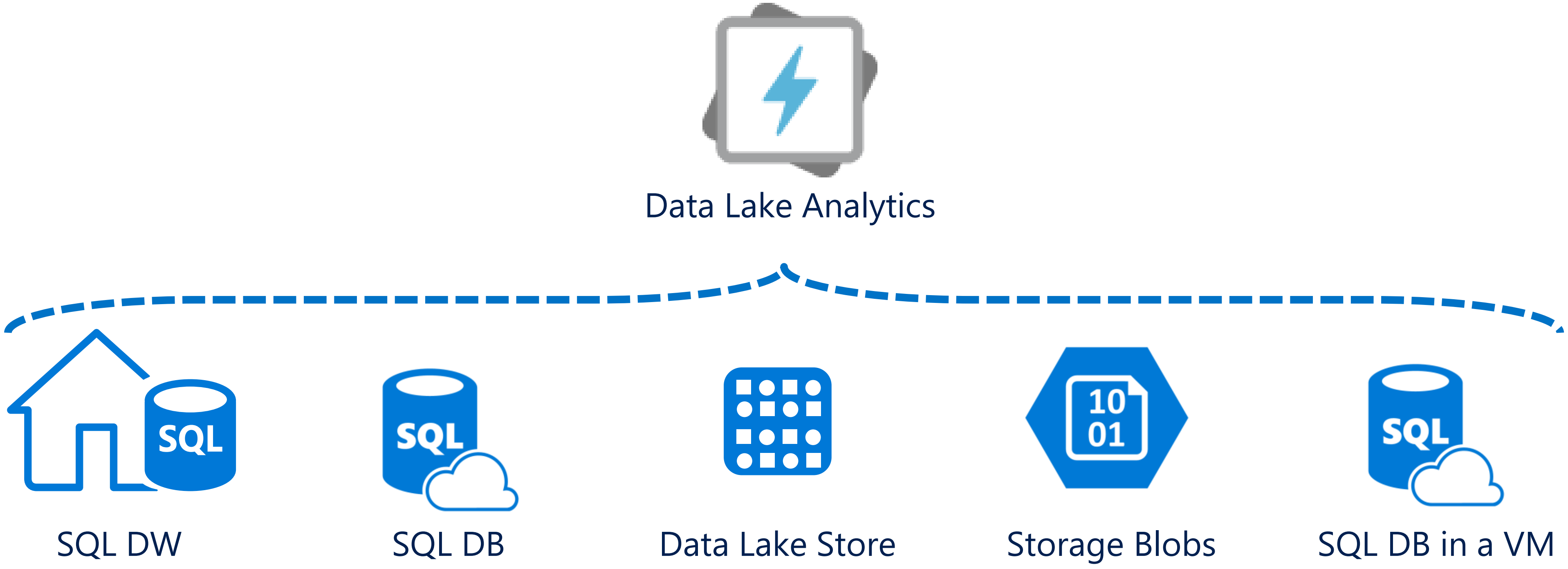


- Simple, scalable, cutting edge. A fully managed cloud service that enables you to easily build, deploy, and share predictive analytics solutions.
- Deploy in minutes. Azure Machine Learning means business. You can deploy your model into production as a web service that can be called from any device, anywhere and that can use any data source.
- Publish, share, monetize. Share your solution with the world in the Gallery or on the Azure Marketplace.

Big data analytics made easy

Machine Learning and Analytics

- Machine Learning
- Data Lake Analytics
- HDInsight (Hadoop and Spark)
- Stream Analytics



- Analyze data of any kind and size
- Develop faster, debug and optimize smarter
- Interactively explore patterns in your data
- No learning curve—use U-SQL, Spark, Hive, HBase and Storm
- Managed and supported with an enterprise-grade SLA
- Dynamically scales to match your business priorities
- Enterprise-grade security with Azure Active Directory
- Built on YARN, designed for the cloud

Comprehensive set of managed Apache big data projects

Machine Learning and Analytics

- Machine Learning
- Data Lake Analytics
- HDInsight (Hadoop and Spark)
- Stream Analytics

Batch	**Script**	**SQL**	**NoSQL**	**Streaming**	**In-Memory**
Map Reduce	Pig	Hive	HBase	Storm	Spark

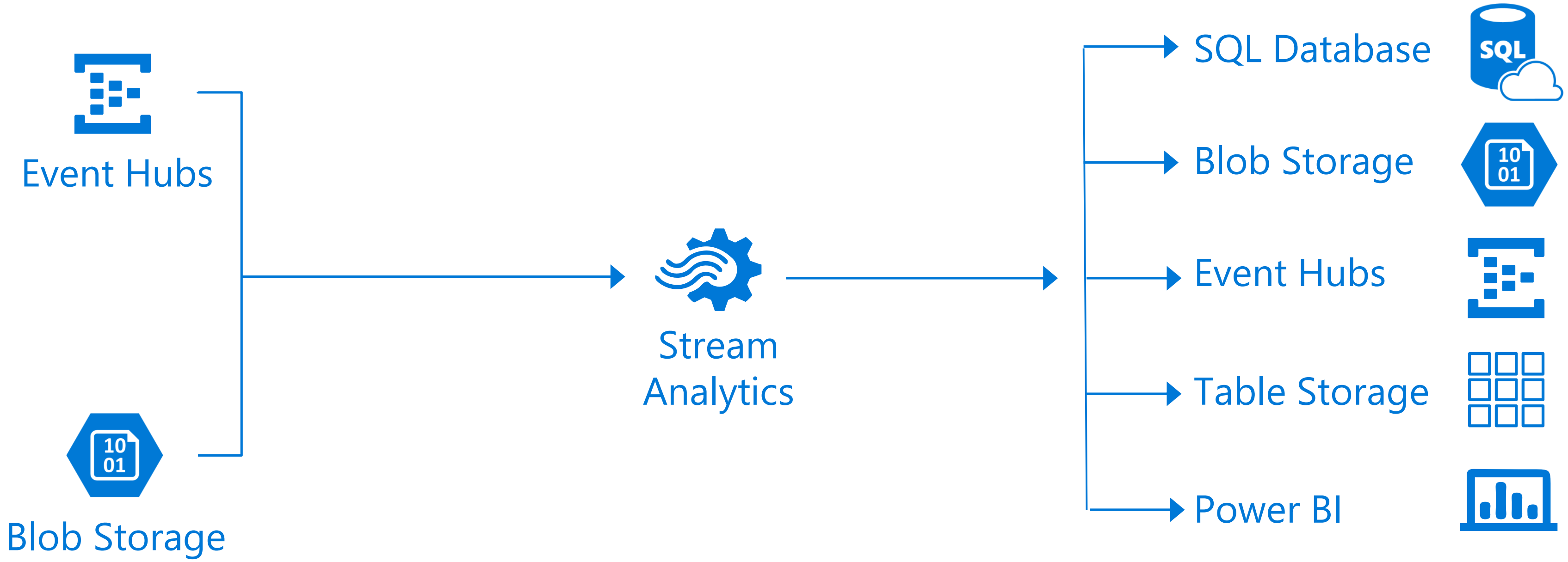
Core Engine

- Scale to petabytes on demand
- Process unstructured and semi-structured data
- Develop in Java, .NET, and more
- Skip buying and maintaining hardware
- Deploy in Windows or Linux
- Spin up an Apache Hadoop cluster in minutes
- Visualize your Hadoop data in Excel
- Easily integrate on-premises Hadoop clusters

Real-time stream processing in the cloud

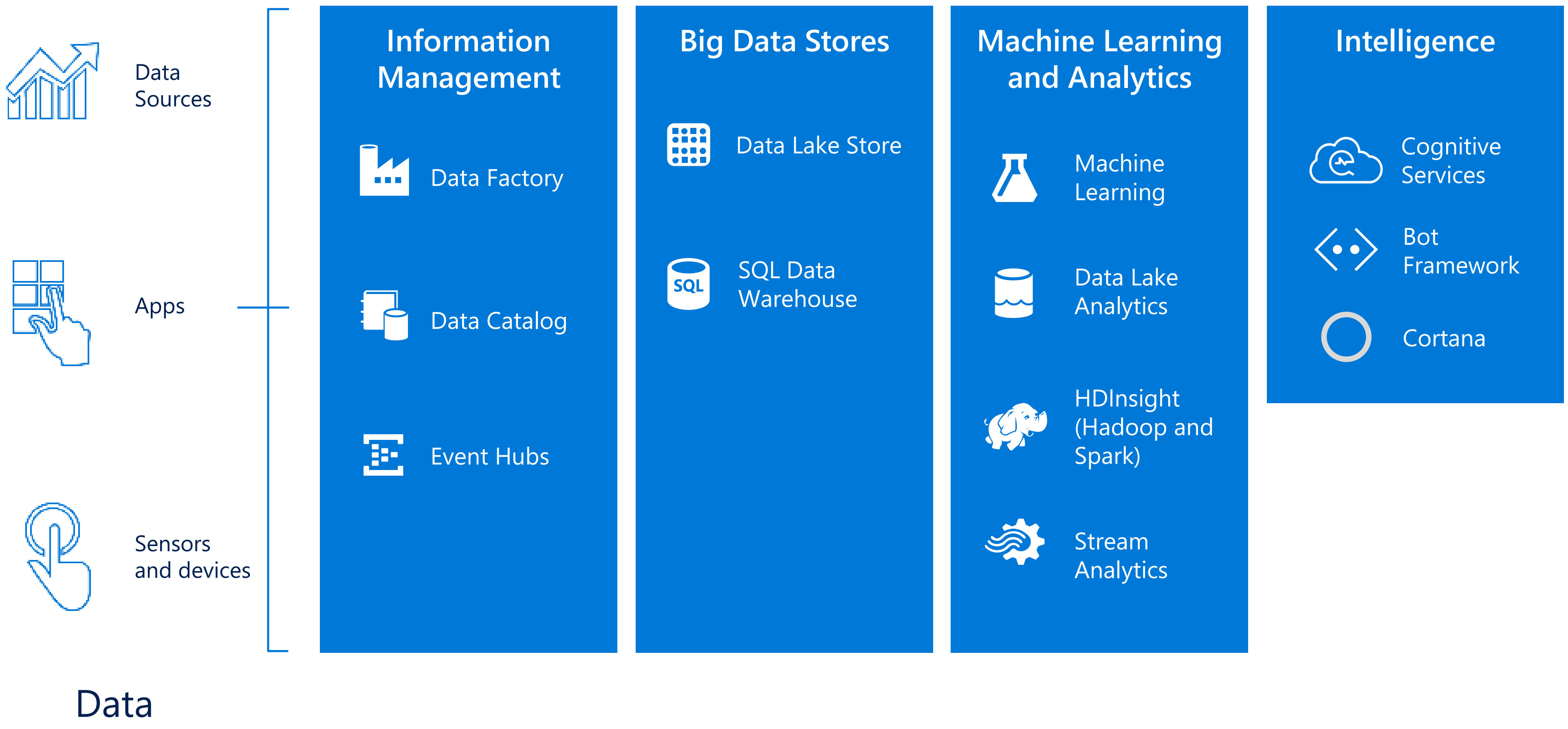
Machine Learning and Analytics

- Machine Learning
- Data Lake Analytics
- HDInsight (Hadoop and Spark)
- Stream Analytics



- Perform real-time analytics for your Internet of Things solutions
- Stream millions of events per second
- Get mission-critical reliability and performance with predictable results
- Create real-time dashboards and alerts over data from devices and applications
- Correlate across multiple streams of data
- Use familiar SQL-based language for rapid development

Intelligence



Build applications that understand people

Intelligence

- Cognitive Services
- Bot Framework
- Cortana

Vision	Speech	Language	Knowledge	Search
Computer Vision	Speaker Recognition	Text Analytics	Academic Knowledge	Bing Search API
Face	Speech	Bing Speller	Entity Linking Service	Bing Image Search API
Emotion	CRIS	Web Language Model	Knowledge Exploration Service	Bing Video Search API
Video		Linguistic Analysis	Recommendations	Bing News Search API
		Language Understanding Intelligent Service		Bing Auto Suggest API

- Faces, images, emotion recognition and video intelligence
- Spoken language processing, speaker recognition, custom speech recognition
- Natural language processing, sentiment and topics analysis, spelling errors
- Complex tasks processing, knowledge exploration, intelligent recommendations
- Bing engine capabilities for Web, Autosuggest, Image, Video and News

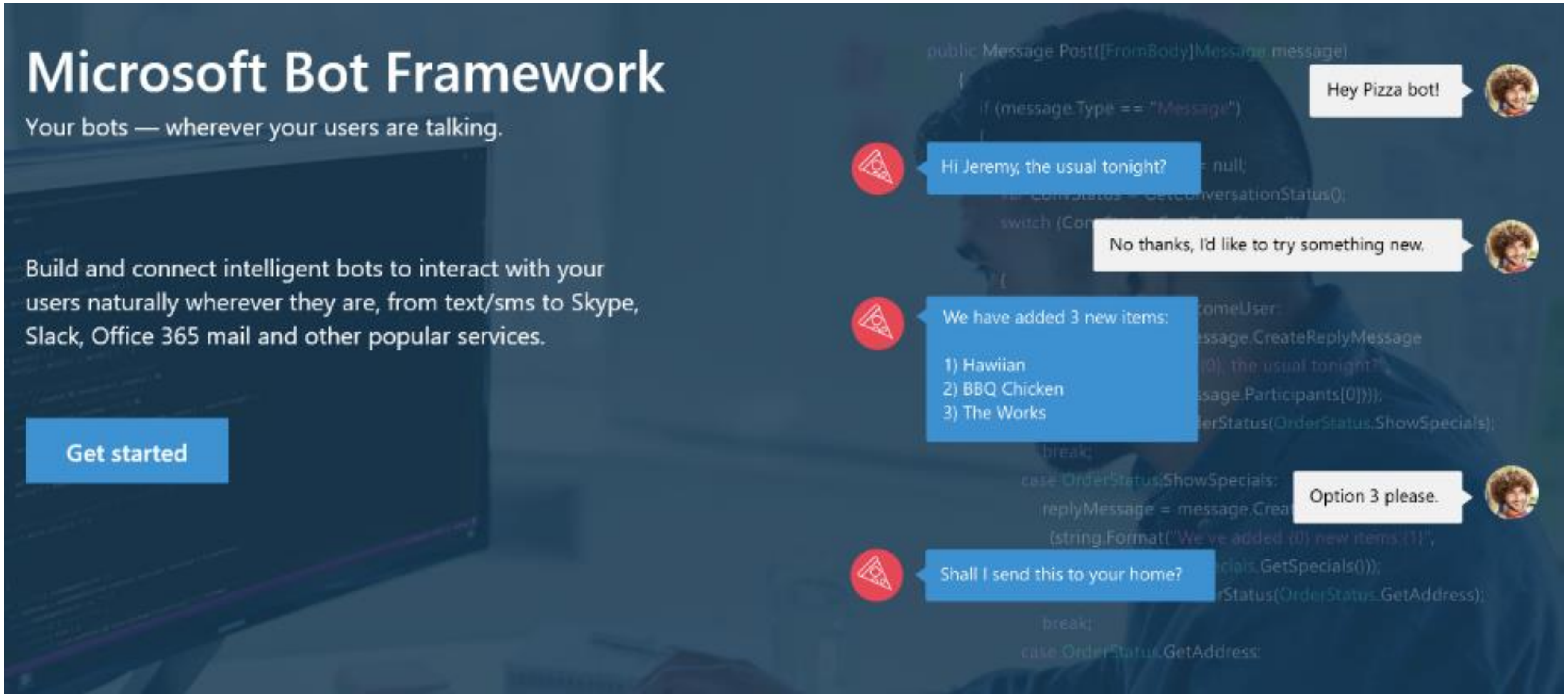
Your bots – wherever your users converse

Intelligence

Cognitive Services

Bot Framework

Cortana



- Bot Connector Service: A service to register your bot, configure channels and publish to the Bot Directory. Connect your bot(s) seamlessly to text/sms, Office 365 mail, Skype, Slack, Twitter, and more.
- Bot Builder SDK: An open source SDK hosted on GitHub. Everything you need to build great dialogs within your Node.js or C# bot
- Bot Directory: A public directory of bots registered through the Bot Connector Service. Discover, try, and add bots to conversation experiences

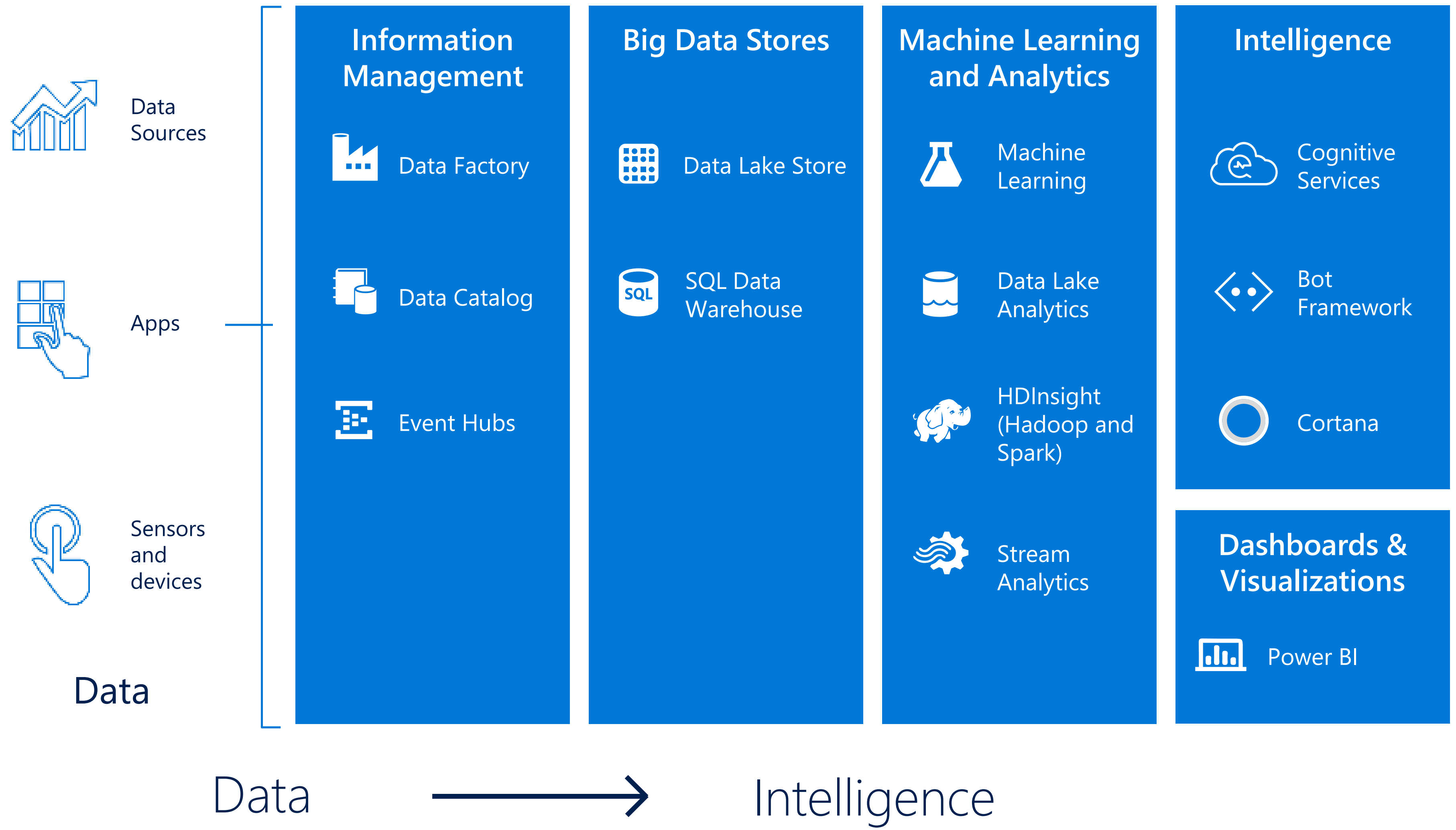
Get things done in more helpful, proactive and natural ways

Intelligence

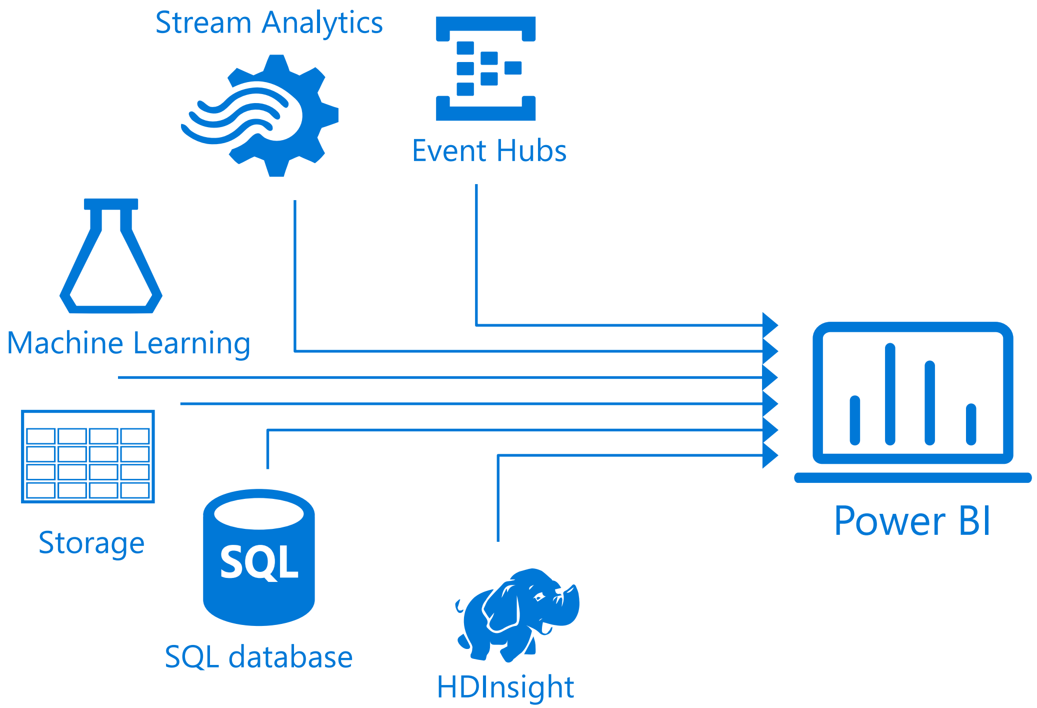
- Cognitive Services
- Bot Framework
- Cortana

Here are some of the things I can help you with…	Cortana for Consumers (today)	With the Cortana Intelligence Suite
Answers	Public reference data answers – *"How far is it from Los Angeles to San Francisco?"*	Answers from organizational data in Power BI *"What were our biggest deals that closed last month?"*
Predictions	Event predictions – *"Who do you think is going to win the Germany Italy game?"*	Integration with prediction solutions *"Which of our customers are most likely to churn in the next quarter?"*
Monitoring & Alerts	Flight status, traffic conditions, changes in weather, …	Monitoring KPIs and preemptive alerting *"Alert me if this customer ever has a 90% chance of churn in the next 30 days"*
Task Completion	Setting reminders, scheduling meetings, getting directions, …	Line of business process integration *Assistance with expense report submission on-time within policy*

Dashboards & Visualizations



Keep a pulse on your business with live, interactive dashboards

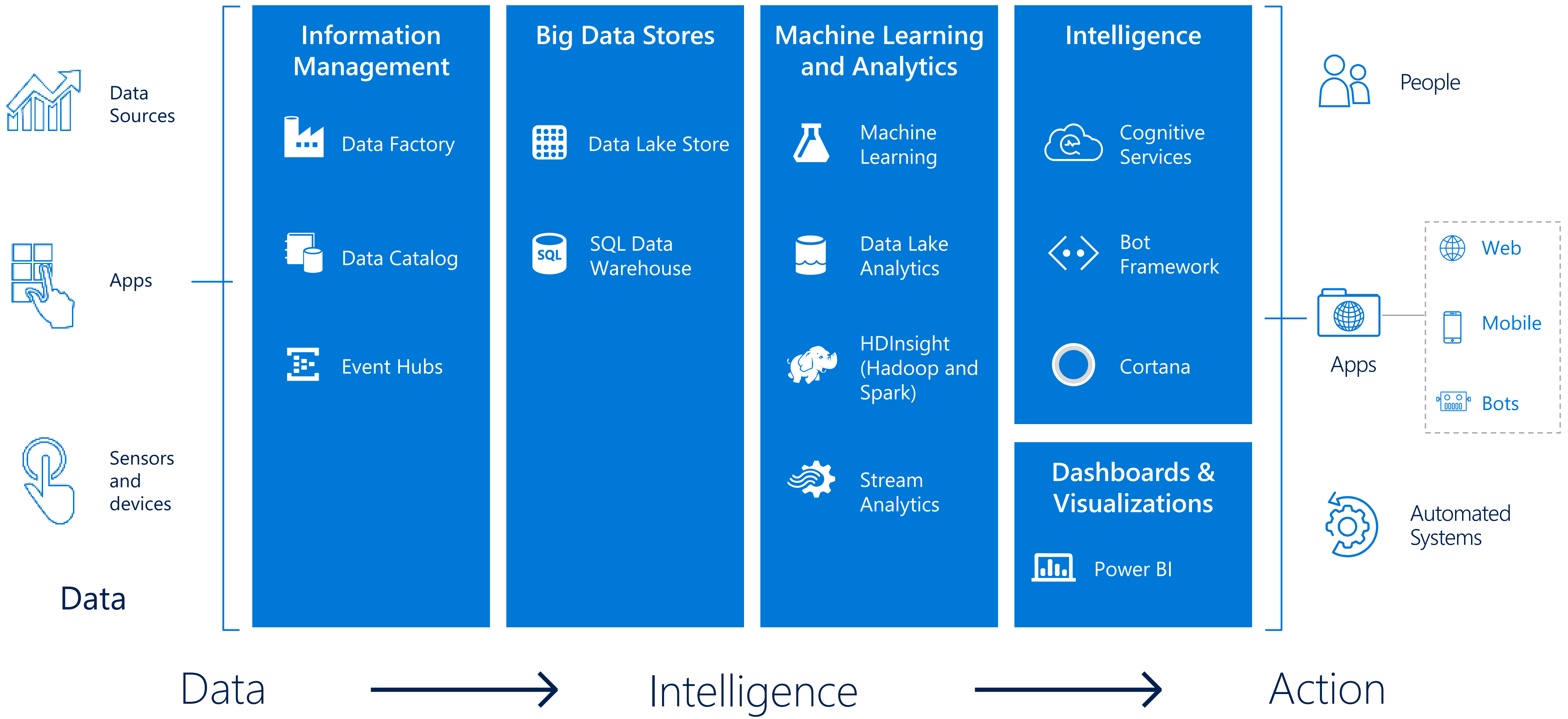


Dashboards & Visualizations

Power BI

- Analytics for everyone, even non-data experts
- Your whole business on one dashboard
- Create stunning, interactive reports
- Drive consistent analysis across your organization
- Embed visuals in your applications
- Get real-time alerts when things change

Transform data into intelligent action



The Cortana Intelligence Platform

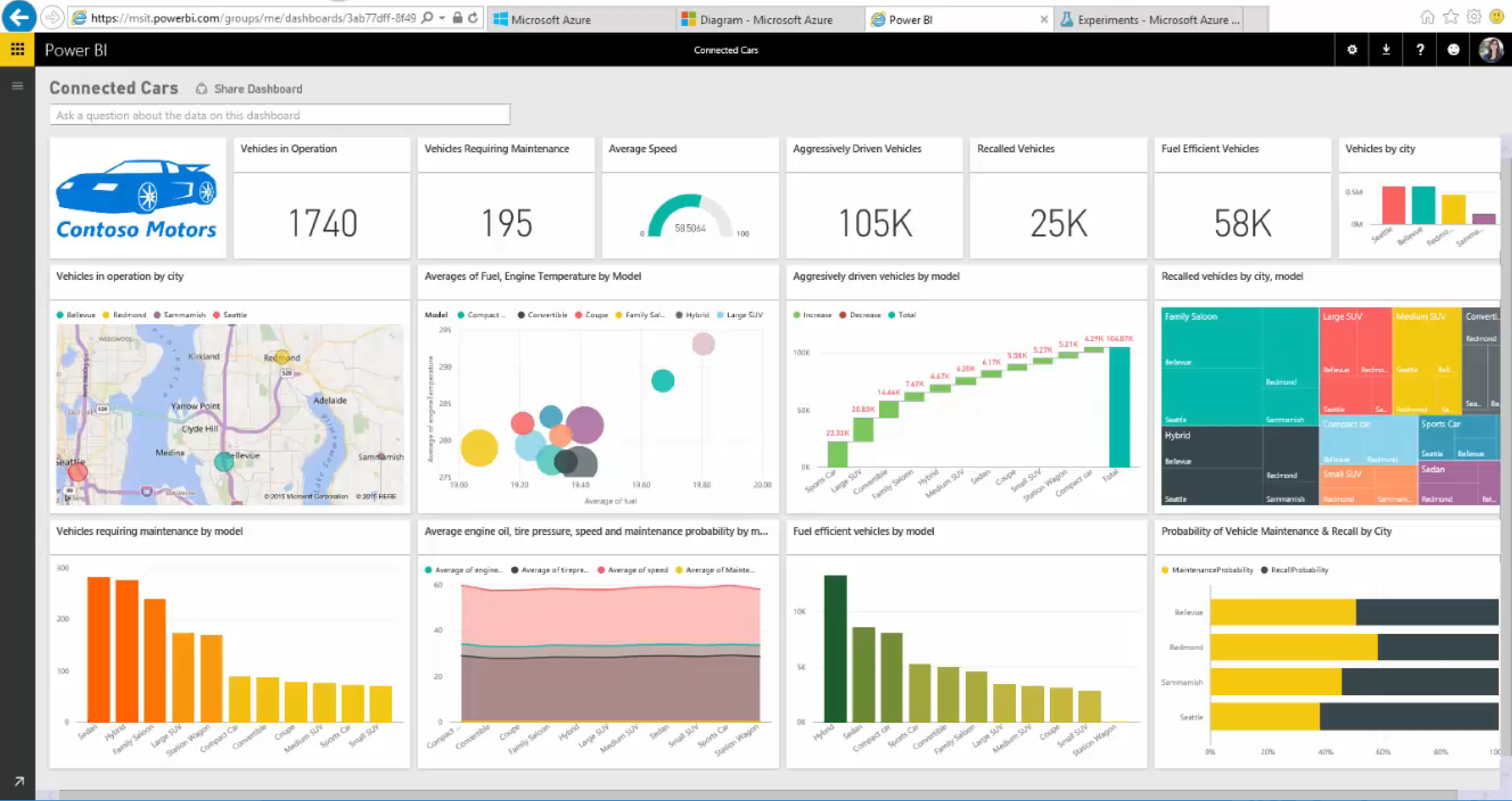
- Cortana, Cognitive Services, Bot Framework
- Power BI
- Azure Stream Analytics
- Azure HDInsight
- Azure Machine Learning and MRS
- Azure SQL DB, Data Warehouse, DocumentDB
- Azure Data Lake
- Azure Event Hubs
- Azure Data Factory
- Azure Data Catalog
- Microsoft Azure

Demo: Vehicle health and driving behavior patterns

Connected Cars

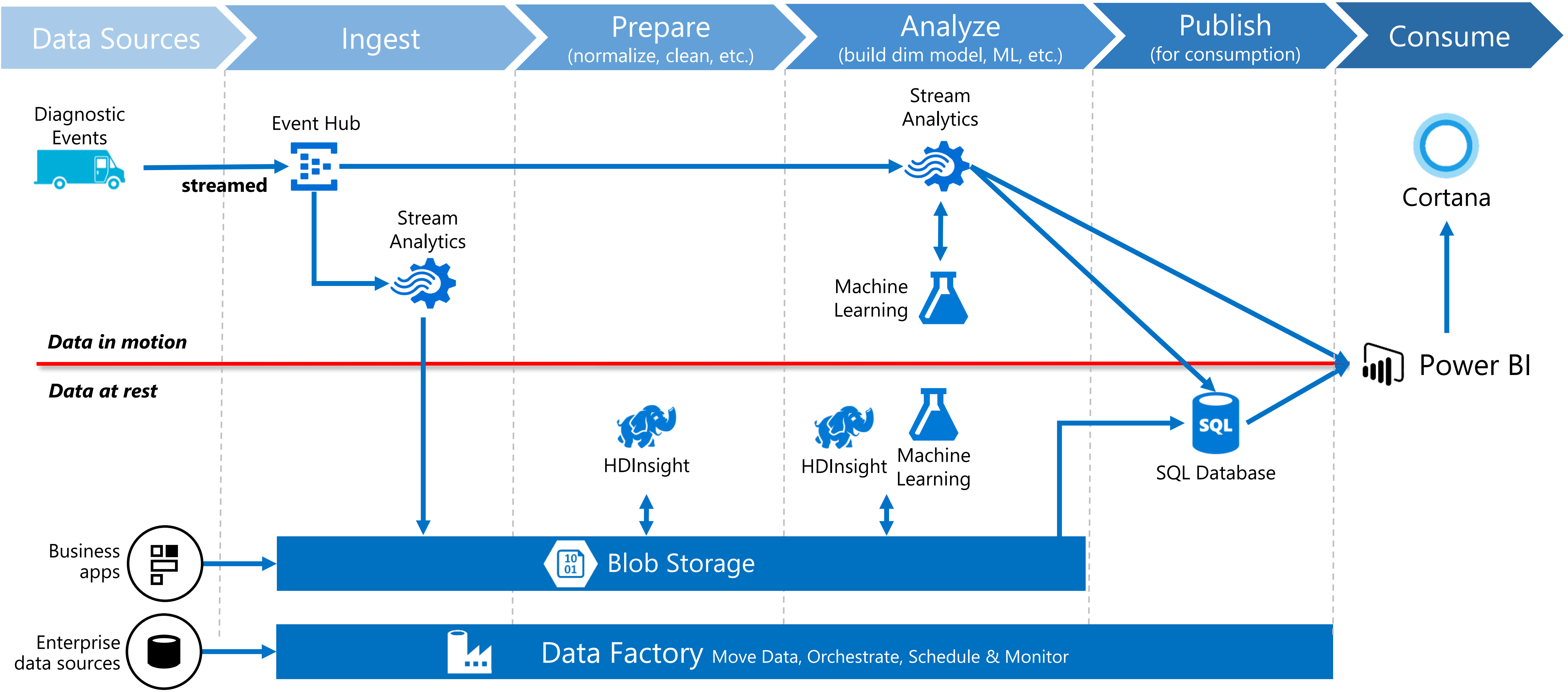
75% of the cars shipped globally by **2020** will be built with the necessary hardware to connect to the Internet

- Vehicle diagnostic
- Usage-based insurance
- Fleet management
- Roadside assistance
- Eco-driving
- Engine performance remapping
- Engine emission control


https://msit.powerbi.com/groups/me/dashboards/3ab77dff-8f49
Microsoft Azure
Diagram - Microsoft Azure
Power BI
Experiments - Microsoft Azure ...
Power BI
Connected Cars
Connected Cars
Share Dashboard
Ask a question about the data on this dashboard
Contoso Motors
Vehicles in Operation
1740
Vehicles Requiring Maintenance
195
Average Speed
58.5064
Aggressively Driven Vehicles
105K
Recalled Vehicles
25K
Fuel Efficient Vehicles
58K
Vehicles by city
Vehicles in operation by city
Bellevue
Redmond
Sammamish
Seattle
Averages of Fuel, Engine Temperature by Model
Average of fuel
Aggresively driven vehicles by model
Increase
Decrease
Total
Recalled vehicles by city, model
Family Saloon
Large SUV
Medium SUV
Hybrid
Compact car
Sports Car
Small SUV
Sedan
Vehicles requiring maintenance by model
Average engine oil, tire pressure, speed and maintenance probability by m...
Fuel efficient vehicles by model
Probability of Vehicle Maintenance & Recall by City
MaintenanceProbability
RecallProbability

Microsoft

Connected Car – Solution Architecture



Big Picture for Hands On Labs

Big Picture – Predictive Maintenance

Preventive Maintenance	Condition-based Maintenance	Predictive Maintenance
Regular and routine inspection	Conditions relevant to attrition	Optimal point in time for maintenance

"Sensor data that are used to predict when equipment is wearing down or needs repair can reduce maintenance costs by as much as 40 percent and cut unplanned downtime in half."

McKinsey Global Institute report, *The Internet of Things: Mapping the value beyond the hype*, June 2015

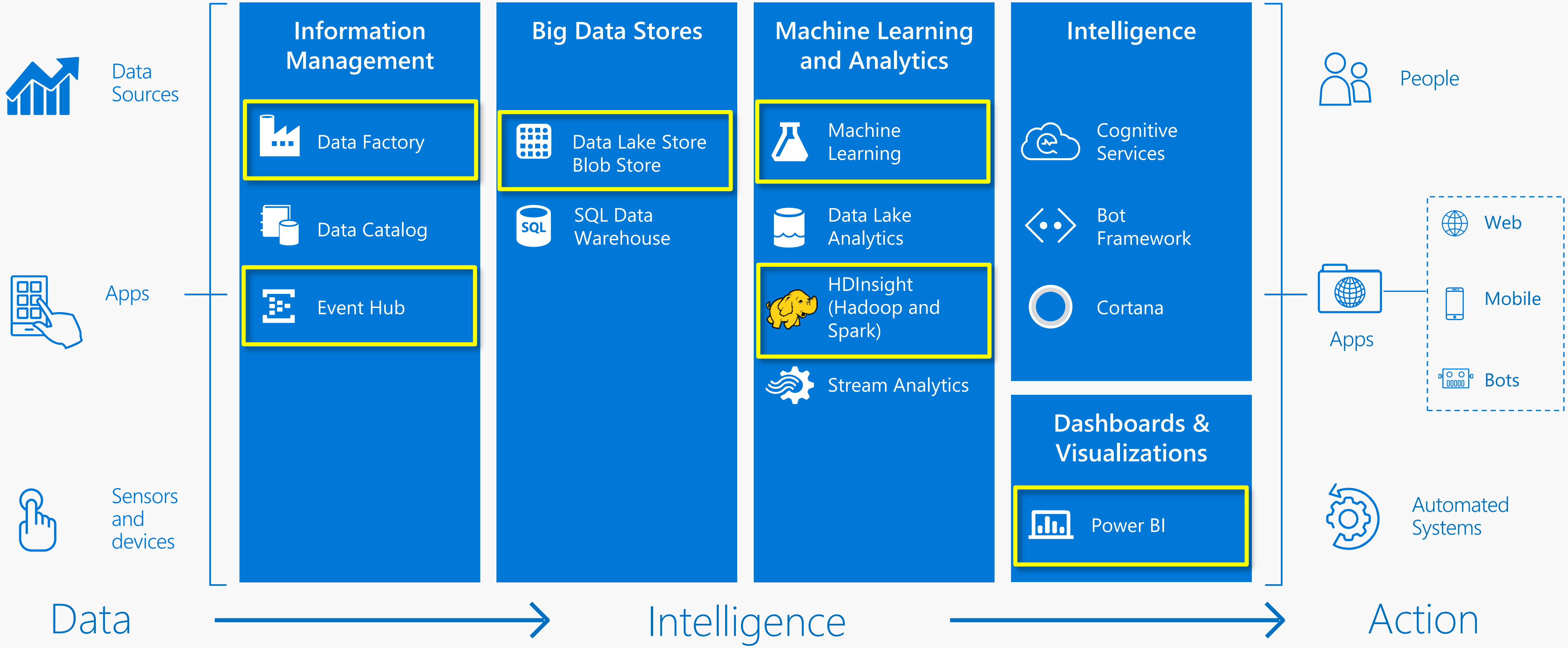
Predictive Maintenance Use Cases

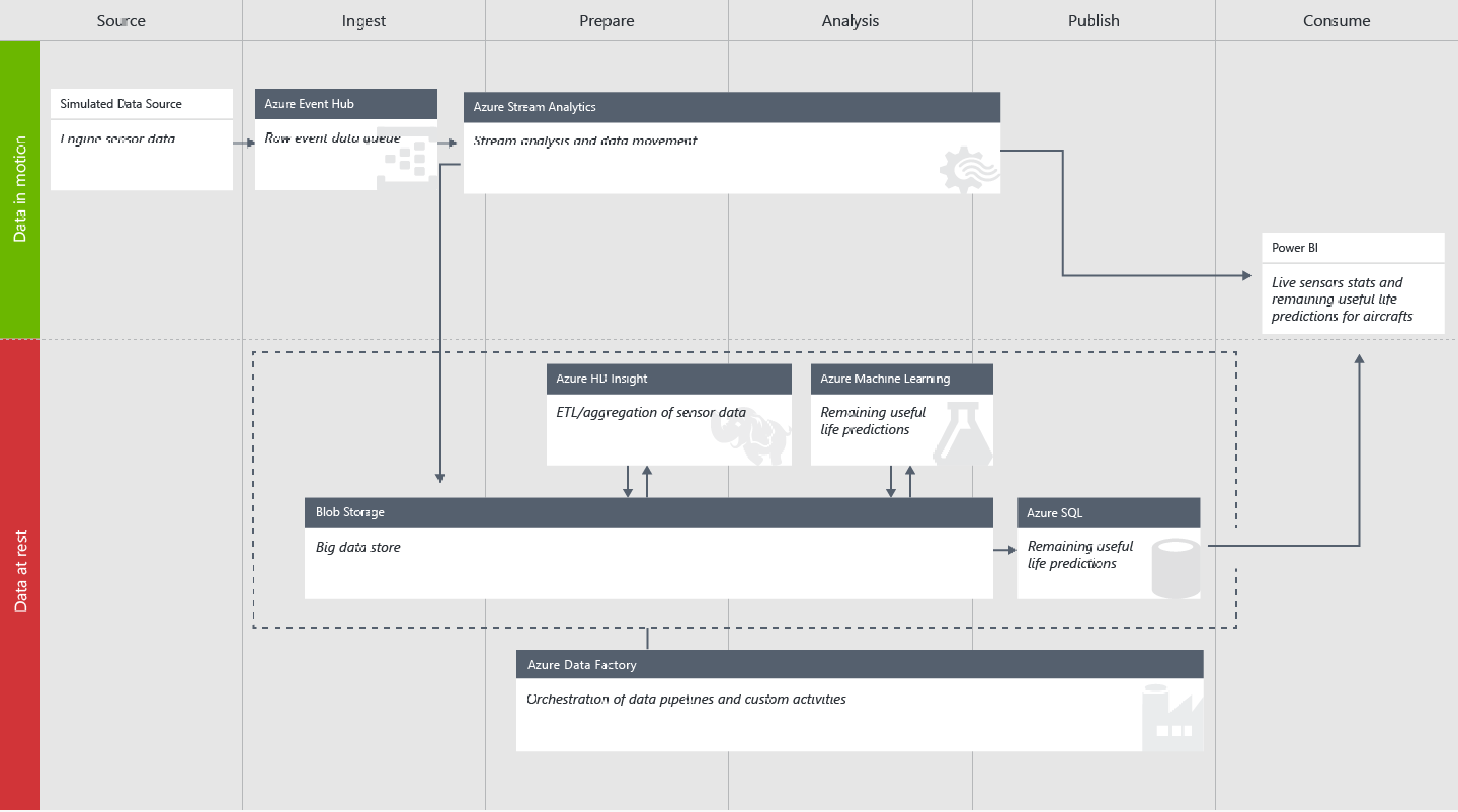
Aerospace	Utilities	Manufacturing	Transportation & Logistics
What is the likelihood of delay due to mechanical issues?	When is my solar panel or wind turbine going to fail next?	Will the component pass the next stage of testing on factory floor or do I need to rework?	Should I replace the break disks in my car or can I wait for another month?
When is this aircraft component likely to fail next?	Which circuit breakers in my system are likely to fail in the next month?	What is the root cause of the test failure?	What maintenance task should I perform on my elevator?
	Is the ATM going to dispense the next 5 notes without failing?		

Cortana Intelligence Suite

Make machine learning *accessible* to
every enterprise, data scientist, developer,
information worker, consumer,
and device anywhere in the world.

Cortana Intelligence Suite




Source
Ingest
Prepare
Analysis
Publish
Consume
Data in motion
Data at rest
Simulated Data Source
Engine sensor data
Azure Event Hub
Raw event data queue
Azure Stream Analytics
Stream analysis and data movement
Power BI
Live sensors stats and remaining useful life predictions for aircrafts
Azure HD Insight
ETL/aggregation of sensor data
Azure Machine Learning
Remaining useful life predictions
Blob Storage
Big data store
Azure SQL
Remaining useful life predictions
Azure Data Factory
Orchestration of data pipelines and custom activities

gallery.cortanaanalytics.com/SolutionTemplate/Predictive-Maintenance-for-Aerospace-1

Cortana Analytics Gallery

Sign in

Browse all | Solution Templates | Experiments | Machine Learning APIs | Notebooks | Tutorials | Collections

SOLUTION TEMPLATE

Predictive Maintenance for Aerospace

Microsoft • published on February 16, 2016

Deploy

Preview

+ Add to Collection

1475 views

506 deployments

Tweet | in Share

SERVICES USED

- Azure Event Hubs
- Azure Stream Analytics
- Azure Machine Learning
- Azure Data Factory
- Azure HDInsight
- Azure Blob Storage
- Azure SQL
- PowerBI

Summary

Air travel is central to modern life, however, aircraft engines are expensive and keeping them up and running requires frequent maintenance by highly skilled technicians. Modern aircraft engines are equipped with highly sophisticated sensors to track the functioning of these machines. By combining the data from these sensors with advanced analytics it's possible to both monitor the aircraft in real time, as well as predict the remaining useful life of an engine component so that maintenance can be scheduled in a way to prevent mechanical failures.

The Cortana Analytics Predictive Maintenance for Aerospace Solution Template monitors aircraft and predicts the remaining useful life of aircraft engine components.

Description

Cortana Analytics opens new possibilities in the predictive maintenance space, including data ingestion, data storage, data processing and advanced analytics components—all the essential elements for building a predictive maintenance solution. While this solution is customised for aircraft monitoring, it can very easily be generalised for other predictive maintenance scenarios.

The solution template uses several Azure services, such as Event Hubs for ingesting aircraft sensor readings into Azure. Stream Analytics provides real-time insights on engine health and stores that data in long-term storage for more complex, compute-intensive batch analytics. HDInsight transforms the sensor

https://gallery.cortanaanalytics.com/SolutionTemplate/Predictive-Maintenance-for-Aerospace-1

GitHub, Inc. [US] github.com/oliviak/technicalsummit-data/tree/master/predictivemaintenance

The guide below provides a full set of instructions on how to put together and deploy a predictive maintenance solution using the Cortana Intelligence Suite. The **Developer Journey Map** walks through the different components created as part of the end-to-end solution.

Hands On Lab: Predictive Maintenance

There is a lot of documentation around the Cortana Intelligence Suite Solution Template for predictive maintenance for aerospace that predicts the remaining useful life of an aircraft engine.

http://aka.ms/msts16-analytics

manual process gives an implementer an inside view on how the solution is built and an understanding of each of the components.

Requirements

This section contains required accounts and software you will need to create this solution.

1. The full contents of this GitHub repo (either cloned or downloaded as a zip file).
2. A Microsoft Azure subscription.
3. A Microsoft Office 365 subscription for Power BI access.
4. A network connection
5. SQL Server Management Studio OR another similar tool to access a SQL server database, e.g. Visual Studio Community or Visual Studio Code with the MSSQL extension
6. Microsoft Azure Storage Explorer

HOL 1: Azure Setup

HOL1: Azure Setup

Azure Subscription

Azure Resource Group

Azure Storage Account

On GitHub: Steps 1 & 2

2. Stream Processing

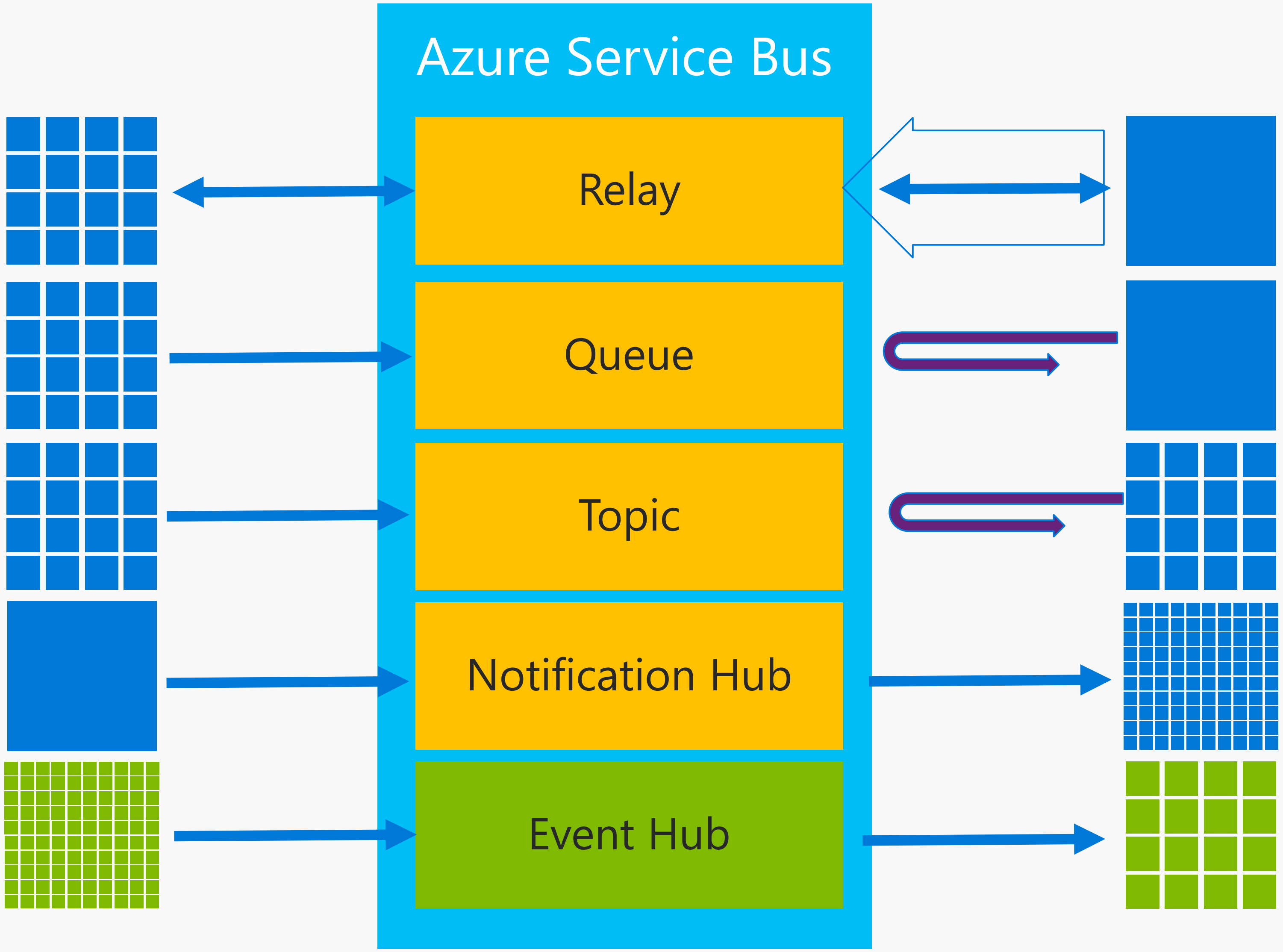
Streaming Data

Event Hub

Stream Analytics

Client App

Device Connectivity: Azure Service Bus



NAT and Firewall Traversal Service
Request/Response Services
Unbuffered with TCP Throttling

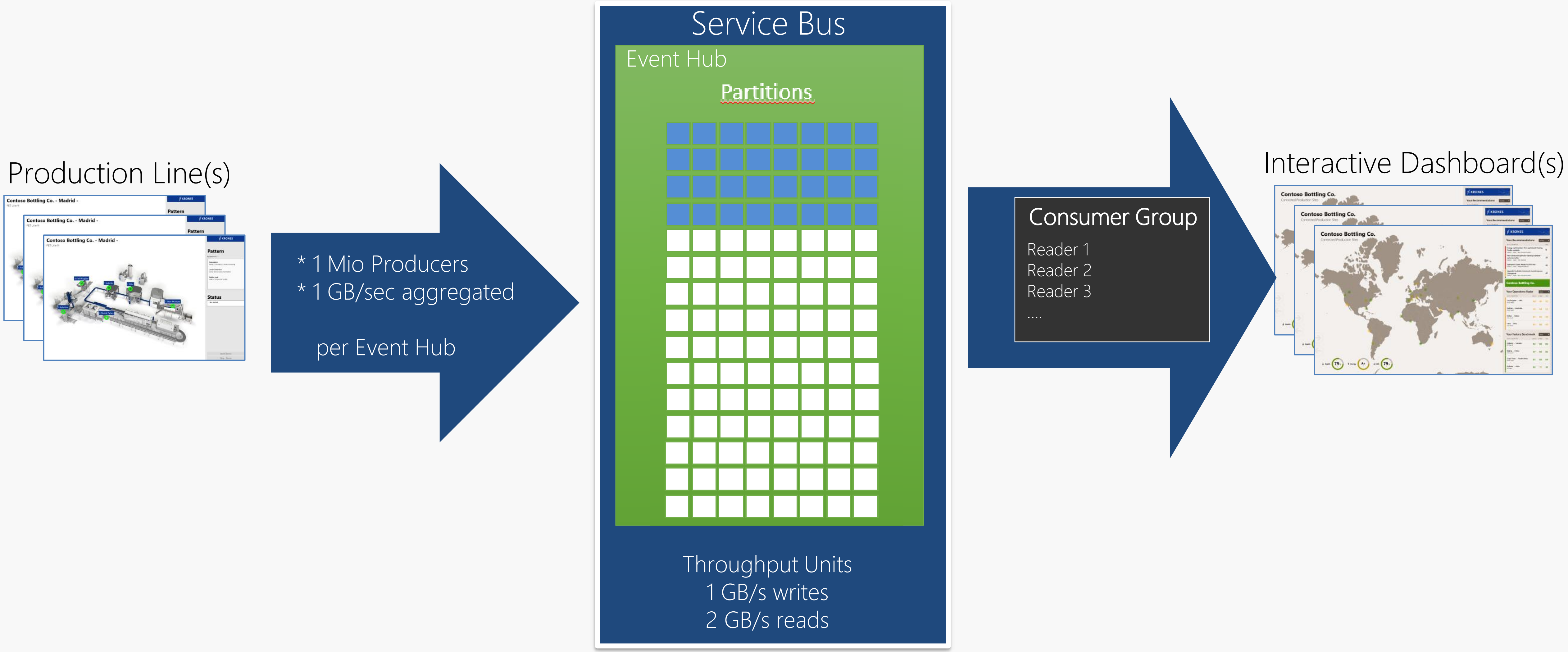
Transactional Cloud AMQP/HTTP Broker
High-Scale, High-Reliability Messaging
Sessions, Scheduled Delivery, etc.

Transactional Message Distribution
Up to 2000 subscriptions per Topic
Up to 2K/100K filter rules per subscription

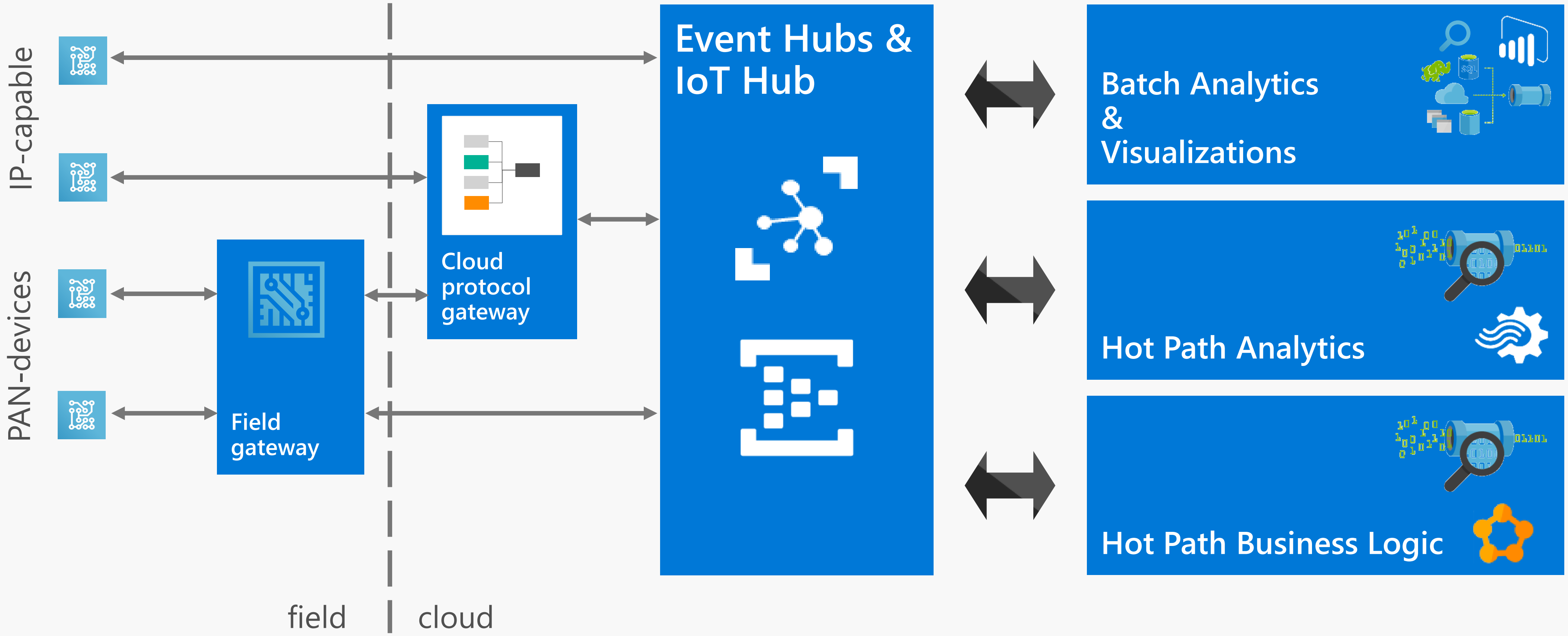
High-scale notification distribution
Most mobile push notification services
Millions of notification targets

… this session …

Data Ingest – Service Bus / Event Hub
Partitions



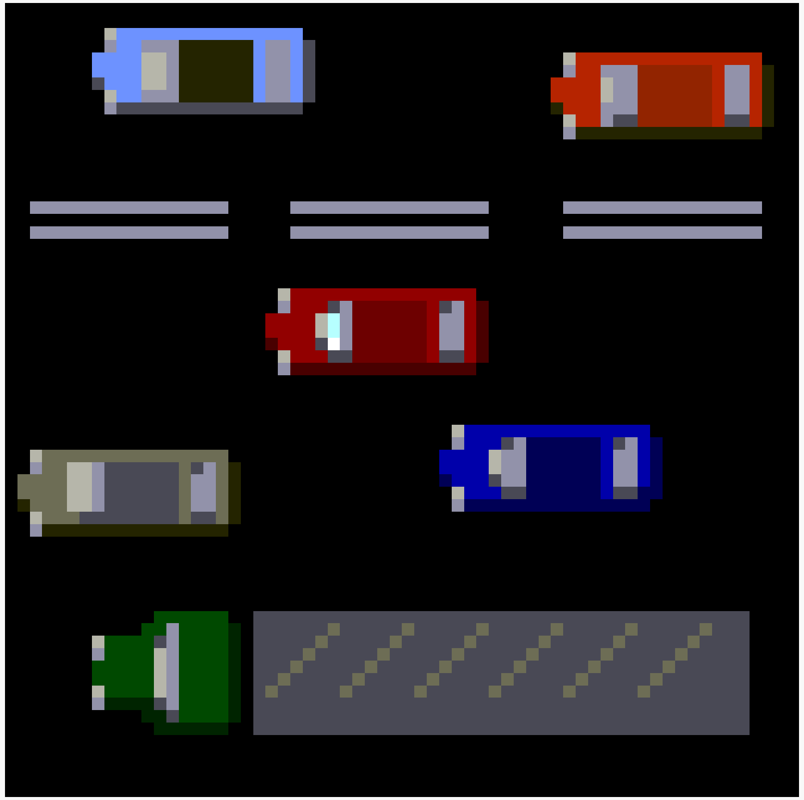
Connect your devices to Azure



What is Streaming Data?

Data at Rest

Data in Motion



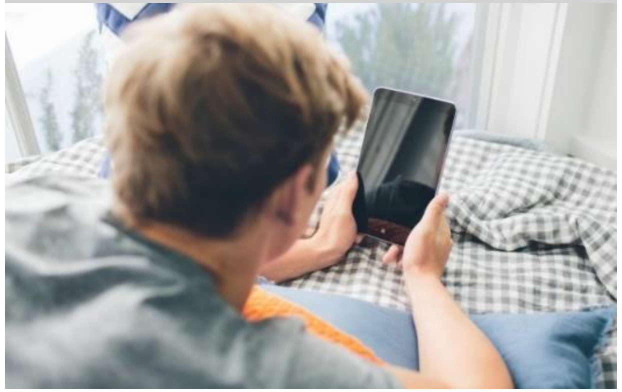
What are customers wanting to do?

Real-time fraud detection

Connected cars
Smart cities

Click-stream analysis

Real-time financial portfolio alerts



Smart grid

CRM alerting sales to customer case

Data and identity protection services

Smart retail

ThyssenKrupp

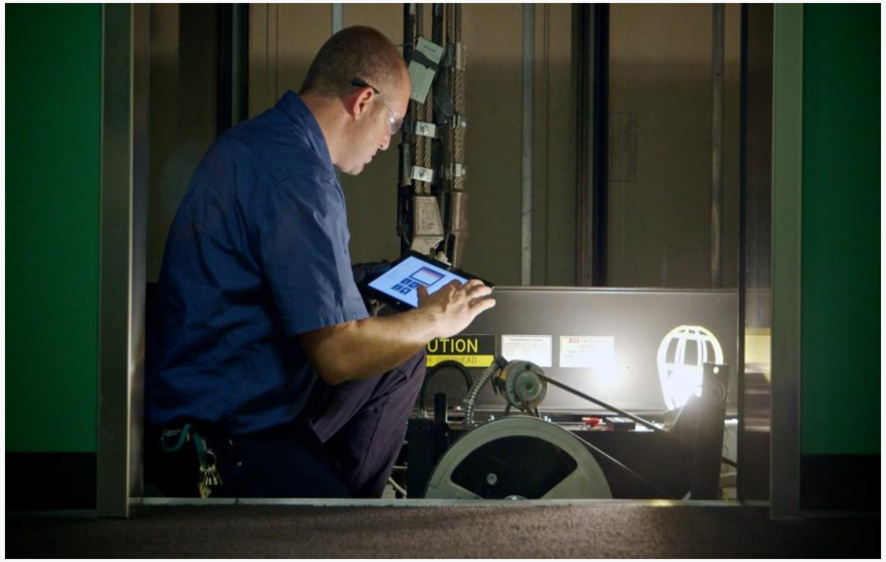
Differentiated Customer Service and Bottomline improvement through Preventive Maintenance and Connection Operations

Event Hubs

Stream Analytics

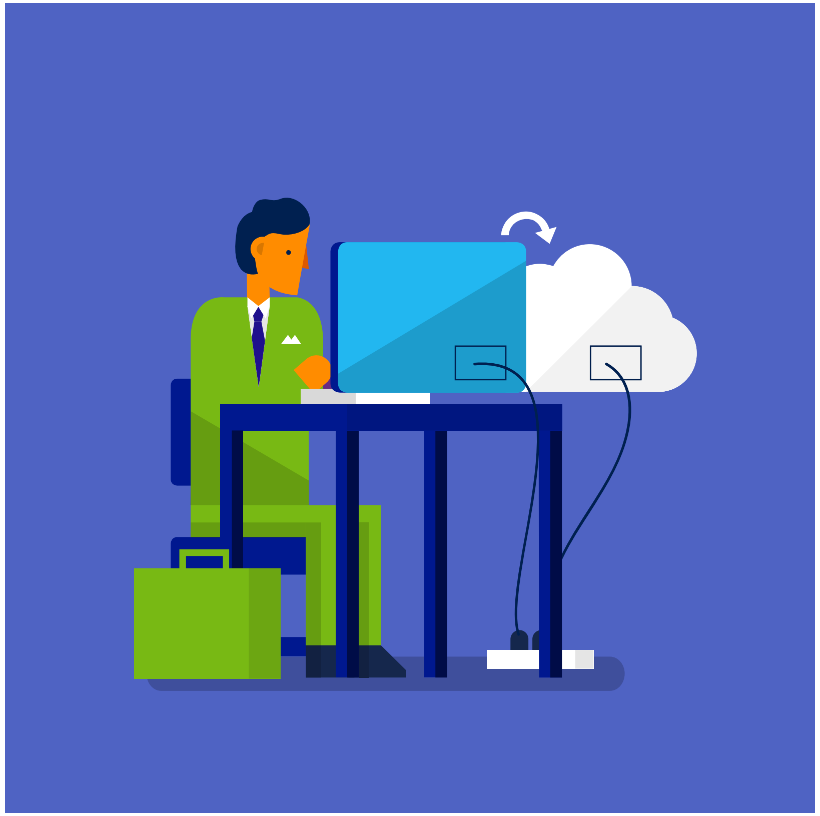
Machine Learning

Alerts & Dashboard helping timely and efficient scheduling of the technicians

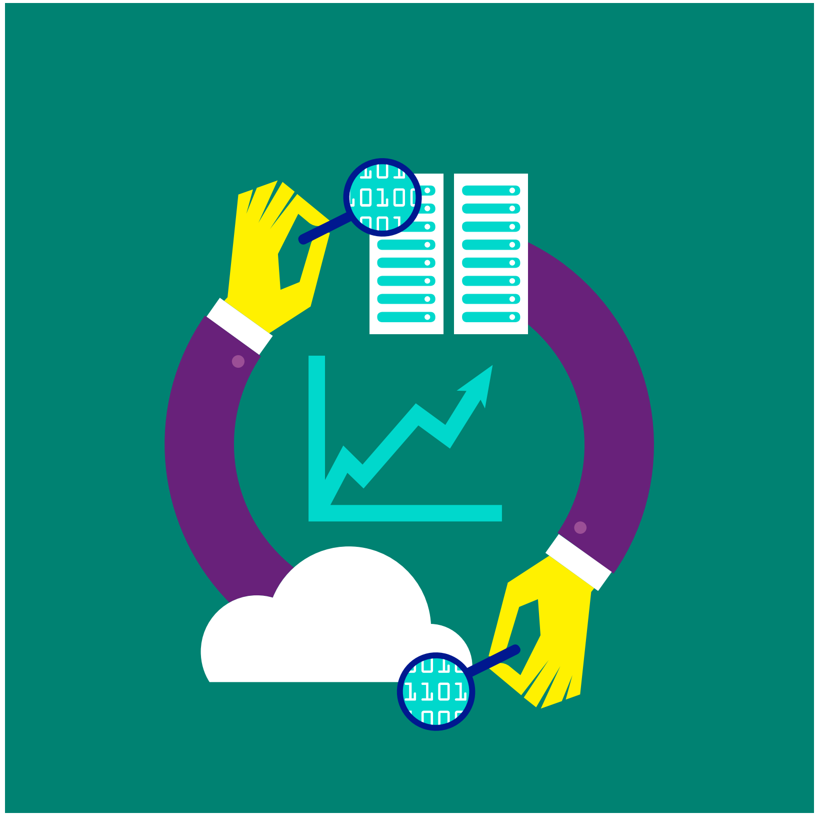


> "We wanted to go beyond the industry standard of preventative maintenance, to offer predictive and even *preemptive* maintenance, so we can guarantee a higher uptime percentage on our elevators."
>
> — CEO of the Company

Canonical scenarios



ARCHIVING



DASHBOARDING

TRIGGERING WORKFLOWS

Introducing Azure Stream Analytics

Fully managed real-time analytics

Mission critical reliability and scale

Enables rapid development

Real-time analytics

Fully managed
real-time analytics

Real-time Analytics

- Intake **millions of events per second** (up to 1 GB/s)
- **Low processing latency**, auto adaptive (sub-second to seconds)
- **Correlate** between different streams, or with reference data
- Find **patterns** or lack of patterns in data in real-time

Fully Managed Cloud Service

- No hardware acquisition and maintenance
- No platform/infrastructure deployment and maintenance
- Easily **expand your business globally** leveraging Azure regions

Mission critical

Mission critical reliability and scale

Mission Critical Reliability

- **Guaranteed event delivery**
- **Guaranteed business continuity**: Automatic and fast recovery

Effective Audits

- **Privacy and security** properties of solutions are evident
- Azure integration for **monitoring and ops alerting**

Easy To Scale

- **Scale** from small to large on demand

Rapid development

Enables rapid development

Rapid Development with SQL like language

- **High-level**: focus on stream analytics solution
- **Concise**: less code to maintain
- **Fast test**: Rapid development and debugging
- First-class support for **event streams and reference data**

Built in temporal semantics

- Built-in **temporal windowing and joining**
- Simple **policy configuration** to manage out-of-order events and late arrivals

SAQL – Language & Library

DML

- SELECT
- FROM
- WHERE
- GROUP BY
- HAVING
- CASE WHEN THEN ELSE
- INNER/LEFT OUTER JOIN
- UNION
- CROSS/OUTER APPLY
- CAST
- INTO
- ORDER BY ASC, DSC

Scaling Extensions

- WITH
- PARTITION BY
- OVER

Date and Time Functions

- DateName, DatePart, Day
- Month, Year, DateDiff
- DateTimeFromParts, DateAdd

Mathematical Functions

- ABS, CEILING, EXP, FLOOR
- POWER, SIGN, SQUARE
- SQRT

Temporal Functions

- Lag, IsFirst, Last
- CollectTop

Windowing Extensions

- TumblingWindow
- HoppingWindow
- SlidingWindow

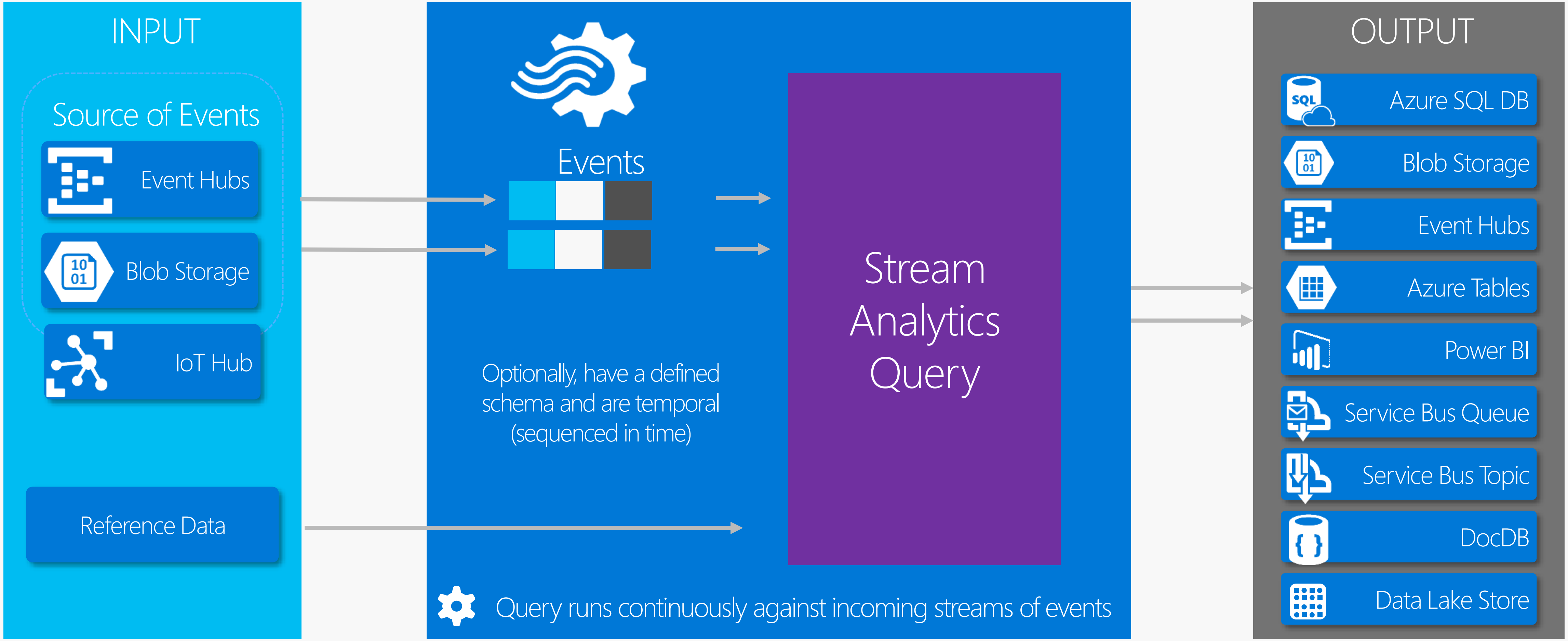
Aggregate Functions

- SUM
- COUNT
- AVG
- MIN
- MAX
- STDEV
- STDEVP
- VAR
- VARP
- TopOne

String Functions

- Len, Concat, CharIndex
- Substring, Lower
- Upper, PatIndex

Azure Stream Analytics Components



Let's count tweets by topic…

```
SELECT count(*), Topic FROM Tweets
GROUP BY Topic, TumblingWindow(second, 5)
```

That's all.
Just 2 (very short) lines of code.

Definitions

Payload/Data

CreatedAt="2015-04-30T20:45:30"
Topic="XBox"
SentimentScore=2
Text="Oh, joy! More forced @Xbox .."
….

Events

System.Timestamp=
2015-04-30T20:45:30

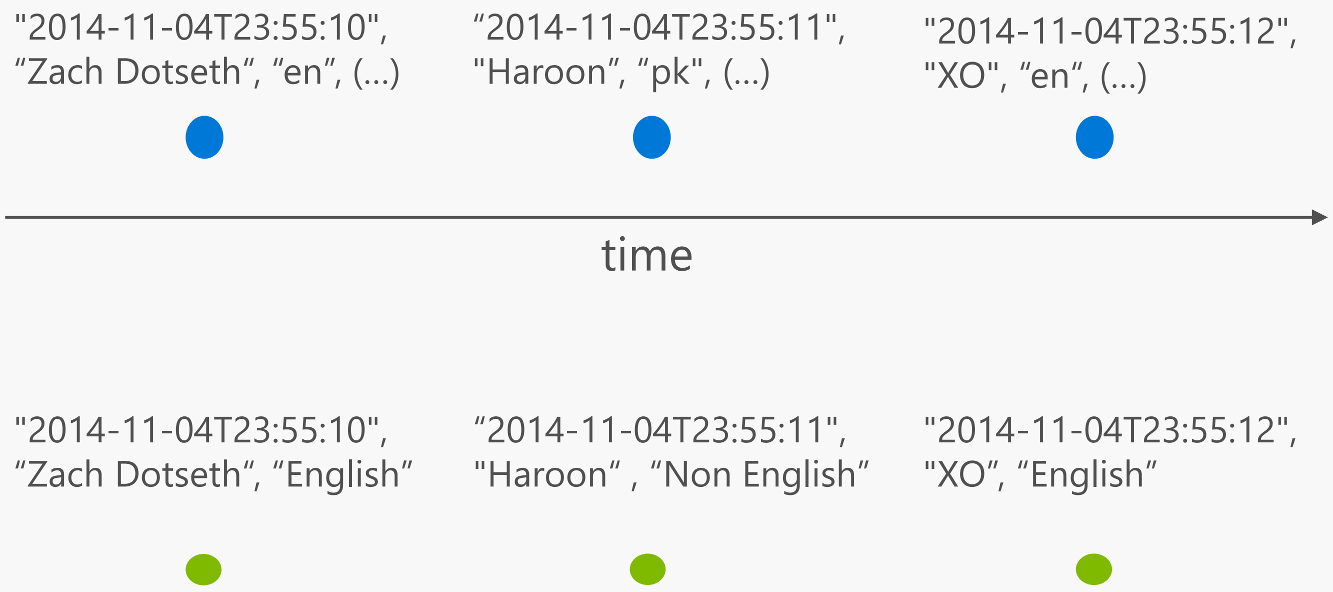
CreatedAt="2015-04-30T20:45:30"
Topic="XBox"
SentimentScore=2
Text="Oh, joy! More forced @Xbox .."
….

Streams

Projections

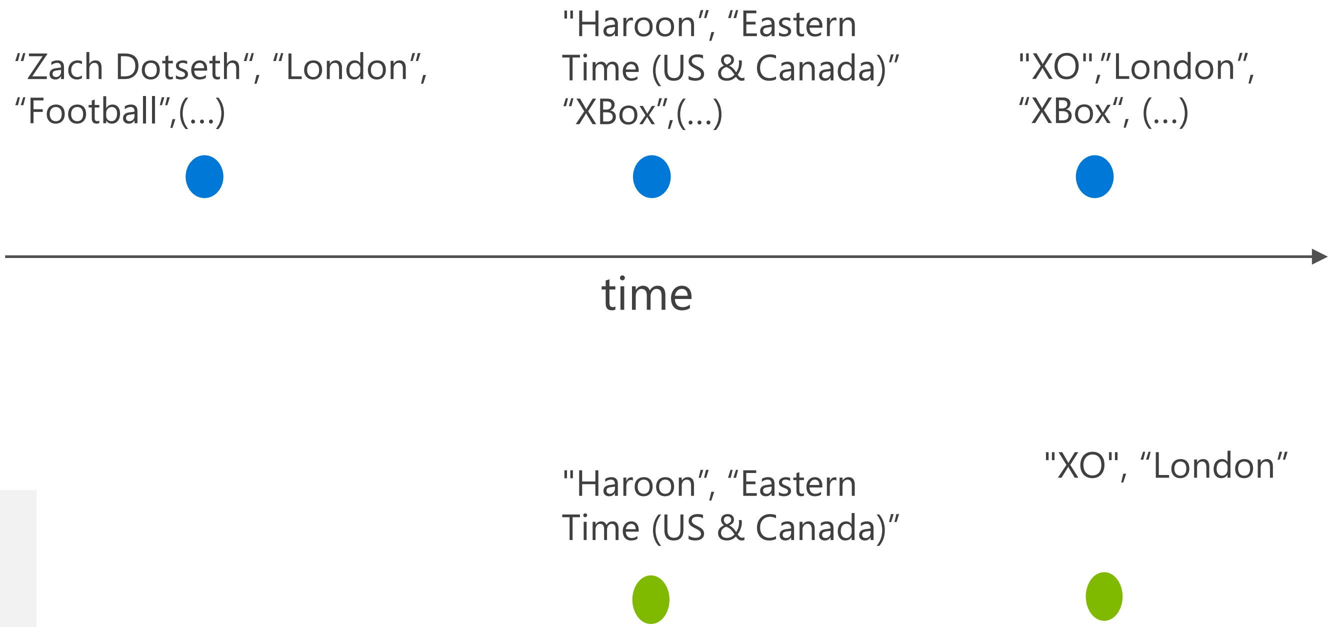
Tell me the time, user name and if the language is English or not

```
SELECT CreatedAt, UserName,
   Language=
   CASE Language
       WHEN 'en' THEN 'English'
       ELSE 'Non English'
   END
FROM TwitterStream
```



Filters

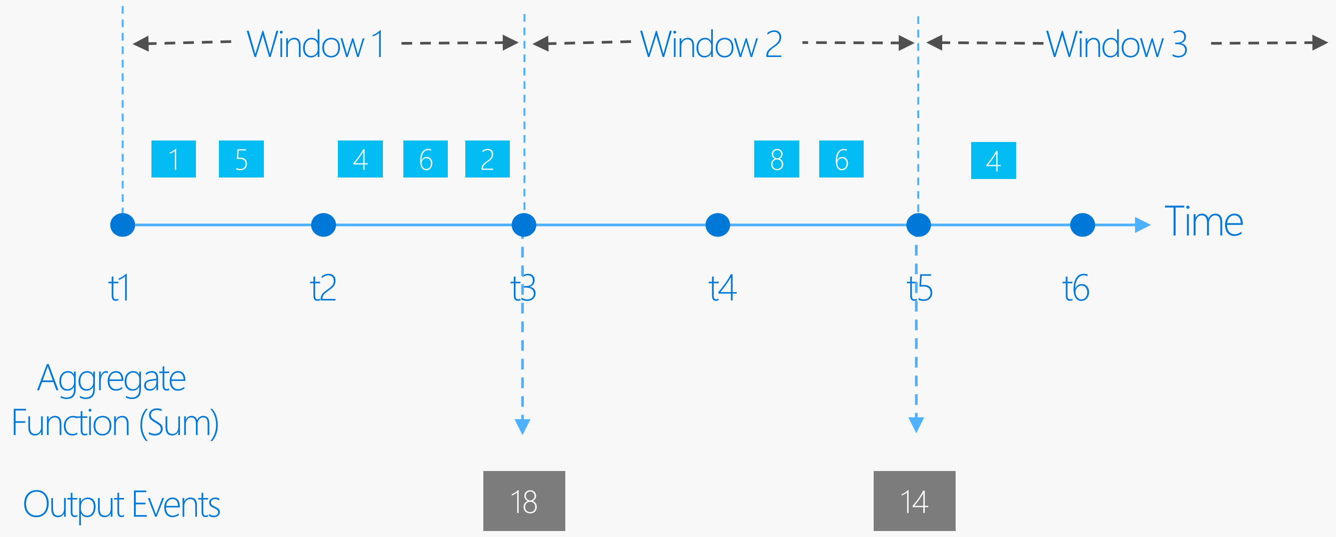
Show me the user name and time zone of tweets on the topic XBox



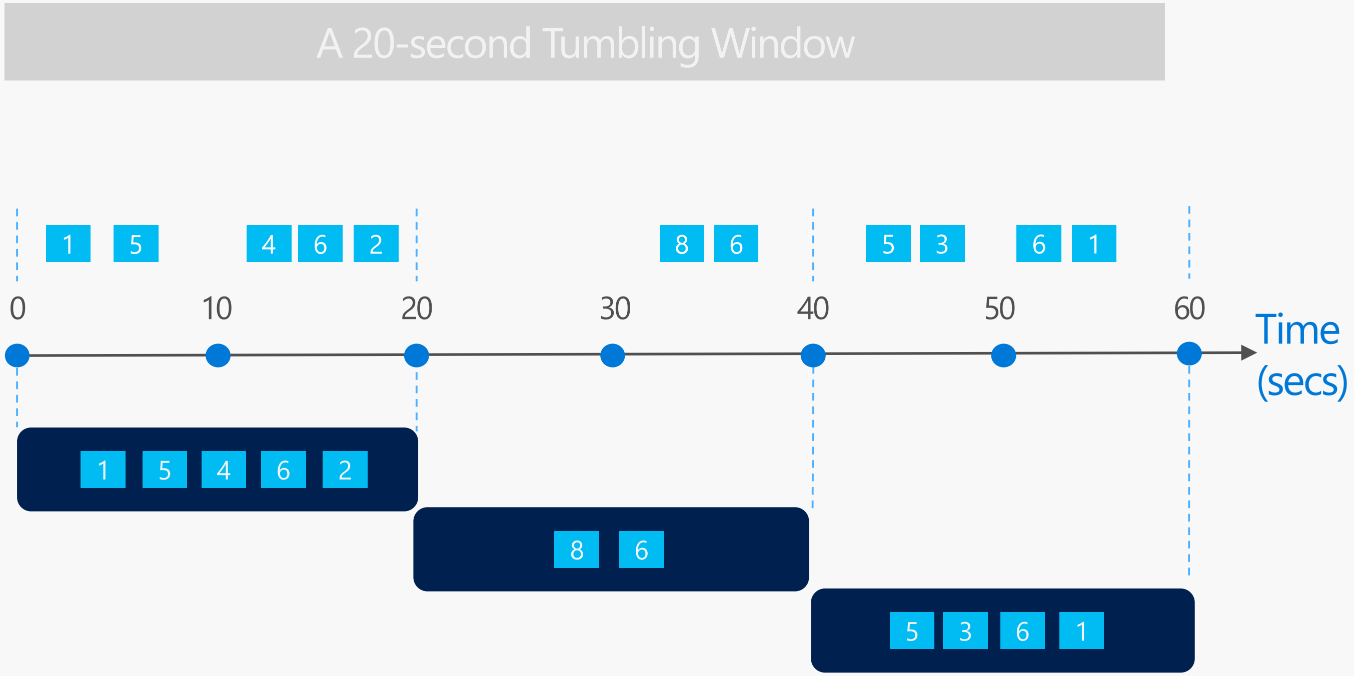

```
SELECT UserName, TimeZone
FROM InputStream
WHERE Topic = 'XBox'
```

Windowing Concepts

- Windows can be – Hopping, Sliding or Tumbling
- Windows are fixed length
- Output event will have the timestamp of the end of the window
- Must be used in a GROUP BY clause



Tumbling Window



Tumbling windows:

- Repeat
- Are non-overlapping

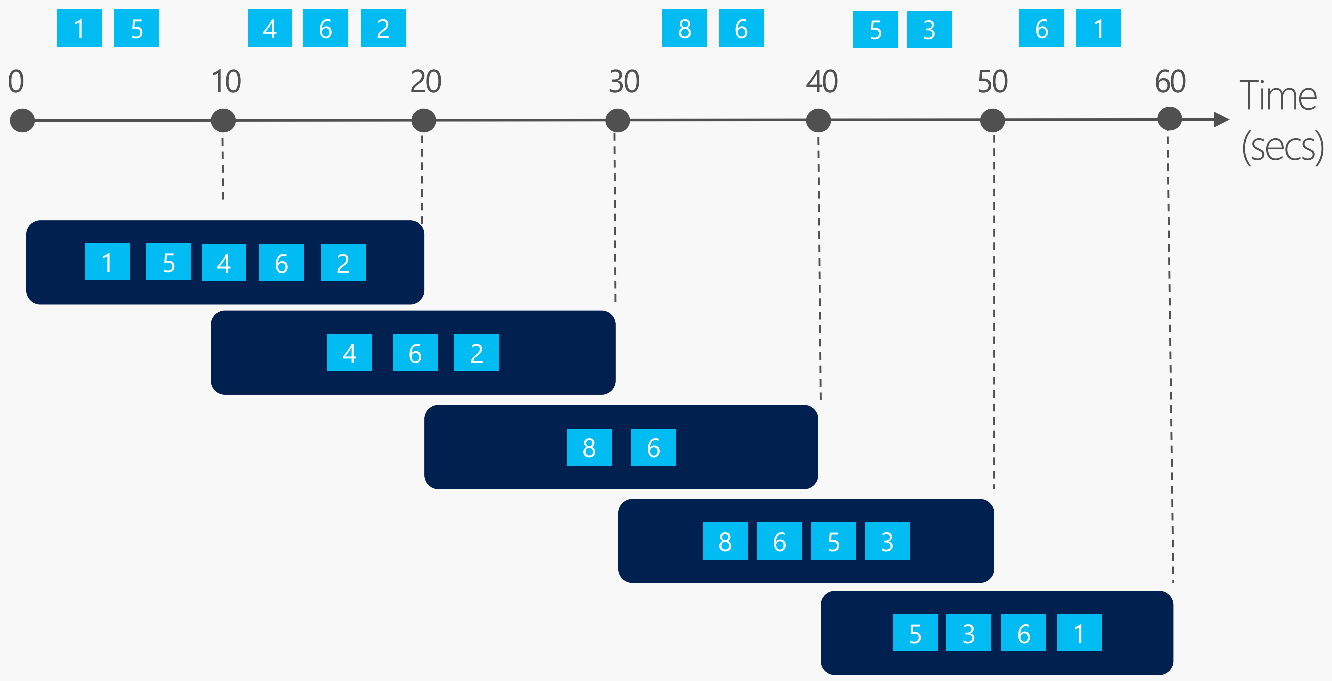
An event can belong to only one tumbling window

Query: Count the total number of vehicles entering each toll booth every interval of 20 seconds.

```
SELECT TollId, COUNT(*)
FROM EntryStream TIMESTAMP BY EntryTime
GROUP BY  TollId, TumblingWindow(second, 20)
```

Hopping Window

A 20-second Hopping Window with a 10 second "Hop"



Hopping windows:

- Repeat
- Can overlap
- Hop forward in time by a fixed period

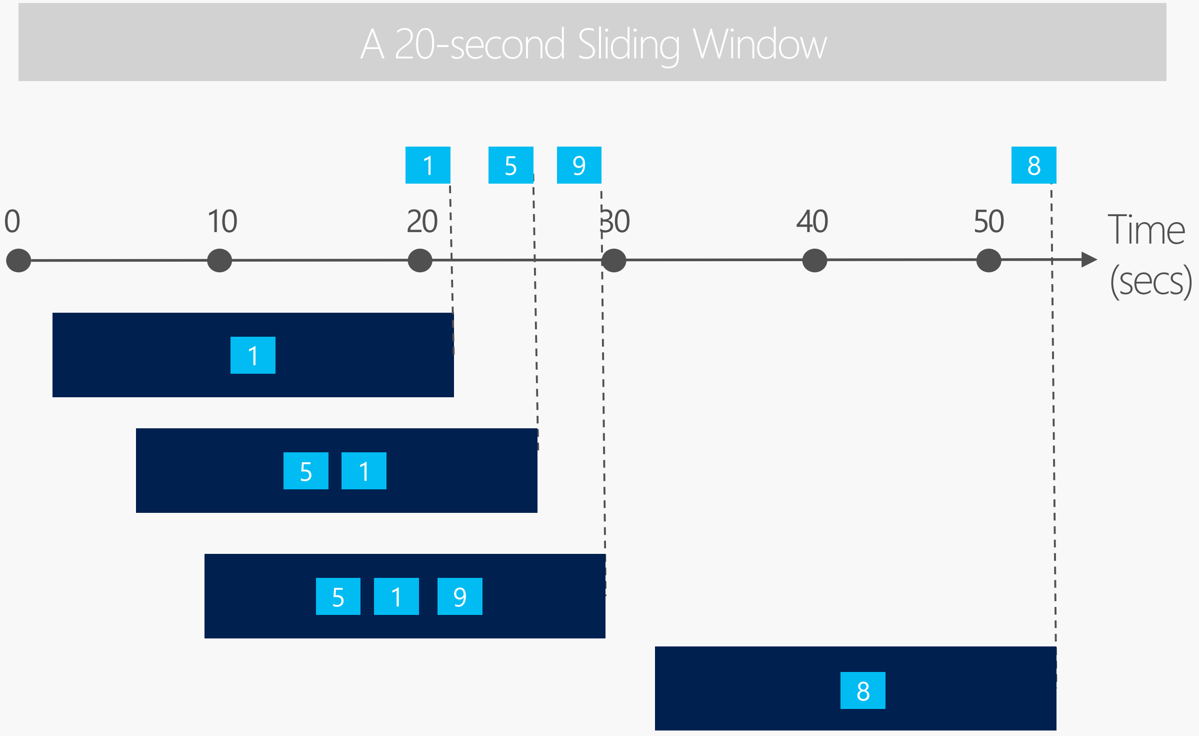
Same as tumbling window if hop size = window size

Events can belong to more than one hopping window

QUERY: Count the number of vehicles entering each toll booth every interval of 20 seconds; update results every 10 seconds

```
SELECT COUNT(*), TollId
FROM EntryStream TIMESTAMP BY EntryTime
GROUP BY  TollId, HoppingWindow (second, 20,10)
```

Sliding Window



Sliding window:

- Continuously moves forward by an € (epsilon)
- Produces an output *only during the occurrence of an event*
- Every windows will have at least one event

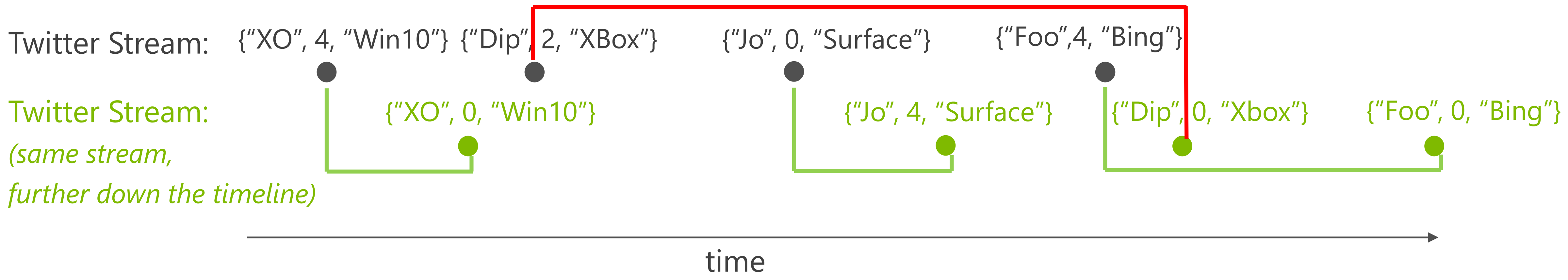
Events can belong to more than one sliding window

Query: Find all the toll booths which have served more than 10 vehicles in the last 20 seconds

```
SELECT TollId, Count(*)
FROM EntryStream ES
GROUP BY  TollId, SlidingWindow (second, 20)
HAVING Count(*) > 10
```

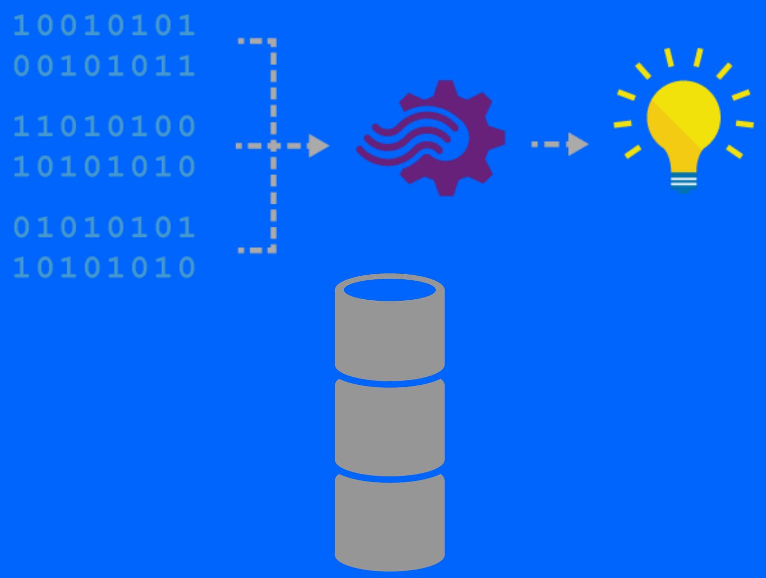
Joining multiple streams

"List all users and the topics on which they switched their sentiment within a minute"




```
SELECT TS1.UserName, TS1.Topic
FROM TwitterStream TS1 TIMESTAMP BY CreatedAt
JOIN TwitterStream TS2 TIMESTAMP BY CreatedAt
      ON TS1.UserName = TS2.UserName AND TS1.Topic = TS2.Topic
      AND DateDiff(second, TS1, TS2) BETWEEN 1 AND 60
WHERE TS1.SentimentScore != TS2.SentimentScore
```

Reference Data



Seamless correlation of event streams with reference data

Static or slowly-changing data stored in blobs

CSV and JSON files in Azure Blobs;
scanned for new snapshots on a settable cadence

JOIN (INNER or LEFT OUTER) between streams and reference data sources

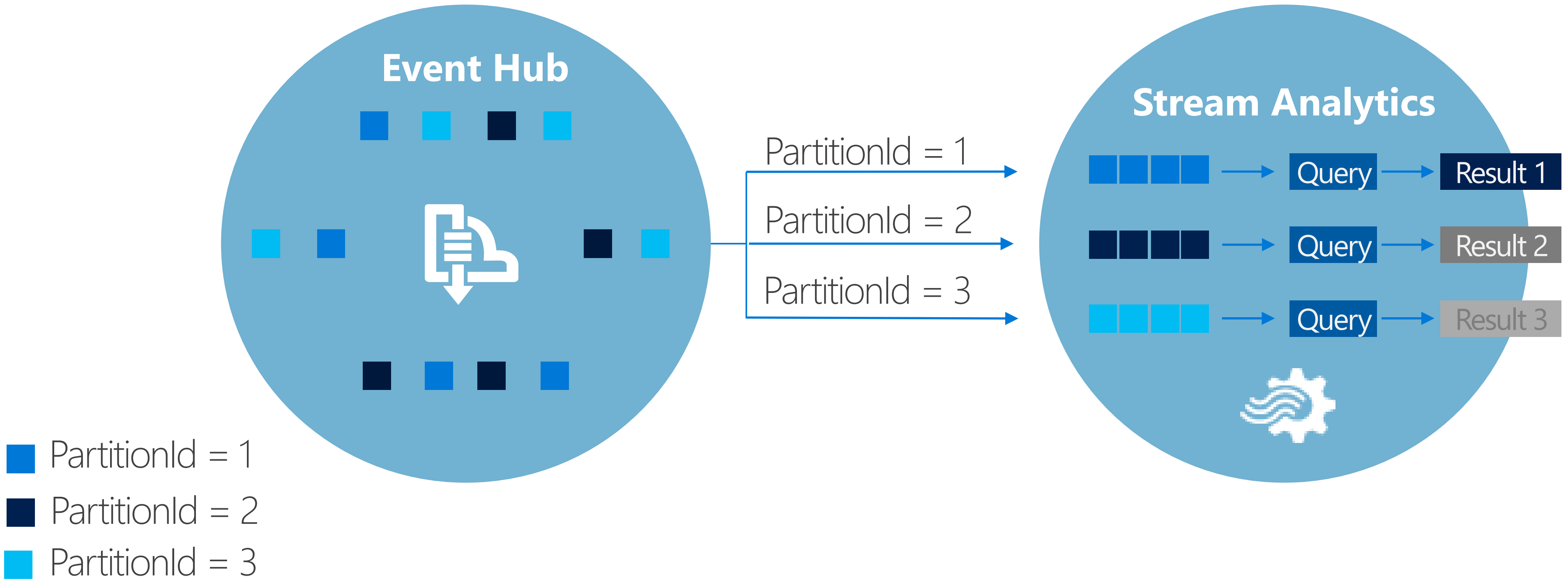
Reference data appears like another input:

```
SELECT myRefData.Name, myStream.Value
FROM myStream
JOIN myRefData
       ON myStream.myKey = myRefData.myKey
```

Scaling using Partitions

Partitioning allows for parallel execution over scaled-out resources

```
SELECT Count(*) AS Count, Topic
FROM TwitterStream PARTITION BY PartitionId
GROUP BY TumblingWindow(minute, 3), Topic, PartitionId
```

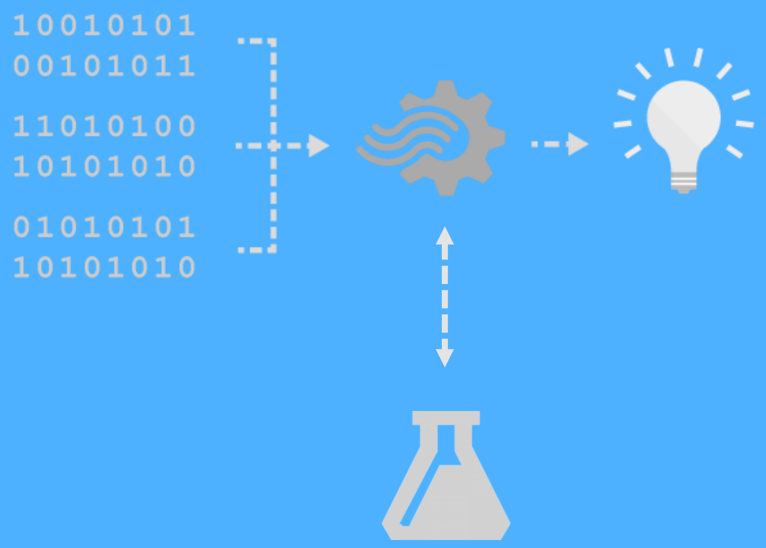


Multiple steps, multiple outputs

```
WITH Step1 AS (
    SELECT Count(*) AS CountTweets, Topic
    FROM TwitterStream PARTITION BY PartitionId
    GROUP BY TumblingWindow(second, 3), Topic, PartitionId
)
,
Step2 AS (
    SELECT Avg(CountTweets)
    FROM Step1
    GROUP BY TumblingWindow(minute, 3)
)
SELECT * INTO Output1 FROM Step1
SELECT * INTO Output2 FROM Step2
SELECT * INTO Output3 FROM Step2
```

- A query can have **multiple steps** to enable pipeline execution
- A step is a sub-query defined using **WITH** ("common table expression")
- Can be used to develop complex queries more elegantly by creating a **intermediary named result**
- Creates unit of execution for scaling out when **PARTITION BY** is used
- Each step's output can be sent to multiple output targets using **INTO**

Machine Learning



Azure ML and Stream Analytics are now integrated

Azure ML can publish web endpoints for operationalized models

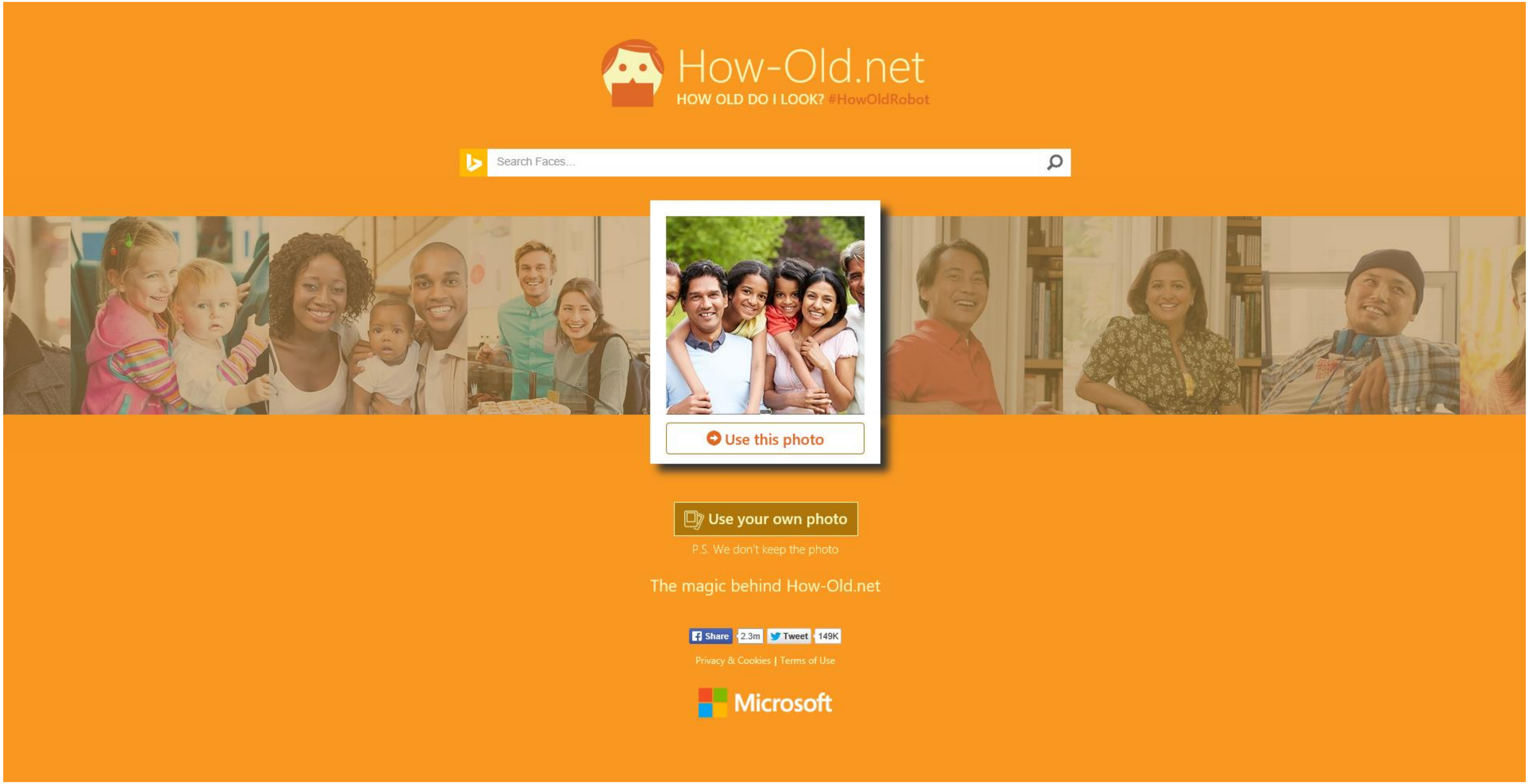
Azure Stream Analytics can bind custom function names to such web endpoints

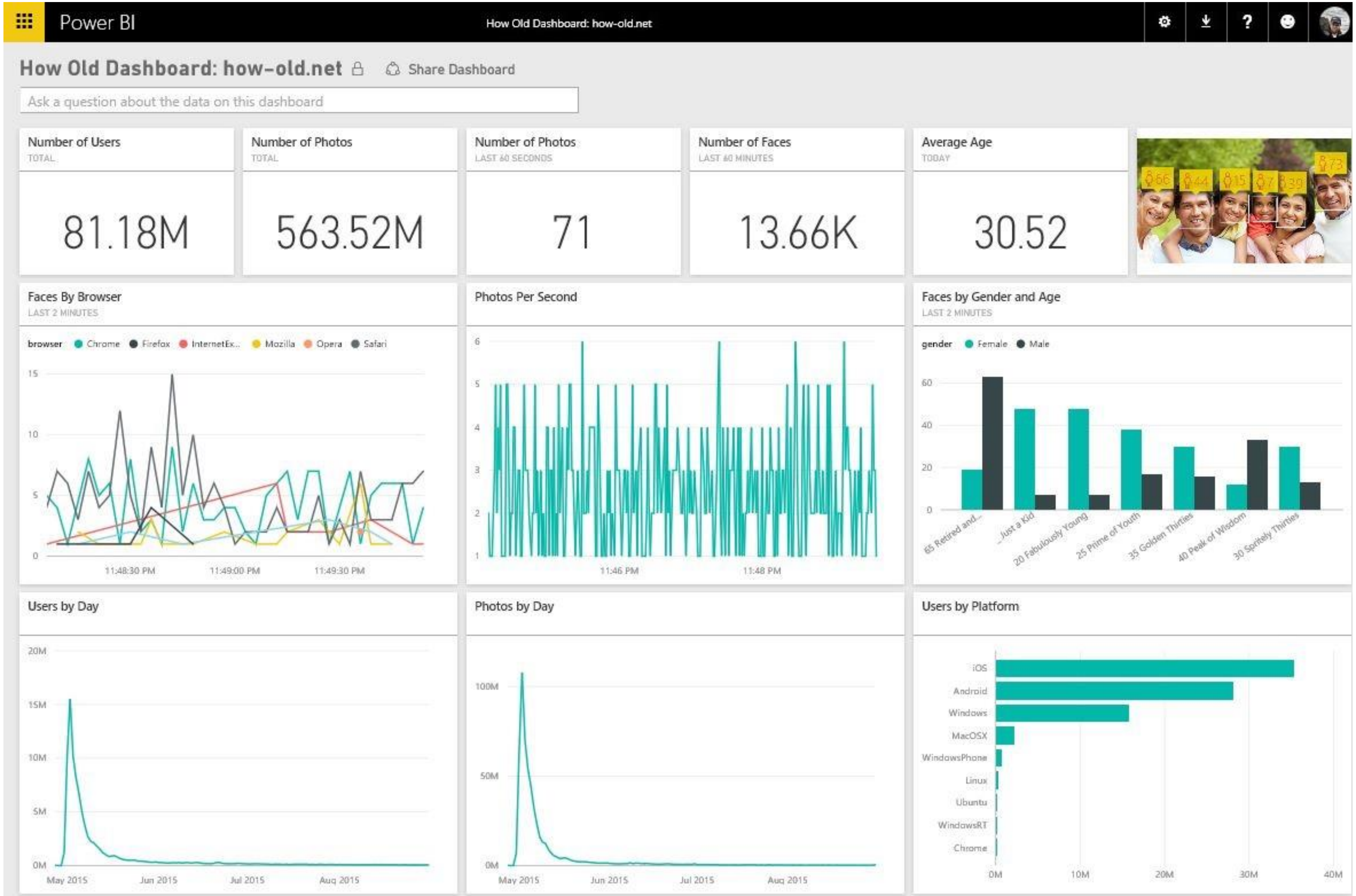
Example: apply bound function event-by-event

```
SELECT text, sentiment(text) AS score
FROM myStream
```

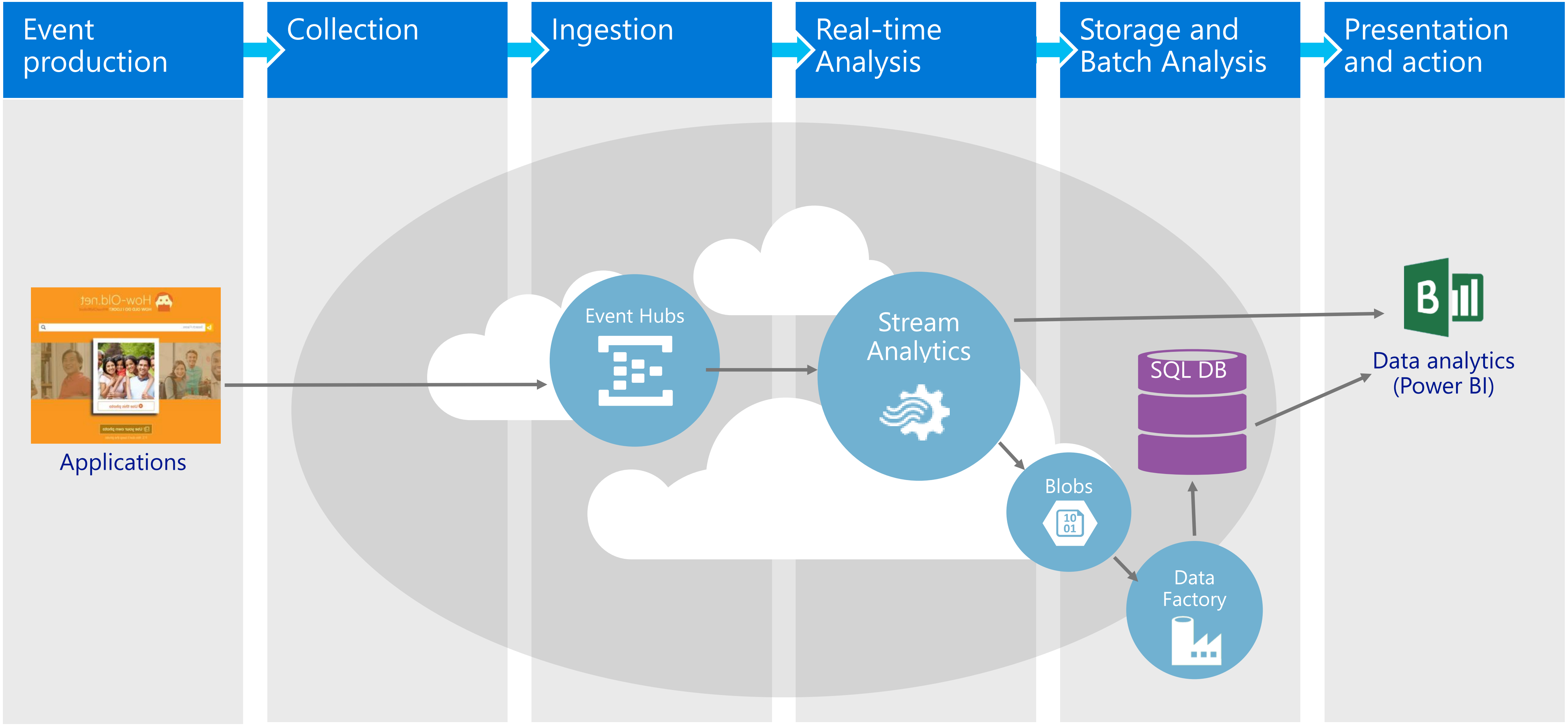
how-old.net

Real-Time Dash-boarding with PowerBI, and Scale

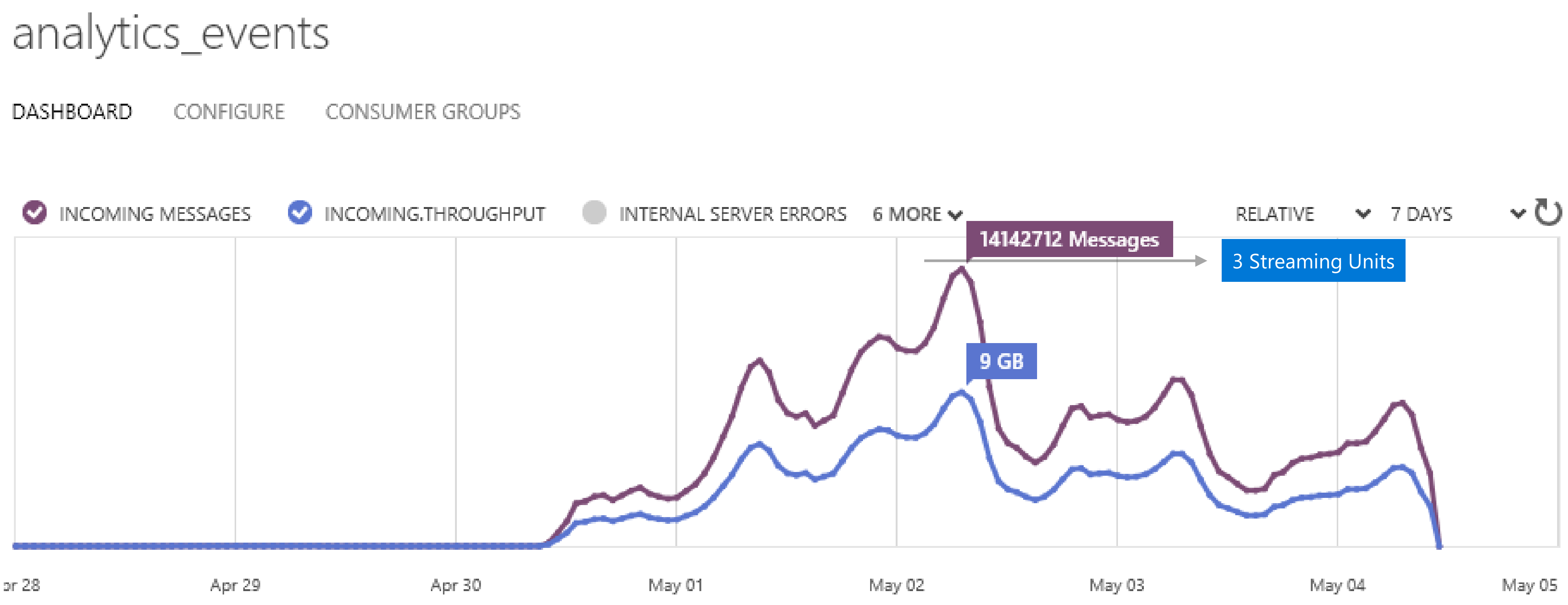

How-Old.net
HOW OLD DO I LOOK? #HowOldRobot
Search Faces...
Use this photo
Use your own photo
P.S. We don't keep the photo
The magic behind How-Old.net
Share 2.3m
Tweet 149K
Privacy & Cookies | Terms of Use
Microsoft


Power BI
How Old Dashboard: how-old.net
How Old Dashboard: how-old.net
Share Dashboard
Ask a question about the data on this dashboard
Number of Users
TOTAL
81.18M
Number of Photos
TOTAL
563.52M
Number of Photos
LAST 60 SECONDS
71
Number of Faces
LAST 60 MINUTES
13.66K
Average Age
TODAY
30.52
Faces By Browser
LAST 2 MINUTES
browser Chrome Firefox InternetEx... Mozilla Opera Safari
11:48:30 PM 11:49:00 PM 11:49:30 PM
Photos Per Second
11:46 PM 11:48 PM
Faces by Gender and Age
LAST 2 MINUTES
gender Female Male
65 Retired and... Just a Kid 20 Fabulously Young 25 Prime of Youth 35 Golden Thirties 40 Peak of Wisdom 30 Spritely Thirties
Users by Day
May 2015 Jun 2015 Jul 2015 Aug 2015
Photos by Day
May 2015 Jun 2015 Jul 2015 Aug 2015
Users by Platform
iOS
Android
Windows
MacOSX
WindowsPhone
Linux
Ubuntu
WindowsRT
Chrome
0M 10M 20M 30M 40M

How-old.net Data Flow



Advantage of Cloud – Scale and Costs



ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository

NATIONAL AERONAUTICS
AND SPACE ADMINISTRATION

+ABOUT NASA | +LATEST NEWS | +MULTIMEDIA | +MISSIONS | +WORK FOR NASA

- + NASA Home
- + Ames Home
- + Intelligent Systems Division
- + Discovery and Systems Health

Data Repository

- + Home
- + Lab
- + Projects
- + Data Repository
- + Publications
- + Roadmap
- + Awards

Search

PCoE Datasets

Overview

The Prognostics Data Repository is a collection of data sets
universities, agencies, or companies. The data repository fo
i.e., data sets that can be used for development of prognost
of data from some nominal state to a failed state. The collec
process.

Publications making use of databases obtained from this rep
the assistance received by using this repository and the don
obtain the same data sets and replicate your experiments. It

Users employ the data at their own risk. Neither NASA nor t
liability for the use of the data or any system developed usin

If you have suggestions concerning the repository send ema
and please come again.

Turbofan Engine Degradation Simulation Data Set
Publications using this data set

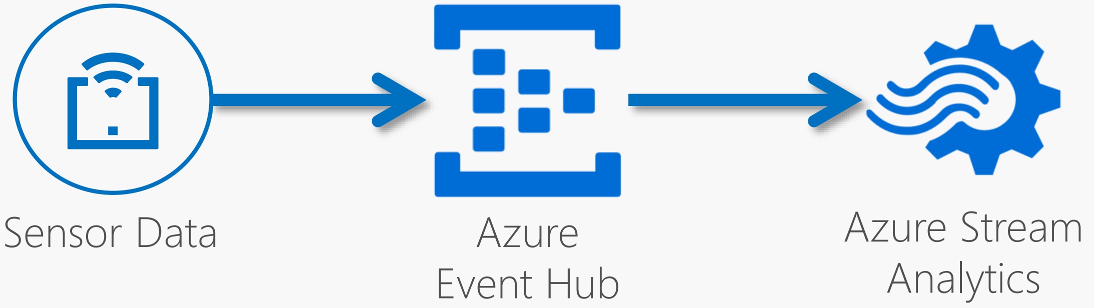
Description	Engine degradation simulation was carried out using C-MAPSS. Four different were sets simulated under different combinations of operational conditions and fault modes. Records several sensor channels to characterize fault evolution. The data set was provided by the Prognostics CoE at NASA Ames.
Format	The set is in text format and has been zipped including a readme file.
Datasets	+ Download Turbofan Engine Degradation Simulation Data Set (9064 downloads)
Dataset Citation	A. Saxena and K. Goebel (2008). "Turbofan Engine Degradation Simulation Data Set", NASA Ames Prognostics Data Repository (http://ti.arc.nasa.gov/project/prognostic-data-repository), NASA Ames Research Center, Moffett Field, CA

http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

HOL 2: Stream Processing

On GitHub

3. Azure Event Hub
4. Azure Stream Analytics Jobs
5. Data Generator



Pricing

Stream Analytics is priced on two variables:

- **Volume of data** processed
- **Streaming units** required to process the data stream

Meter	Price (USD)
Volume of Data Processed • **Volume of data processed by the streaming job (in GB)**	$.001 per GB
Streaming Unit • **Blended measure of cores, memory, and bandwidth**	$0.031 per hour

** Streaming unit is a unit of compute capacity with a maximum throughput of 1MB/s*

Example Pricing

Daily Azure Stream Analytics cost for 1 MB/sec of average processing

Volume of Data Processed Cost -

$0.001 /GB * 84.375 GB = **$0.08** ***per day***, *streaming max 1 MB/s non-stop*

Streaming Unit Cost -

$.031 /hr * 24 hrs = **$0.74** ***per day***, *for 1 MB/sec max. throughput*

Total cost -

$0.74 + $0.08 = **$0.82** ***per day*** -or- **$24.60 per month**

Example Pricing

Azure Stream Analytics cost for **How-Old.net** at its peak

Volume of Data Processed Cost -

$0.001 /GB * 84.375 GB = **$0.08** ***per day***, *streaming max 1 MB/s non-stop*

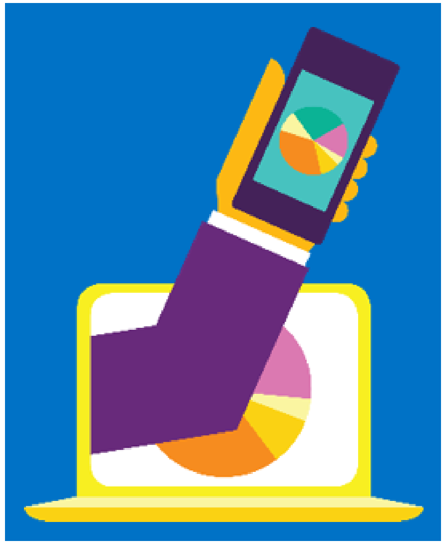
Streaming Unit Cost -

$.031 /hr * 3 units * 24 hrs = **$2.23** ***per day***, *for 3 MB/sec max. throughput*

Total cost -

$2.23 + $0.08 = **$2.39** ***per day*** -or- **$74.15 per month**

Resource Library



Business Overview	http://azure.microsoft.com/en-us/services/stream-analytics/
Documentation	http://azure.microsoft.com/en-us/documentation/services/stream-analytics/
Samples	https://github.com/streamanalytics/samples
ASA Blog	http://blogs.msdn.com/b/streamanalytics/rss.aspx
Follow us on Twitter	https://twitter.com/AzureStreaming (follow @AzureStreaming)
ASA Forum	https://social.msdn.microsoft.com/Forums/en-US/home?forum=AzureStreamAnalytics
Vote for ideas	http://feedback.azure.com/forums/270577-azure-stream-analytics
Email ASA Team	azstream@microsoft.com

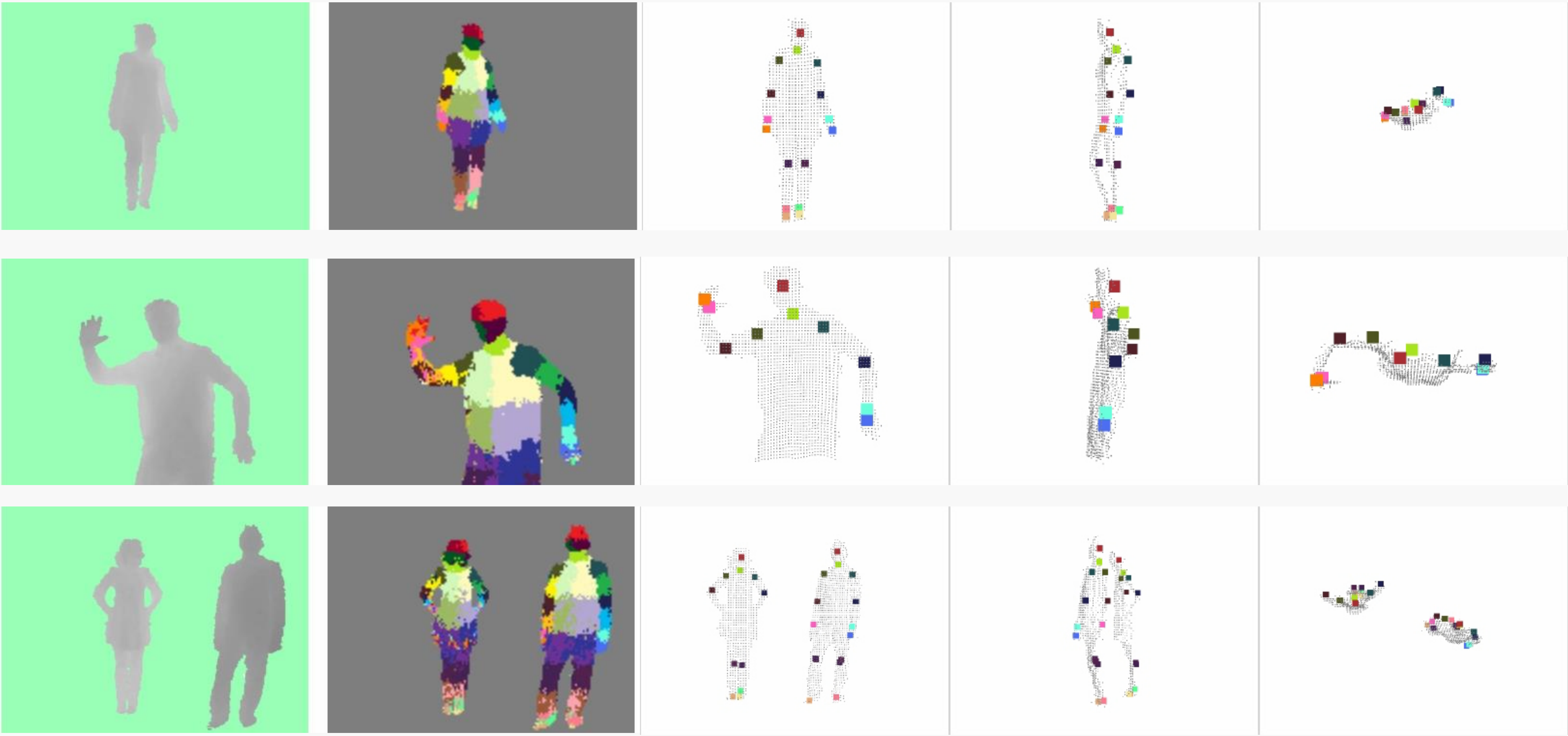
HOL 3: Azure SQL Database

Step 6 (GitHub): Azure SQL Server and Database

Machine Learning

Decision forest detector

(no tracking or smoothing)



input depth | inferred body parts | front view | side view | top view

Kinect Training Data

Discriminative & Generative

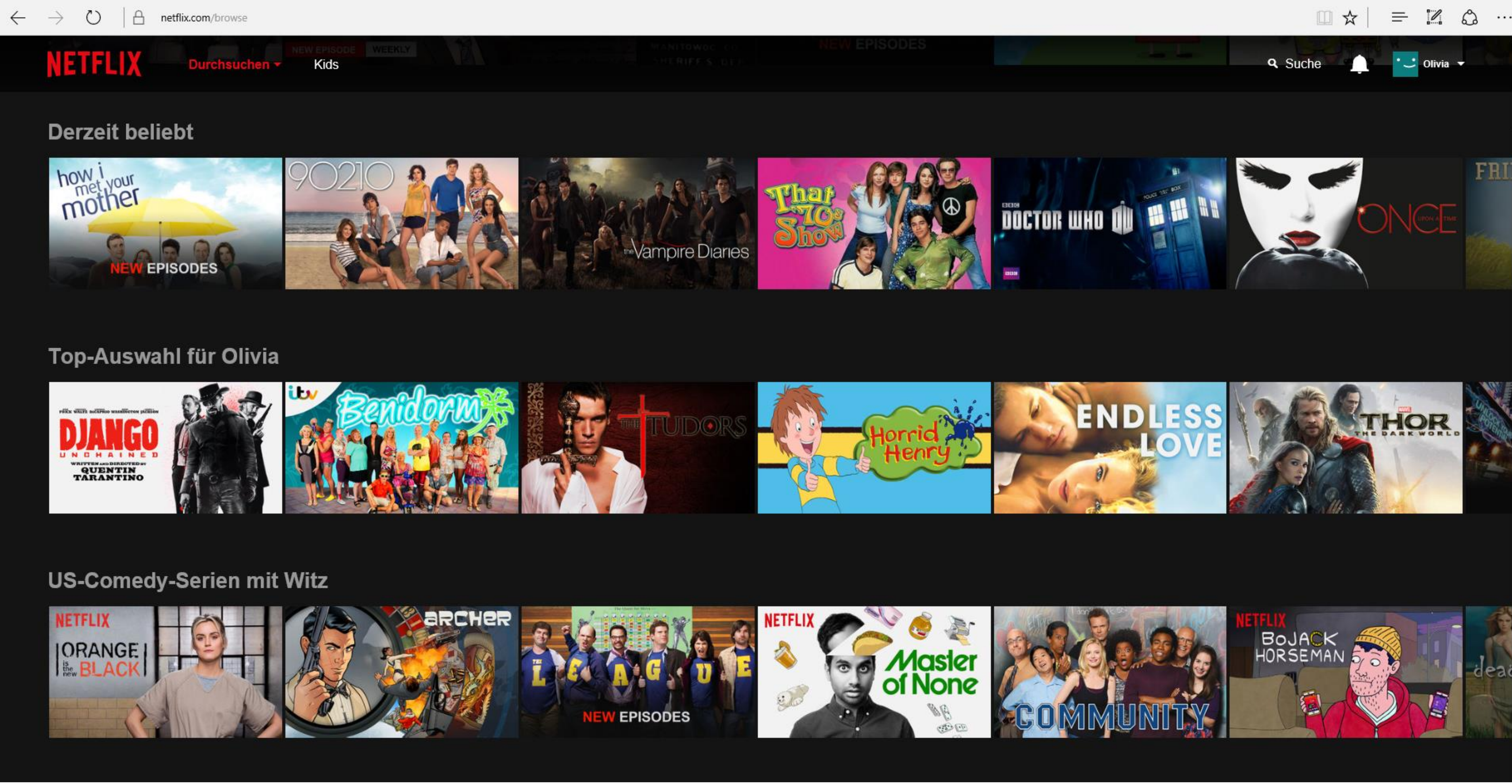
Needs lots of labelled data

Discriminative

'bicycle'

Generative

Huge search problem

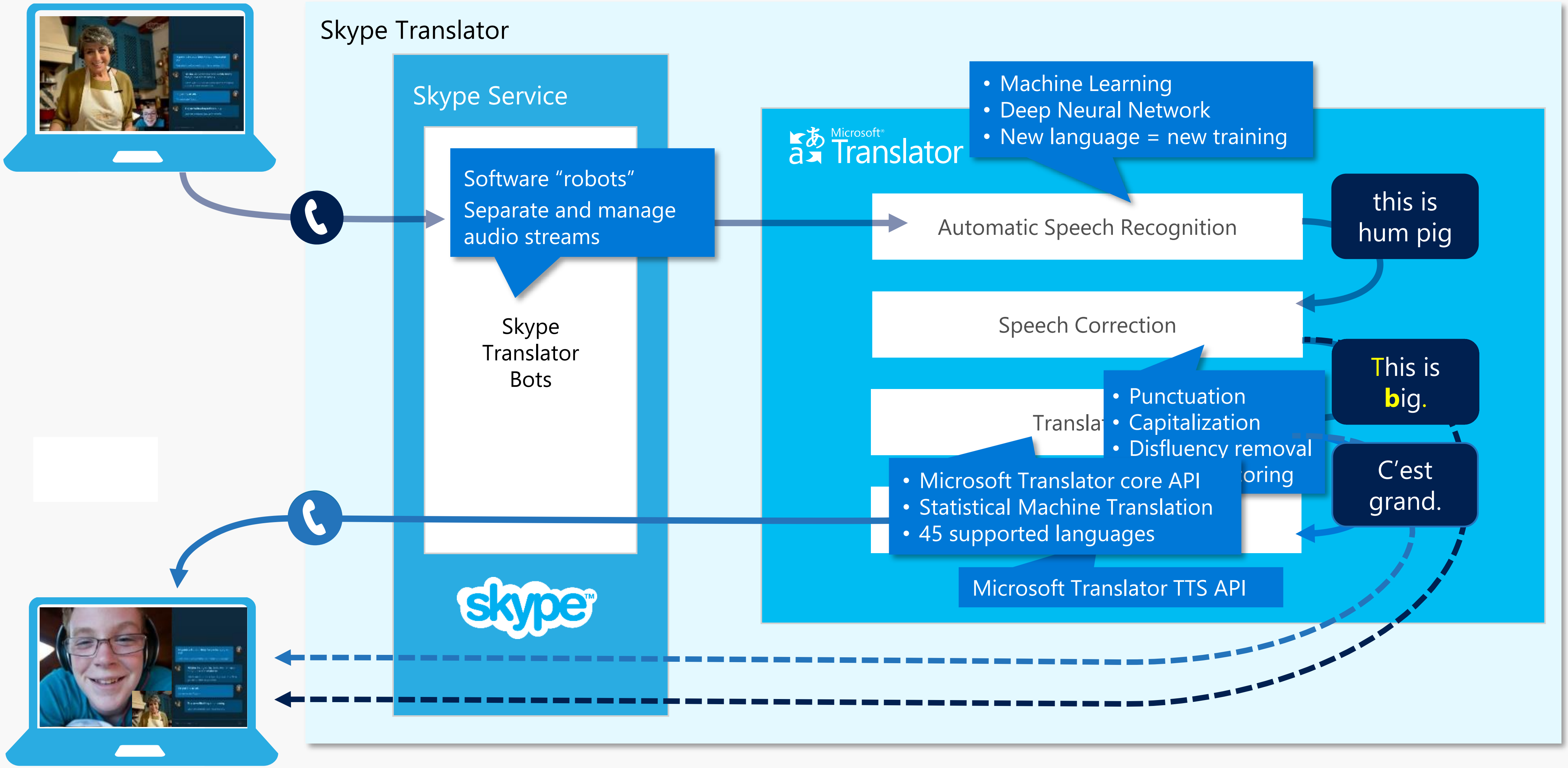


http://techblog.netflix.com/2012/04/netflix-recommendations-beyond-5-stars.html
http://www.newyorker.com/magazine/2014/02/03/outside-the-box-2

Translator

Rich Rashid in Tianjin, October, 25, 2012, https://youtu.be/Nu-nlQqFCKg?t=7m11s

Skype Translator Pipeline



What is Machine Learning?

What is Machine Learning?

"The goal of machine learning is
to program computers
to use **example data** or **past experience**
to solve a given problem."

Introduction to Machine Learning, 2nd Edition, MIT Press

What is Machine Learning?

A computer program is said to
learn from ***experience E***
with respect to some class of ***tasks T***
and performance ***measure P****,*
if its performance at tasks in T,
as measured by P improves with experience E."

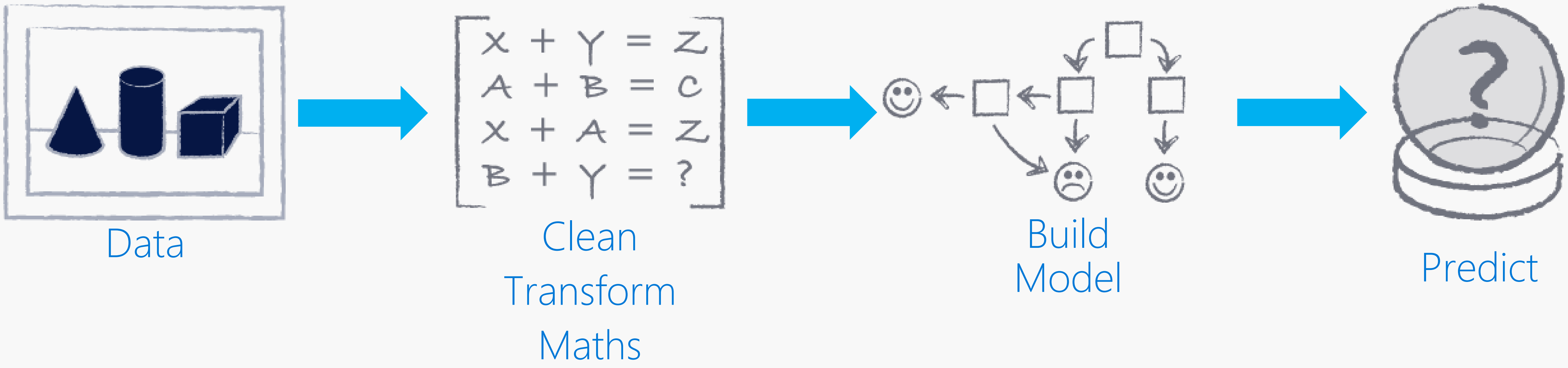
Tom M. Mitchell, Machine Learning, 1997

"We want to do machine learning!"

Sharp / Well-Defined Question	**Relevant** Data	**Accurate** Data	**Connected** Data	**A lot** of data.
Predict whether component X will fail in the next Y days	Identifiers at the level they are predicting	Failures are really failures, human labels on root causes	Machine information linkable to usage information	Will be difficult to predict failure accurately with few examples

http://blogs.technet.com/b/machinelearning/archive/2015/08/26/what-can-data-science-do-for-me.aspx

Machine Learning Process



Ahm – what?

$$f(\mathrm{X}) = \mathrm{y}$$

Input
Matrix / Table

Output
Vector / Column

Ahm – what?

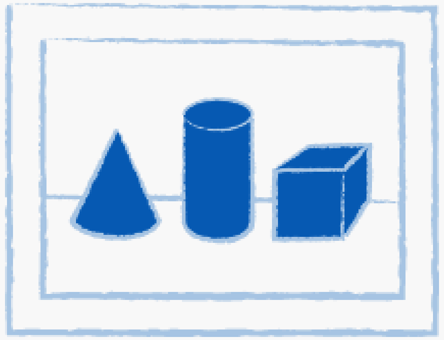
Hypothesis

$$h(\mathrm{X}) = \hat{\mathrm{y}}$$

Input
Matrix / Table

Predicted Output
Vector / Column

Data



Forecast	Temperature	Windy	Play tennis?
Sunny	Low	Ja	Play
Sunny	High	Yes	Don't Play
Sunny	High	No	Don't Play
Cloudy	Low	Yes	Play
Cloudy	High	No	Play
Cloudy	Low	No	Play
Rainy	Low	Yes	Don't Play
Rainy	Low	No	Play
Sunny	**Low**	**No**	**?**

Play / Don't Play

$$f(\mathrm{x}) = y$$

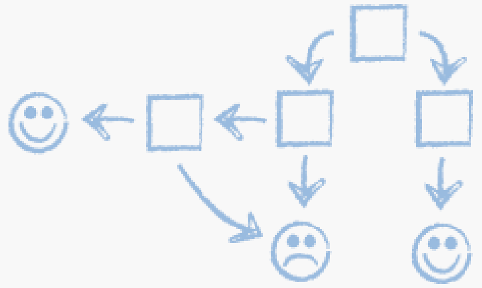
Features / Input:
(Forecast, Temperature, Windy)
e.g. $\mathbf{x} = (\text{sunny, low, yes})$

Clean, transform, maths

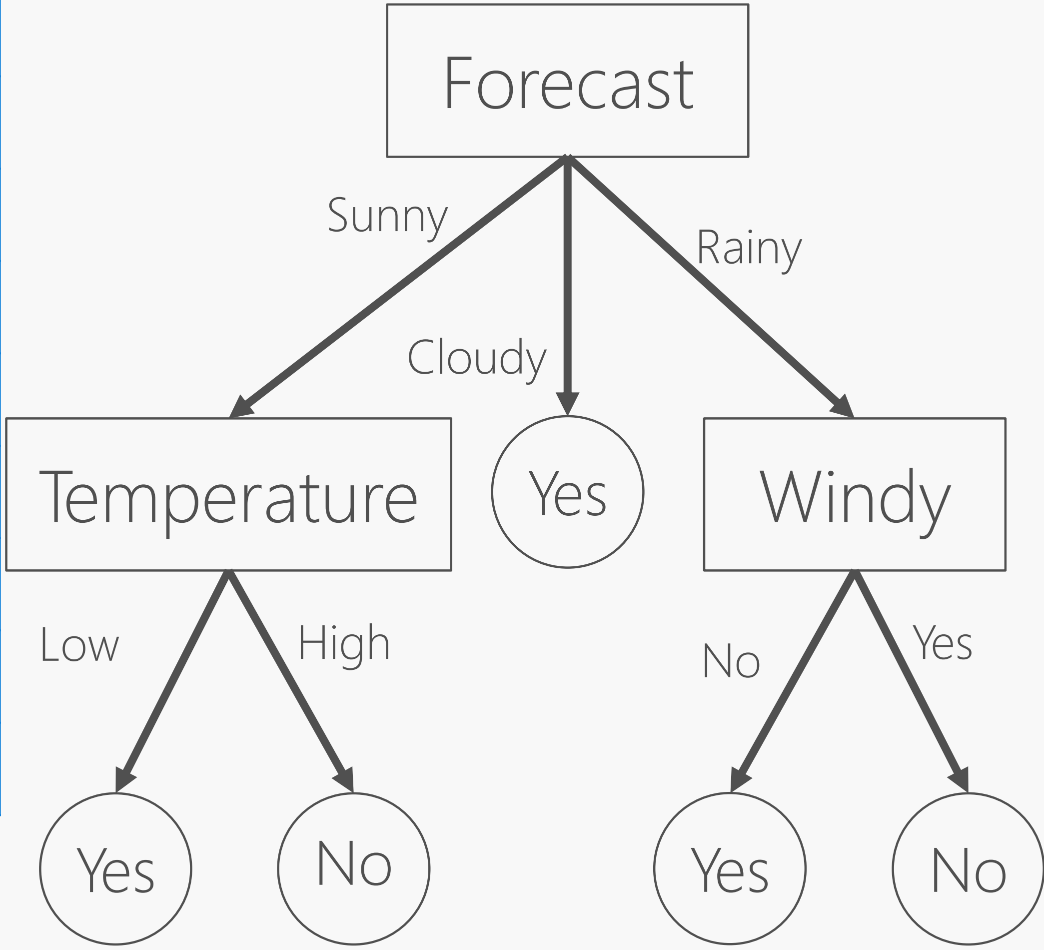
$$\begin{bmatrix} X + Y = Z \\ A + B = C \\ X + A = Z \\ B + Y = ? \end{bmatrix}$$

Forecast	Temperature	Windy	Play tennis?
Sunny	Very Low	Yes	Play
Sunny	High	Yes	Don't Play
Sunny	High	Kinda	Don't Play
Cloudy	?	Yes	One plays
Cotton-wool Clouds	High	No	Play
Cloudy	Low	No	Play
Rainy	?	Yes	Don't Play
Rainy	Low	No	Play
Sunny	**Low**	**No**	**?**

Build Model

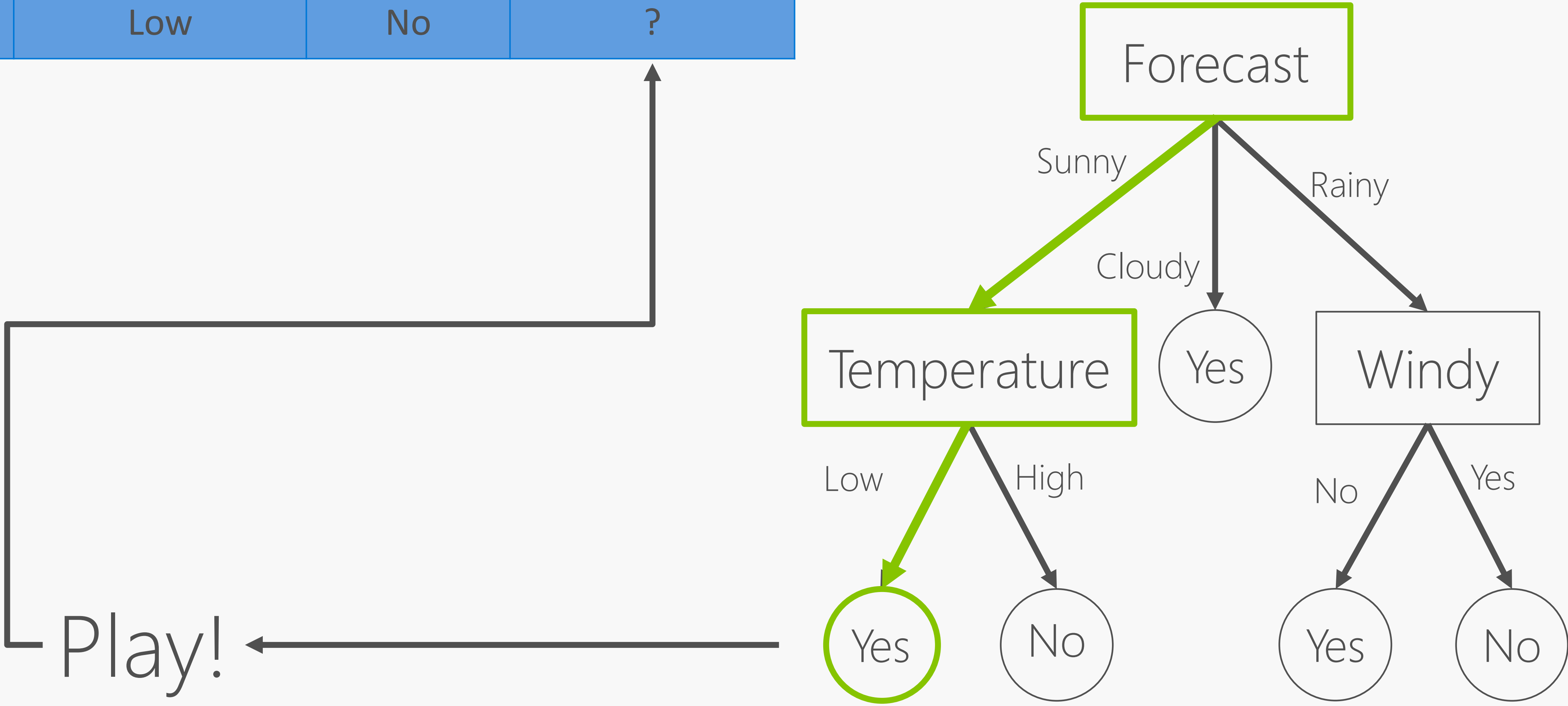


Forecast	Temperature	Windy	Play tennis?
Sunny	Low	Yes	Play
Sunny	High	Yes	Don't Play
Sunny	High	No	Don't Play
Cloudy	Low	Yes	Play
Cloudy	High	No	Play
Cloudy	Low	No	Play
Rainy	Low	Yes	Don't Play
Rainy	Low	No	Play
Sunny	**Low**	**No**	**?**



Predict i.e. Apply Model

Forecast	Temperature	Windy	Play tennis?
Sunny	Low	No	?



Is the model any good? Confusion Matrix

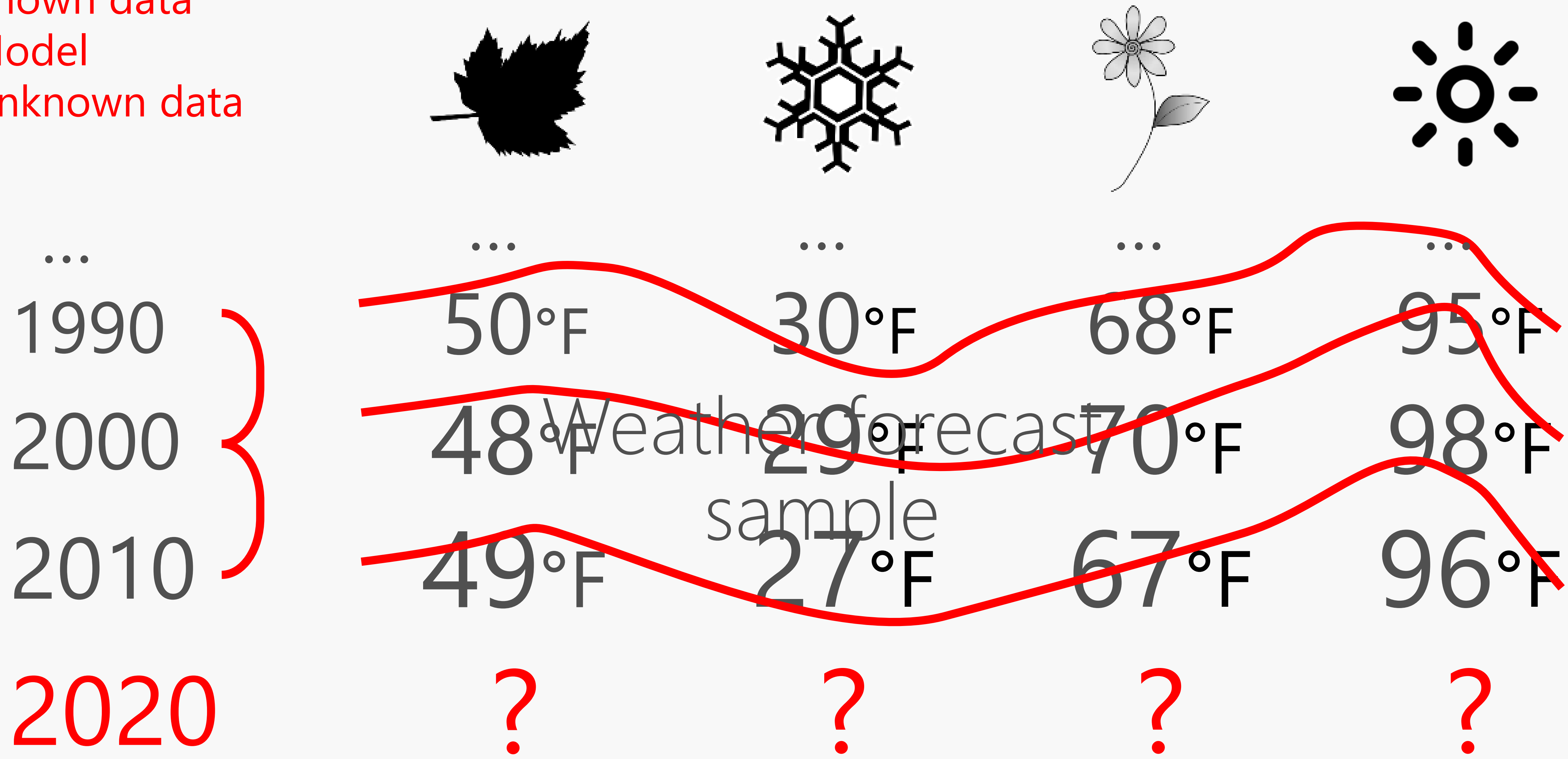
		True Label	
		Positive	Negative
Predicted Label	Positive	True positive (TP)	False positive (FP)
	Negative	False negative (FN)	True negative (TN)

$$Precision = \frac{tp}{tp + fp}$$

$$Recall = \frac{tp}{tp + fn}$$

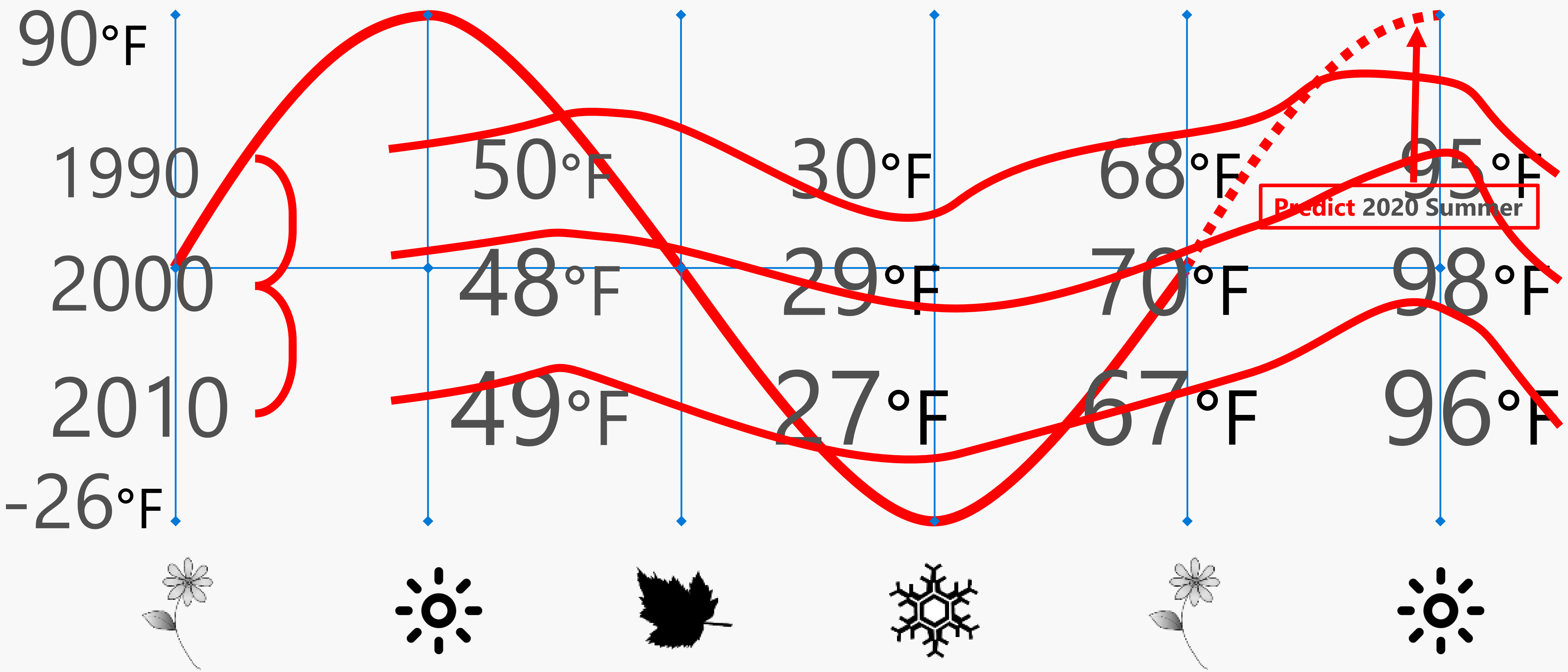
$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

Known data
Model
Unknown data



Using **known data**, develop a **model** to predict **unknown data**.

Model (Regression)

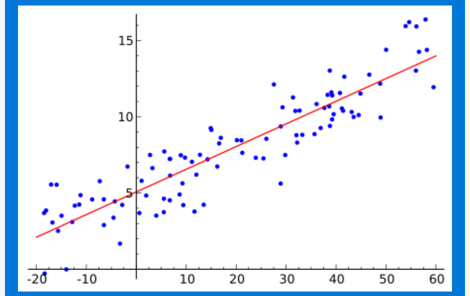


Using **known data**, develop a **model** to **predict** **unknown data**.

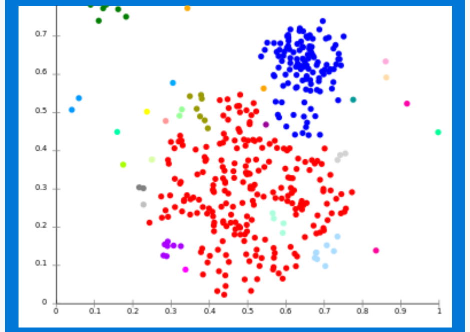
Machine Learning Problems

Classification

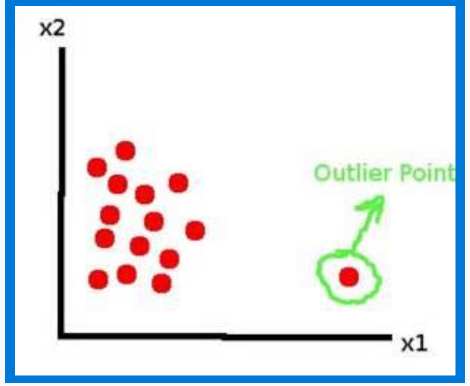
Regression



Clustering

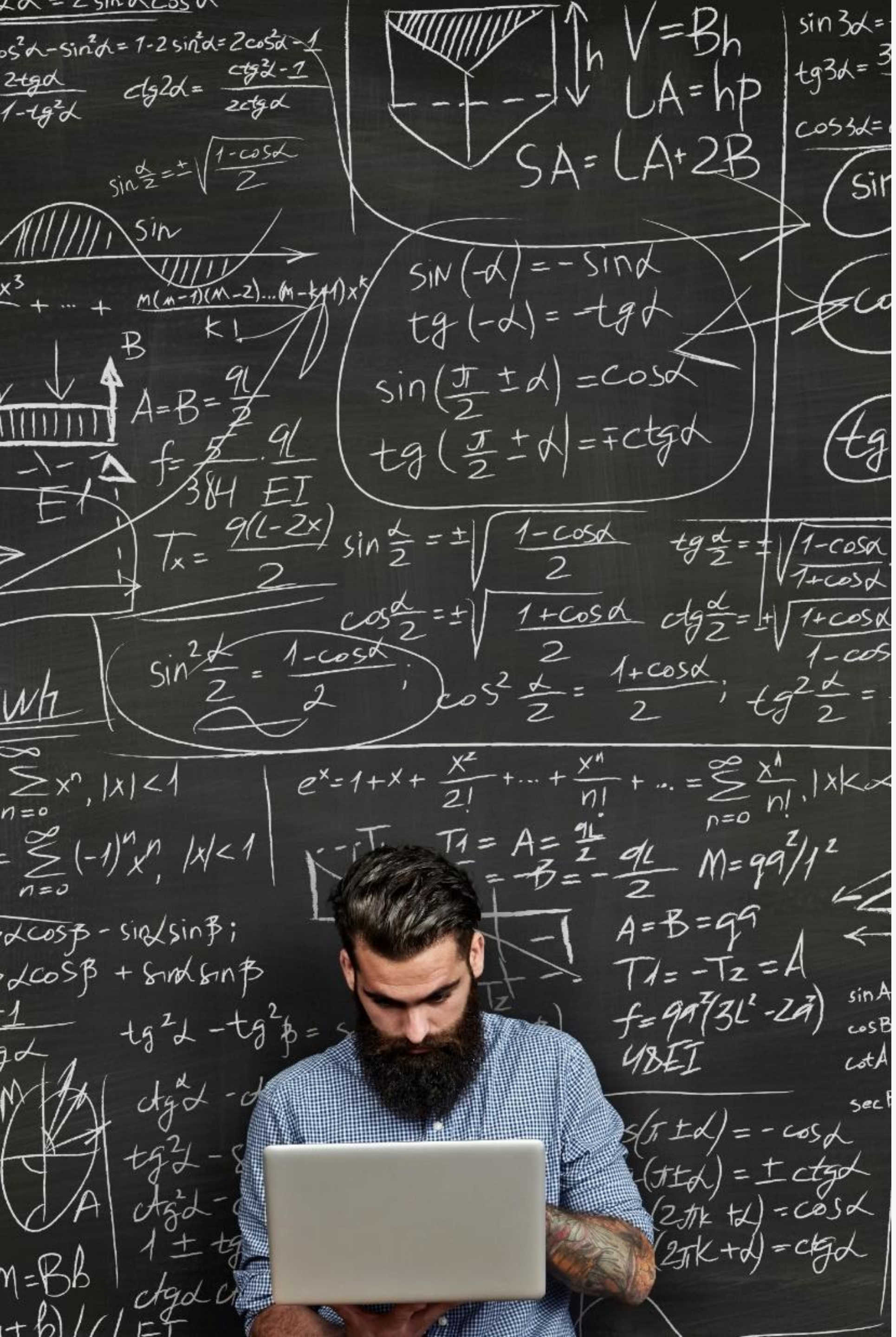


Anomaly Detection



Recommender

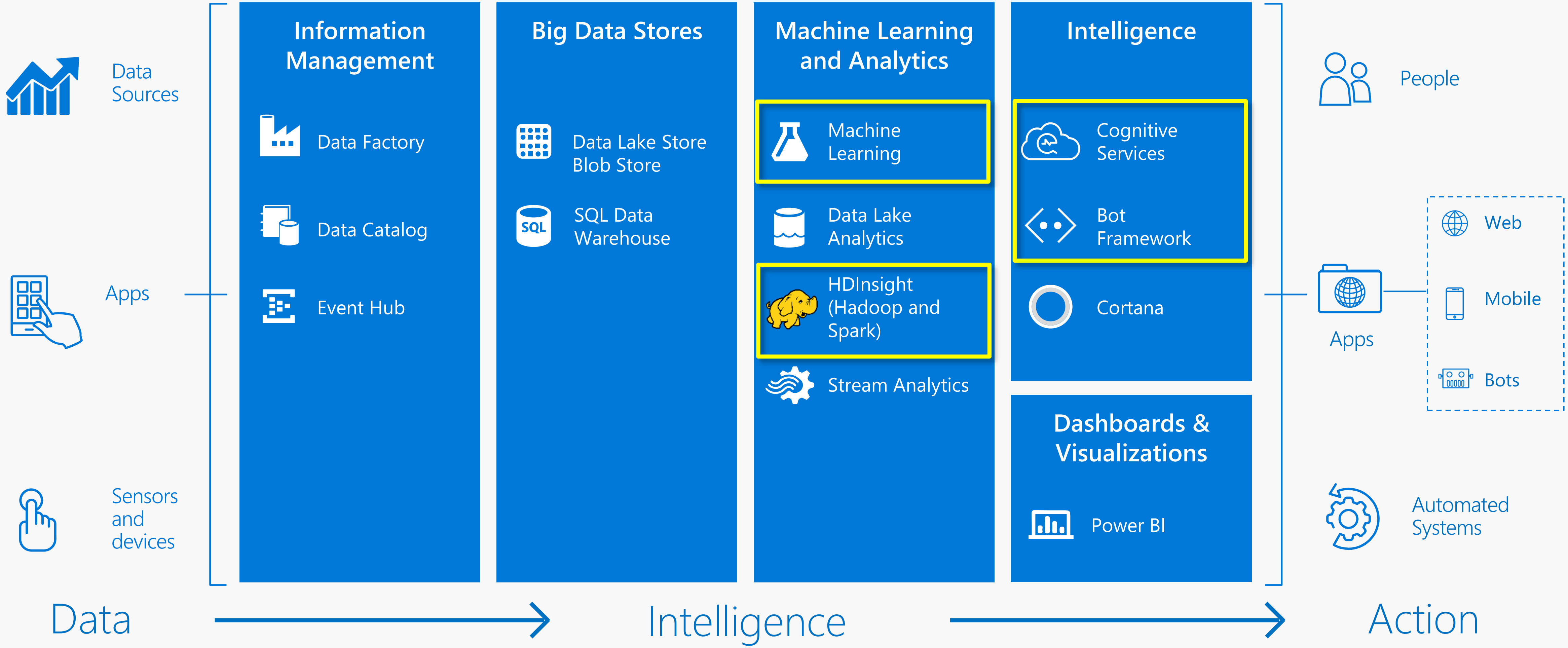
Machine Learning @ Microsoft



ML Adoption Challenges

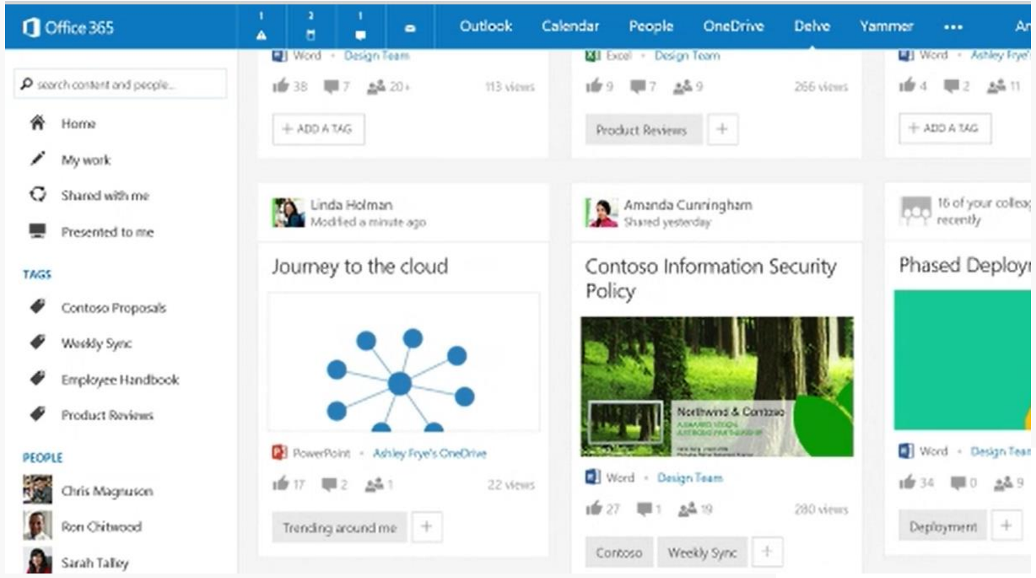
- Specialized skills
- Infrastructure
- Long time to insight
- Productization complexity

Cortana Intelligence Suite

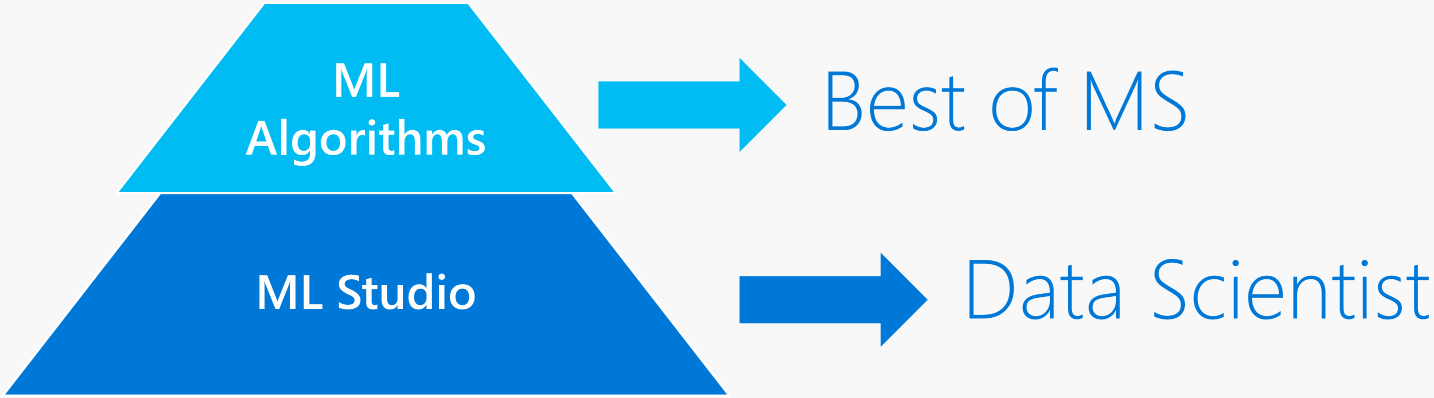


Azure Machine Learning

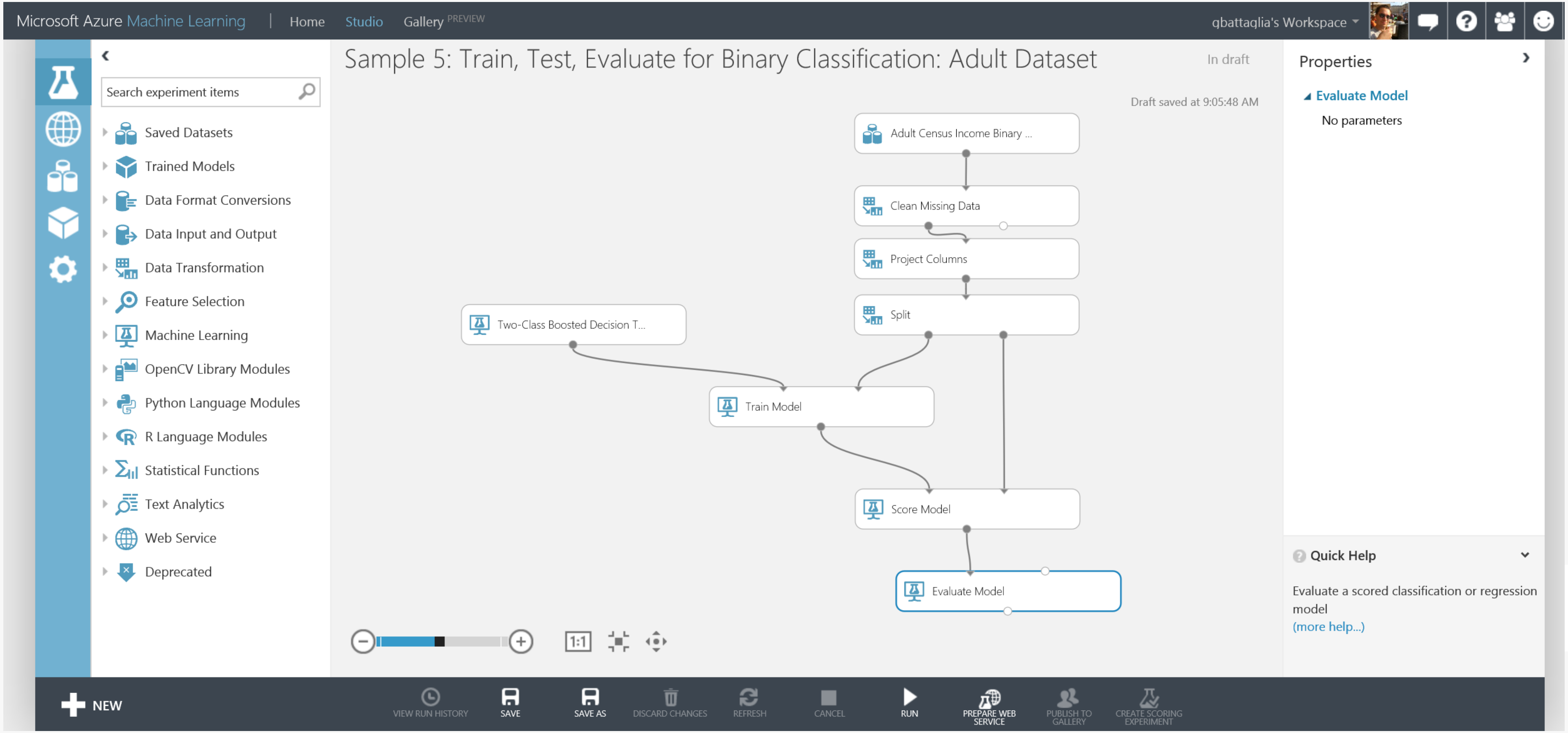
ML Algorithms → State of the Art Algorithms



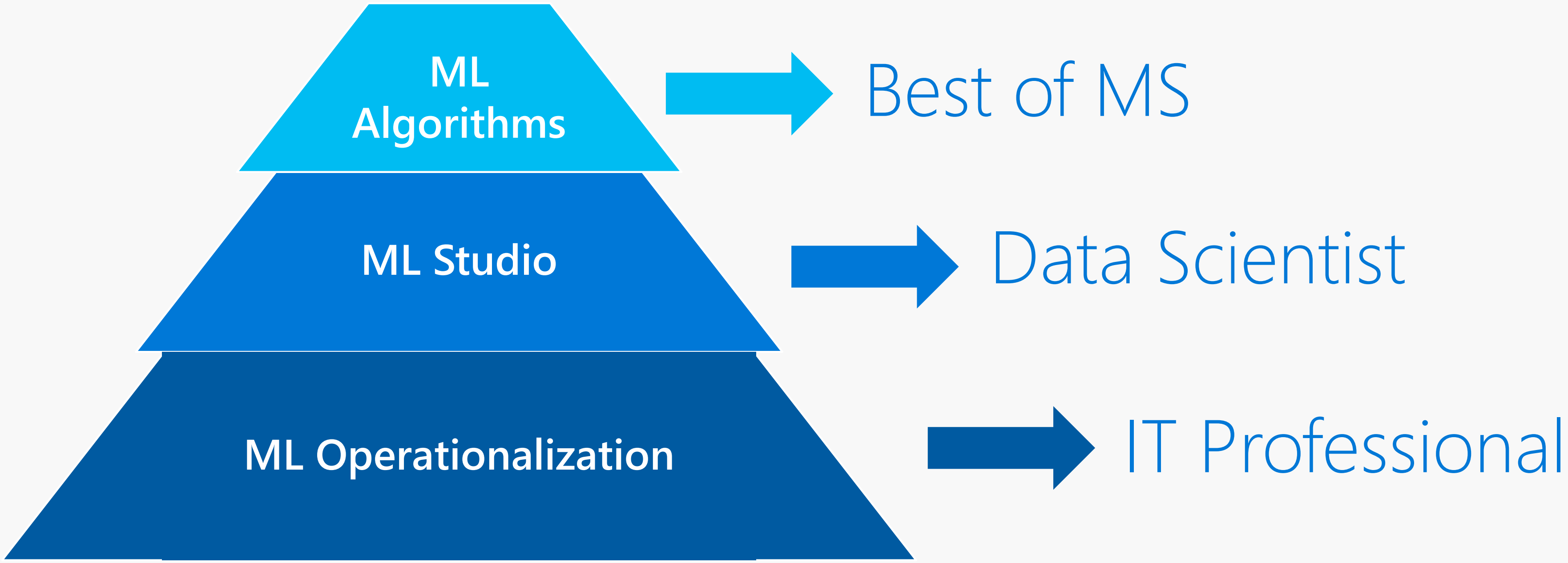
Azure Machine Learning



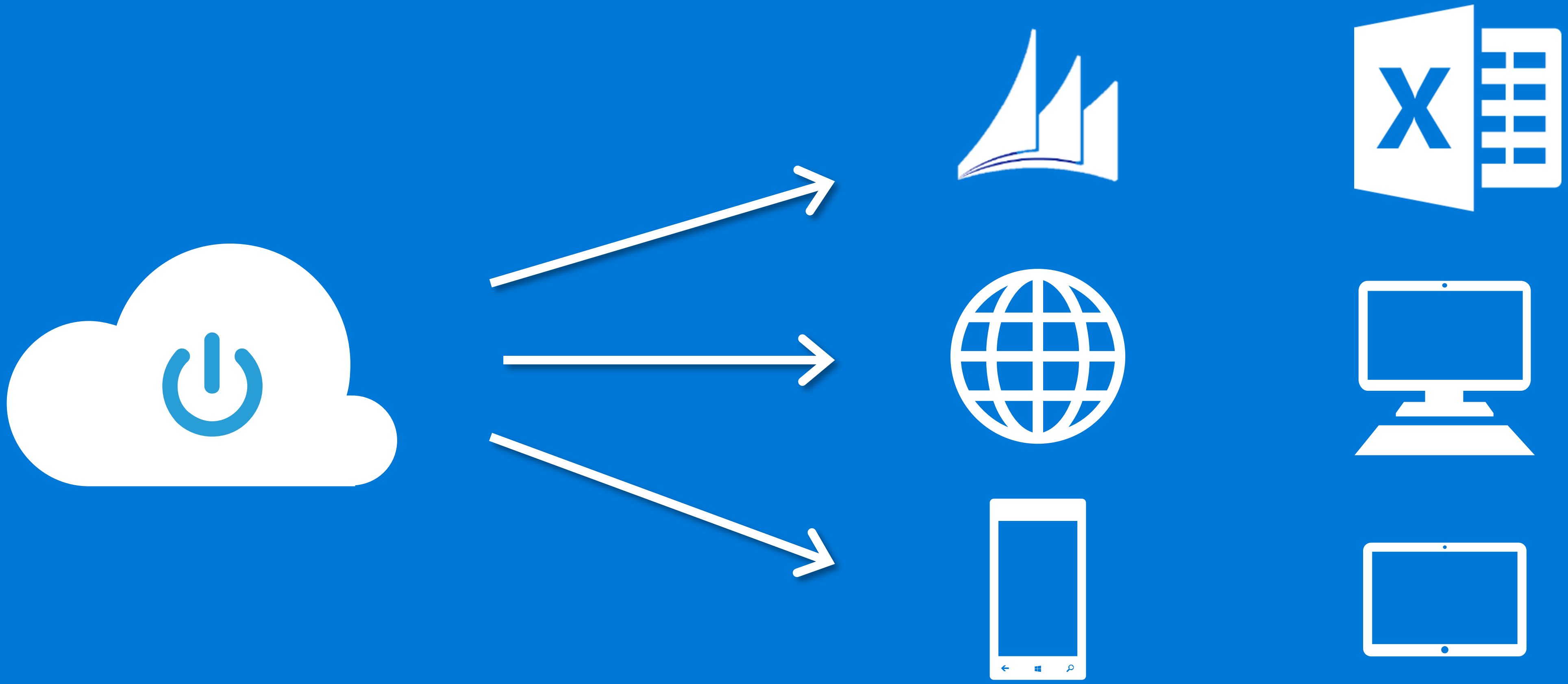
Azure Machine Learning



Azure Machine Learning

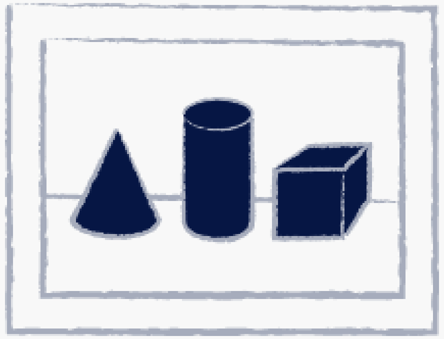


Azure Machine Learning

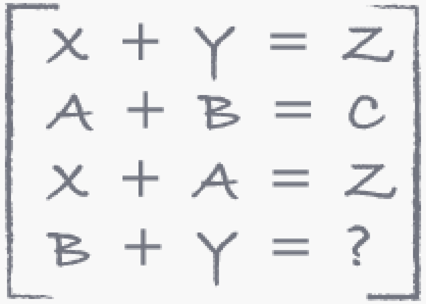


Consume anywhere

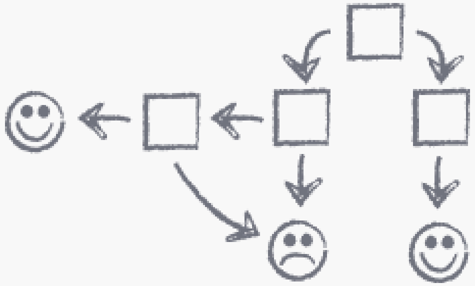
Machine Learning Process



Data



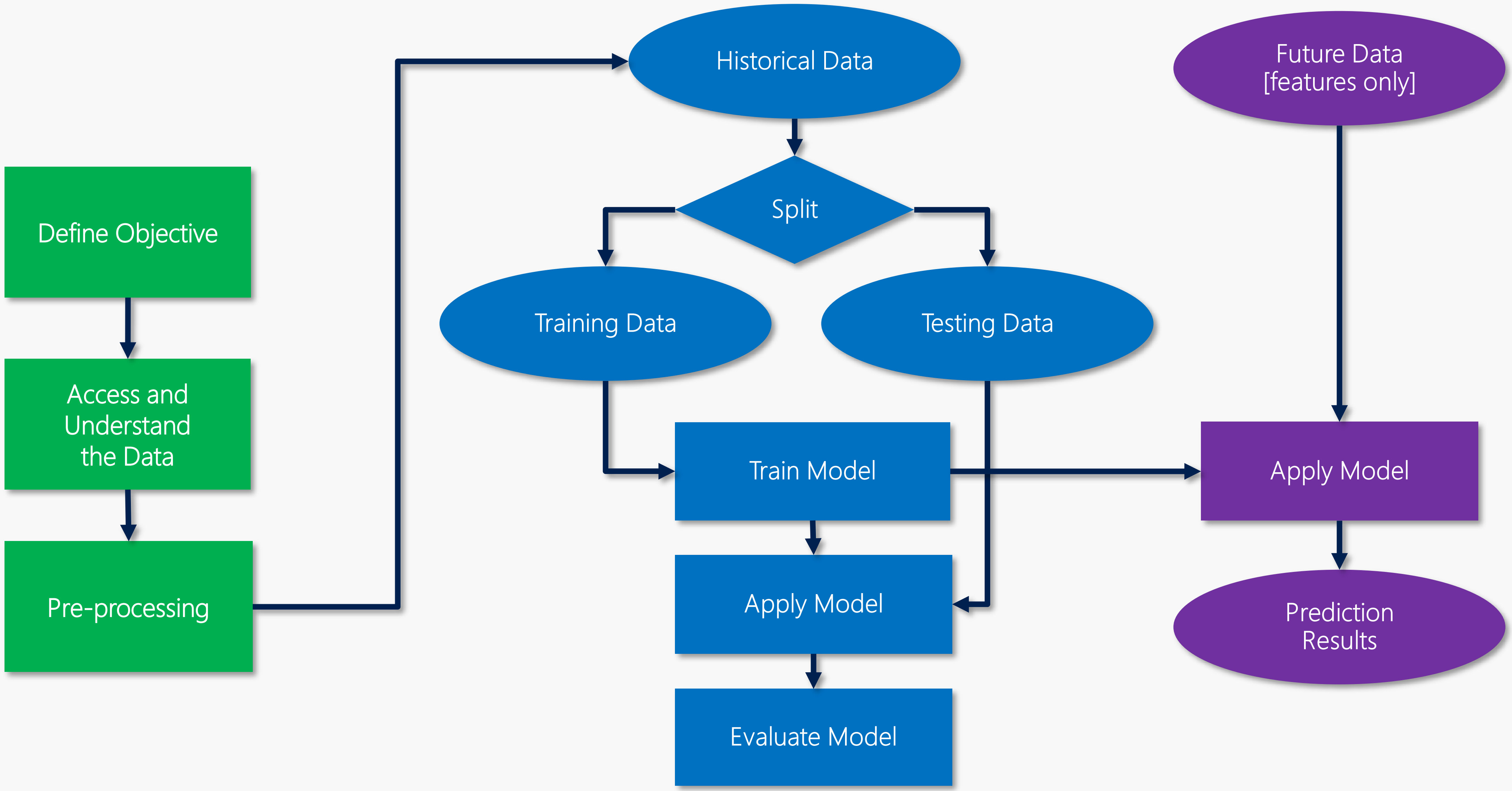
Clean
Transform
Maths



Build
Model

Predict

ML Process



Define Objective
Access and Understand the Data
Pre-processing
Historical Data
Split
Training Data
Testing Data
Train Model
Apply Model
Evaluate Model
Future Data [features only]
Apply Model
Prediction Results

“We want to do machine learning!”

Sharp / Well-Defined Question	Relevant Data	Accurate Data	Connected Data	A lot of data.
Predict whether component X will fail in the next Y days	Identifiers at the level they are predicting	Failures are really failures, human labels on root causes	Machine information linkable to usage information	Will be difficult to predict failure accurately with few examples

http://blogs.technet.com/b/machinelearning/archive/2015/08/26/what-can-data-science-do-for-me.aspx

Sample training data

~20k rows,
100 unique engine id

id	cycle	setting1	setting2	setting3	s1	s2	s3	...	s19	s20	s21
1	1	-0.0007	-0.0004	100	518.67	641.82	1589.7		100	39.06	23.419
1	2	0.0019	-0.0003	100	518.67	642.15	1591.82		100	39	23.4236
1	3	-0.0043	0.0003	100	518.67	642.35	1587.99		100	38.95	23.3442
...	...										
1	191	0	-0.0004	100	518.67	643.34	1602.36		100	38.45	23.1295
1	192	0.0009	0	100	518.67	643.54	1601.41		100	38.48	22.9649
2	1	-0.0018	0.0006	100	518.67	641.89	1583.84		100	38.94	23.4585
2	2	0.0043	-0.0003	100	518.67	641.82	1587.05		100	39.06	23.4085
2	3	0.0018	0.0003	100	518.67	641.55	1588.32		100	39.11	23.425
...	...										
2	286	-0.001	-0.0003	100	518.67	643.44	1603.63		100	38.33	23.0169
2	287	-0.0005	0.0006	100	518.67	643.85	1608.5		100	38.43	23.0848

Sample testing data

~13k rows,
100 unique engine id

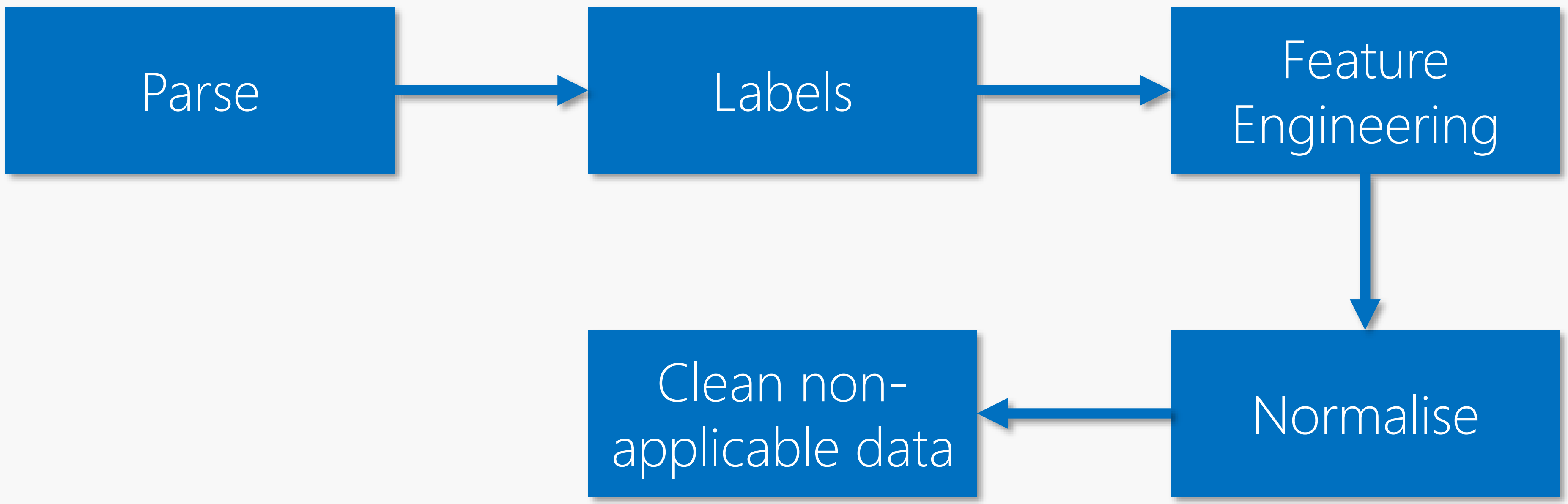
id	cycle	setting1	setting2	setting3	s1	s2	s3	...	s19	s20	s21
1	1	0.0023	0.0003	100	518.67	643.02	1585.29		100	38.86	23.3735
1	2	-0.0027	-0.0003	100	518.67	641.71	1588.45		100	39.02	23.3916
1	3	0.0003	0.0001	100	518.67	642.46	1586.94		100	39.08	23.4166
...	...										
1	30	-0.0025	0.0004	100	518.67	642.79	1585.72		100	39.09	23.4069
1	31	-0.0006	0.0004	100	518.67	642.58	1581.22		100	38.81	23.3552
2	1	-0.0009	0.0004	100	518.67	642.66	1589.3		100	39	23.3923
2	2	-0.0011	0.0002	100	518.67	642.51	1588.43		100	38.84	23.2902
2	3	0.0002	0.0003	100	518.67	642.58	1595.6		100	39.02	23.4064
...	...										
2	48	0.0011	-0.0001	100	518.67	642.64	1587.71		100	38.99	23.2918
2	49	0.0018	-0.0001	100	518.67	642.55	1586.59		100	38.81	23.2618
3	1	-0.0001	0.0001	100	518.67	642.03	1589.92		100	38.99	23.296
3	2	0.0039	-0.0003	100	518.67	642.23	1597.31		100	38.84	23.3191
3	3	0.0006	0.0003	100	518.67	642.98	1586.77		100	38.69	23.3774
...	...										
3	125	0.0014	0.0002	100	518.67	643.24	1588.64		100	38.56	23.227
3	126	-0.0016	0.0004	100	518.67	642.88	1589.75		100	38.93	23.274

Sample ground truth data

100 rows

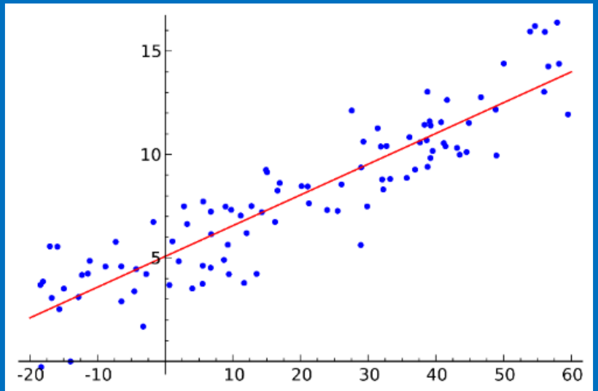
RUL
112
98
69
82
91

Pre-Processing

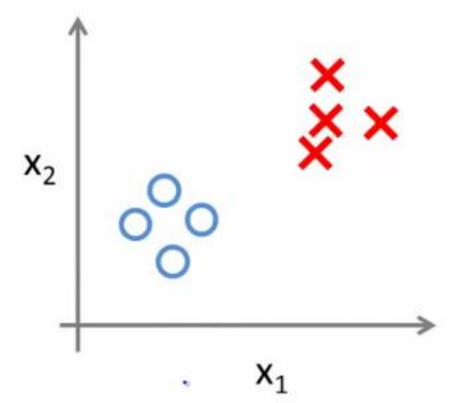


Modelling Techniques

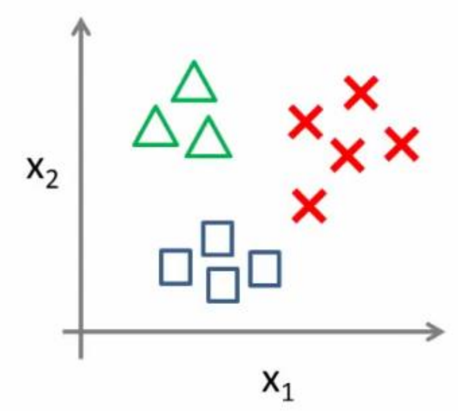
Regression / Survival Analysis



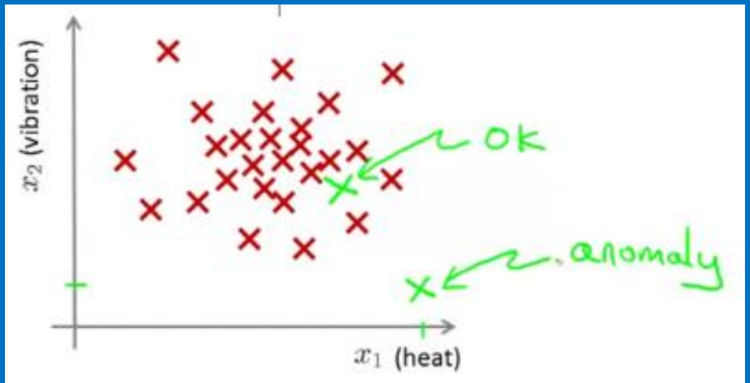
Binary Classification



Multi-Class Classification



Anomaly Detection



https://upload.wikimedia.org/wikipedia/commons/thumb/3/3a/Linear_regression.svg/2000px-Linear_regression.svg.png
http://www.holehouse.org/mlclass/06_Logistic_Regression_files/Image%20[23].png
http://dnene.bitbucket.org/docs/mlclass-notes/_images/aircraft_engines.png

Modelling Techniques

Regression / Survival Analysis

Predict failures within a future period of time

Binary Classification

Predict remaining useful life, the amount of time before the next failure

Multi-Class Classification

Predict failures with their causes within a future time period.

Predict remaining useful life within ranges of future periods

Anomaly Detection

Identify change in normal trends to find anomalies

Data Labeling

Regression

Binary Classification

Multi-Class Classification

id	cycle	...	RUL	label1	label2
1	1		191	0	0
1	2		190	0	0
1	3		189	0	0
1	4		188	0	0
...			...	...	...
1	160		32	0	0
1	161		31	0	0
1	162		30	1	1
1	163		29	1	1
1	164		28	1	1
1	165		27	1	1
1	166		26	1	1
1	167		25	1	1
1	168		24	1	1
1	169		23	1	1
1	170		22	1	1
1	171		21	1	1
1	172		20	1	1
1	173		19	1	1
1	174		18	1	1
1	175		17	1	1
1	176		16	1	1
1	177		15	1	2
1	178		14	1	2
1	179		13	1	2
1	180		12	1	2
1	181		11	1	2
1	182		10	1	2
1	183		9	1	2
1	184		8	1	2
1	185		7	1	2
1	186		6	1	2
1	187		5	1	2
1	188		4	1	2
1	189		3	1	2
1	190		2	1	2
1	191		1	1	2

w1

w0

Predefined window size for classification models

w1 = 30
w0 = 15

Example Feature Engineering Methods

Create features that capture degradation over time

Rolling Aggregates

For each labelled record of an asset, pick a rolling window of size w, compute rolling aggregate features for the periods before the labelling date and time of that record.

Lag features for short term

For each labelled record of an asset, pick a window of size w and use tumbling windows to create aggregate features for the periods before the labelling date and time.

Lag features for long term

For each labelled record, find aggregated features for a larger window than w reflecting the long term effects.

Feature Engineering

The process of creating features that provide better or additional predictive power to the learning algorithm.

id	cycle	setting1	setting2	setting3	s1	s2	s3	...	s19	s20	s21
1	1	-0.0007	-0.0004	100	518.67	641.82	1589.7		100	39.06	23.419
1	2	0.0019	-0.0003	100	518.67	642.15	1591.82		100	39	23.4236
1	3	-0.0043	0.0003	100	518.67	642.35	1587.99		100	38.95	23.3442
...	...										
1	191	0	-0.0004	100	518.67	643.34	1602.36		100	38.45	23.1295
1	192	0.0009	0	100	518.67	643.54	1601.41		100	38.48	22.9649
2	1	-0.0018	0.0006	100	518.67	641.89	1583.84		100	38.94	23.4585
2	2	0.0043	-0.0003	100	518.67	641.82	1587.05		100	39.06	23.4085
2	3	0.0018	0.0003	100	518.67	641.55	1588.32		100	39.11	23.425
...	...										
2	286	-0.001	-0.0003	100	518.67	643.44	1603.63		100	38.33	23.0169
2	287	-0.0005	0.0006	100	518.67	643.85	1608.5		100	38.43	23.0848

a1	a2	...	a21	sd1	sd2	...	sd21	RUL	label1	label2

40+ engineered features

Other potential features: change from initial value, velocity of change, frequency count over a predefined threshold

COLLECTION

Predictive Maintenance Template

Microsoft • published on September 28, 2015

2380 views

Tweet | Share

TAGS

Predictive Maintenance | Failure Prediction | Regression | Classification

Published September 28, 2015
Last updated September 28, 2015

Report Abuse

Summary

This template demonstrate how to build and deploy predictive maintenance models to predict asset failures.

Description

Predictive maintenance encompasses a variety of topics, including but not limited to: failure prediction, failure diagnosis (root cause analysis), failure detection, failure type classification, and recommendation of mitigation or maintenance actions after failure. As part of the Azure Machine Learning offering, Microsoft provides a template that helps data scientists easily build and deploy a predictive maintenance solution. **This predictive maintenance template focuses on the techniques used to predict when an in-service machine will fail, so that maintenance can be planned in advance.** The template includes a collection of pre-configured machine learning modules, as well as custom R scripts in the *Execute R Script* module, to enable an end-to-end solution from data processing to deploying of the machine learning model.

Read full description

7 Items

EXPERIMENT

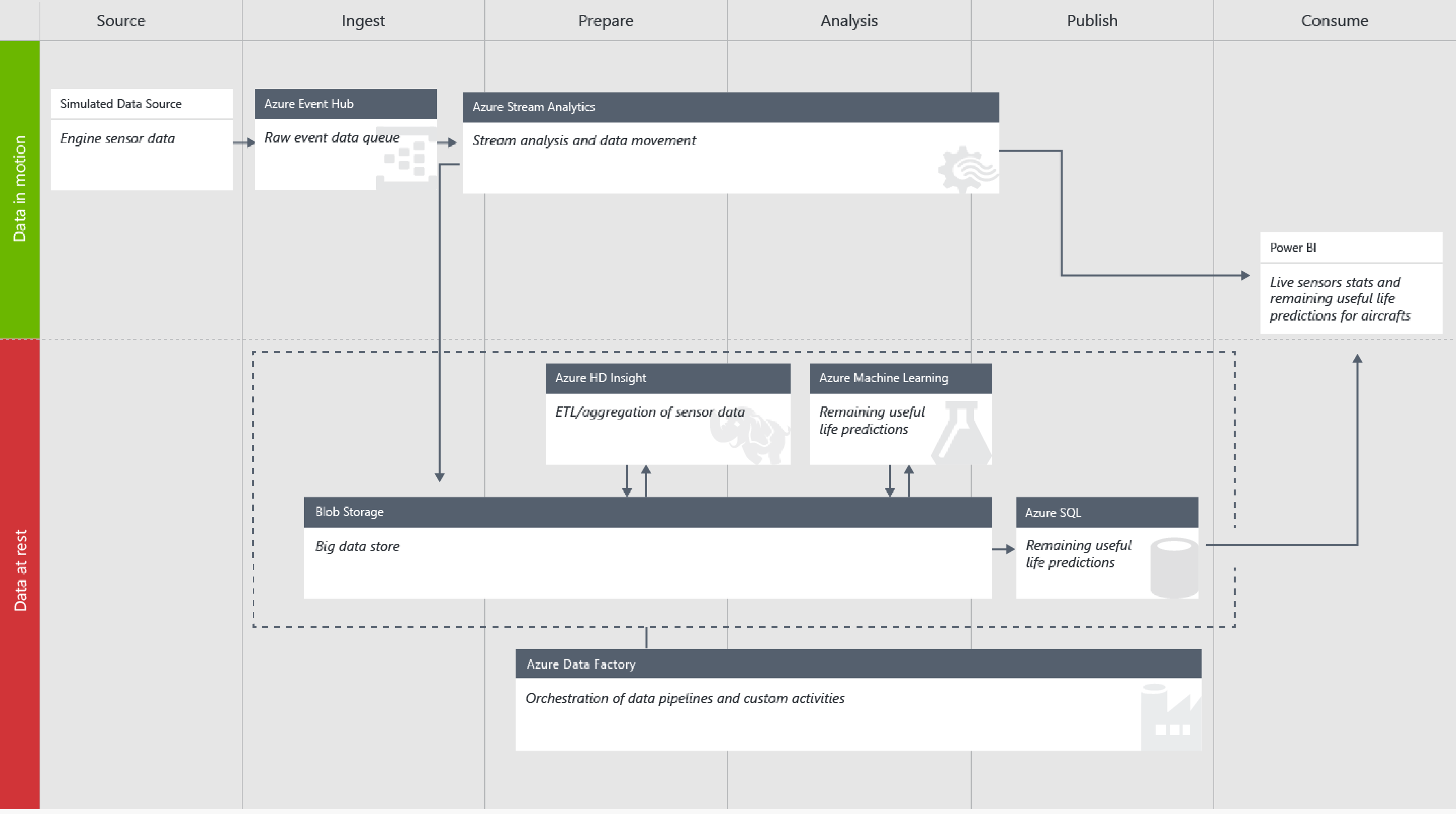
Predictive Maintenance: Step 1 of 3, data preparation and feature engineering

by Microsoft

https://gallery.cortanaanalytics.com/Collection/Predictive-Maintenance-Template-3

HOL 4: Machine Learning

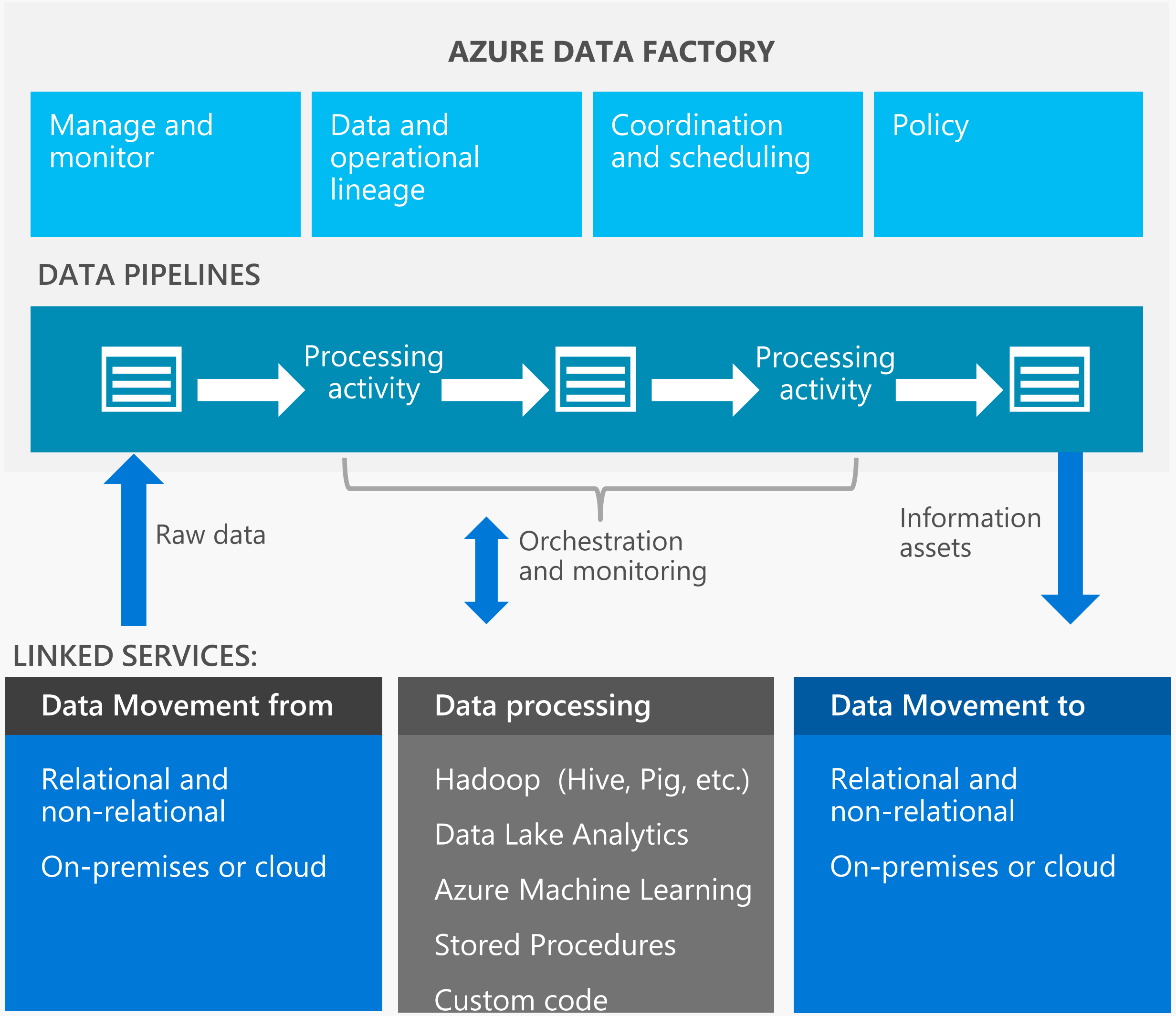
Step 7 (GitHub): Create Azure ML Workspace & Experiment


Source
Ingest
Prepare
Analysis
Publish
Consume
Data in motion
Data at rest
Simulated Data Source
Engine sensor data
Azure Event Hub
Raw event data queue
Azure Stream Analytics
Stream analysis and data movement
Power BI
Live sensors stats and remaining useful life predictions for aircrafts
Azure HD Insight
ETL/aggregation of sensor data
Azure Machine Learning
Remaining useful life predictions
Blob Storage
Big data store
Azure SQL
Remaining useful life predictions
Azure Data Factory
Orchestration of data pipelines and custom activities

Pre-Processing & Orchestration

Azure Data Factory

Compose services to transform data into actionable intelligence



Linked Services

- Connect data factories to the resources and services you want to use
- Connect to data stores like Azure Storage and on premises SQL Server
- Connect to compute services like Azure ML, Azure HDI, and Azure Batch

Data Sets

- A named reference/pointer to data you want to use as an input or output of an activity

Activities

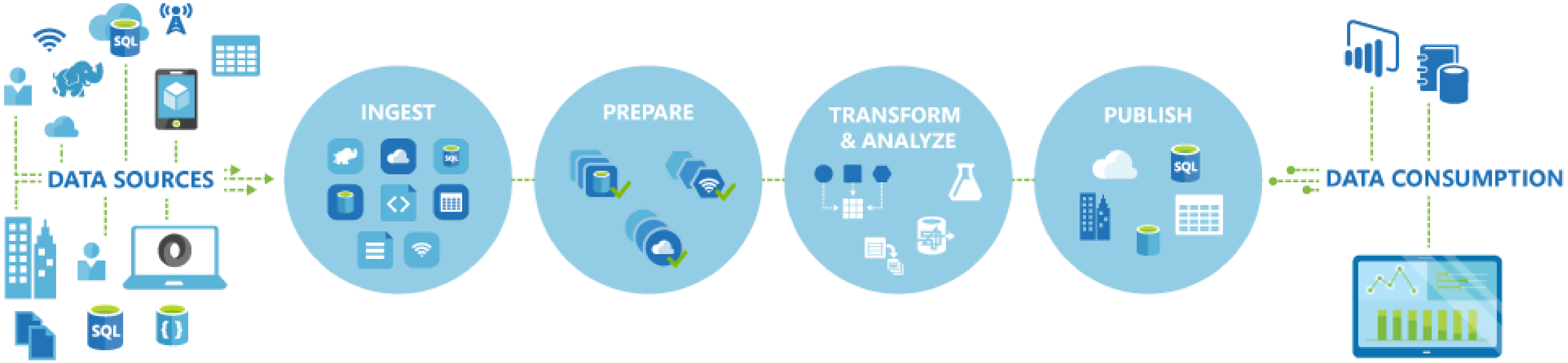
- Actions you perform on your data
- Takes inputs and produce outputs

Pipelines

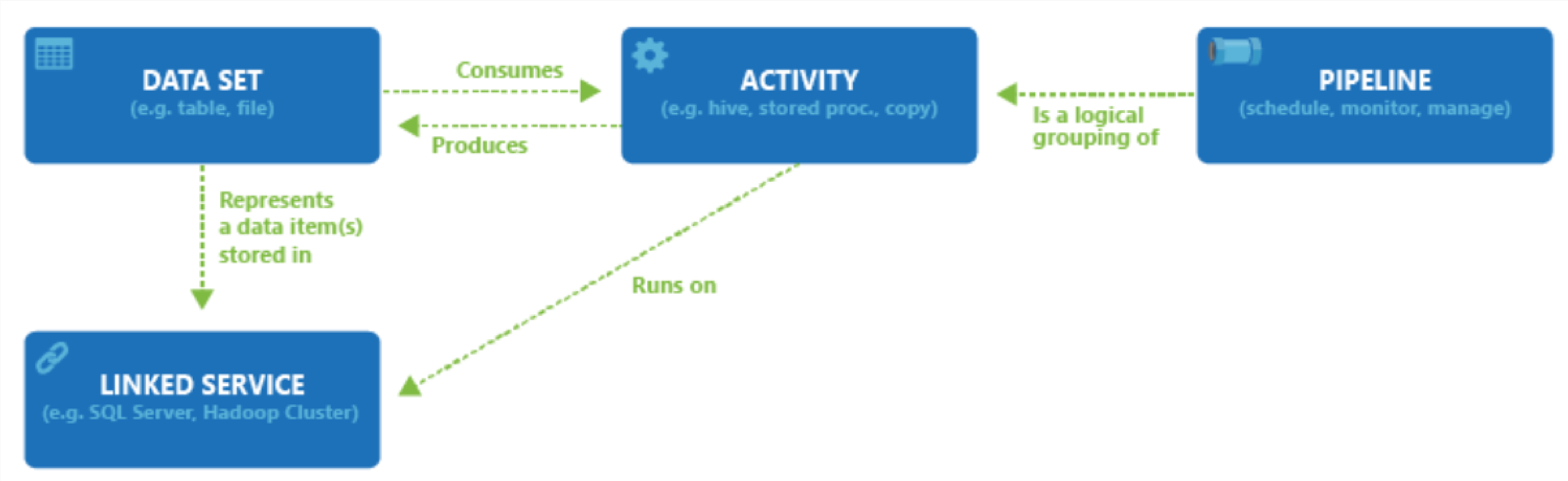
- Logical grouping of activities for group operations

ADF Logical Flow

Overview diagram

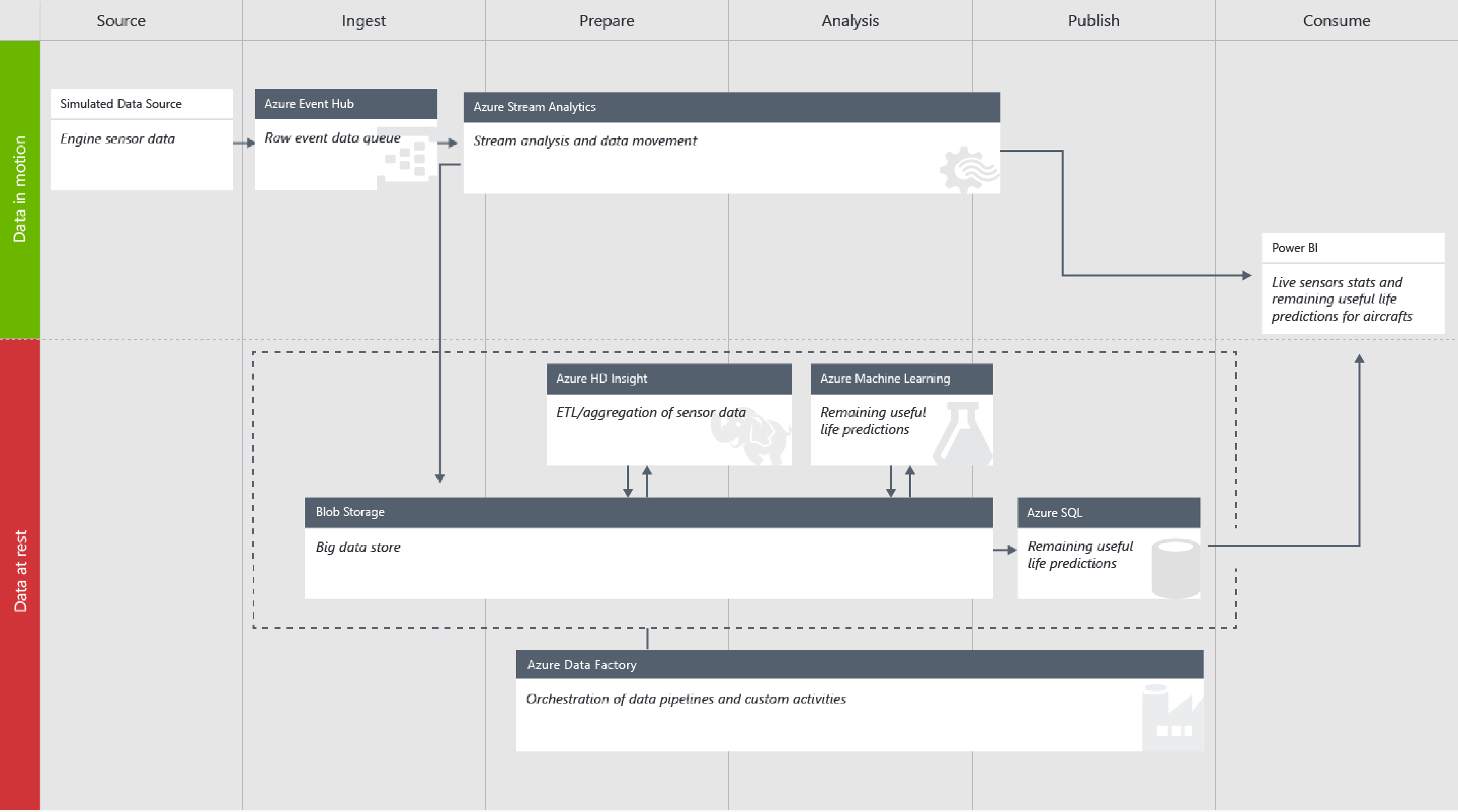


ADF Components

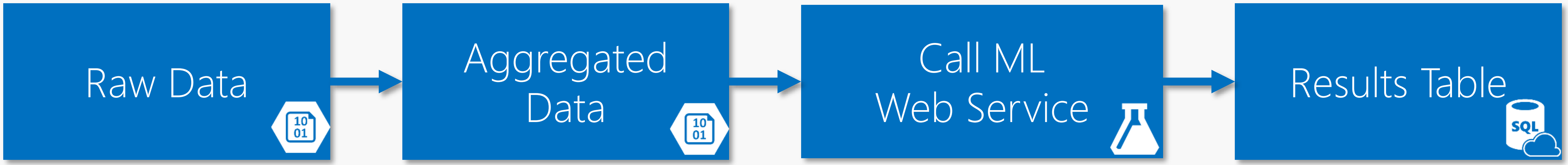


ADF Process

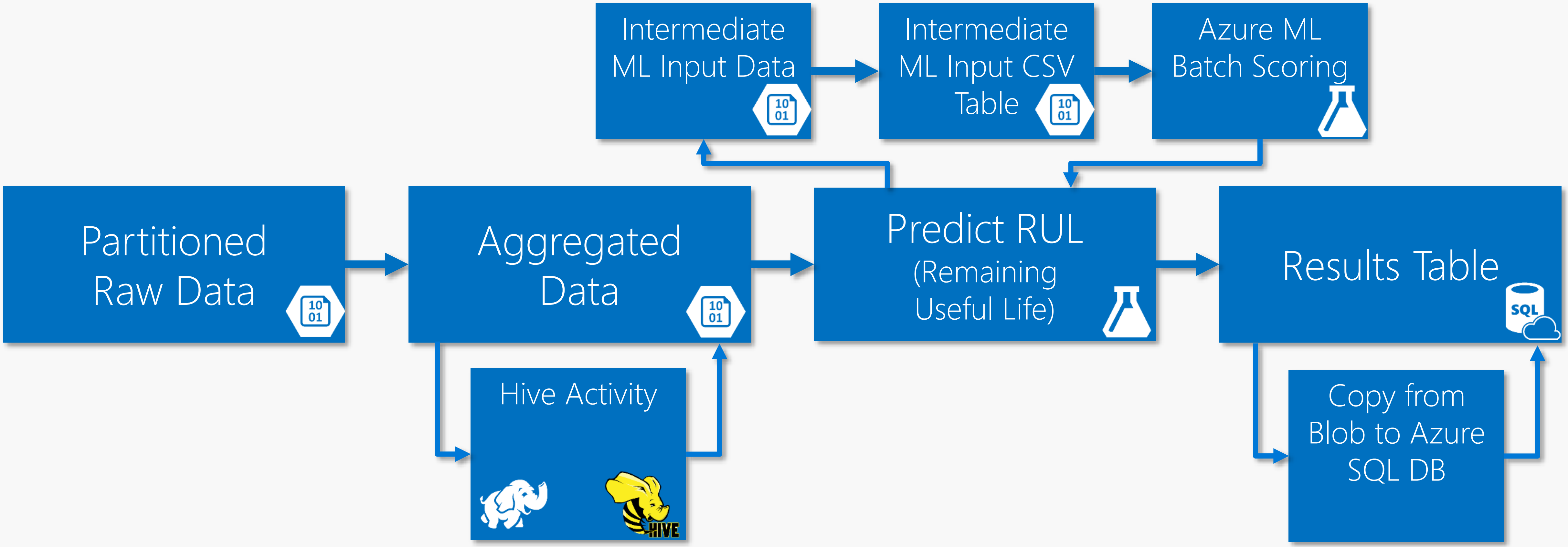
1. Define Architecture: Set up objectives and flow
2. Create the Data Factory: Portal or PowerShell
3. Create Linked Services: Connections to Data and Services
4. Create Datasets: Input and Output
5. Create Pipeline: Define Activities
6. Monitor and Manage: Portal or PowerShell, Alerts and Metrics


Source
Ingest
Prepare
Analysis
Publish
Consume
Data in motion
Data at rest
Simulated Data Source
Engine sensor data
Azure Event Hub
Raw event data queue
Azure Stream Analytics
Stream analysis and data movement
Power BI
Live sensors stats and remaining useful life predictions for aircrafts
Azure HD Insight
ETL/aggregation of sensor data
Azure Machine Learning
Remaining useful life predictions
Blob Storage
Big data store
Azure SQL
Remaining useful life predictions
Azure Data Factory
Orchestration of data pipelines and custom activities

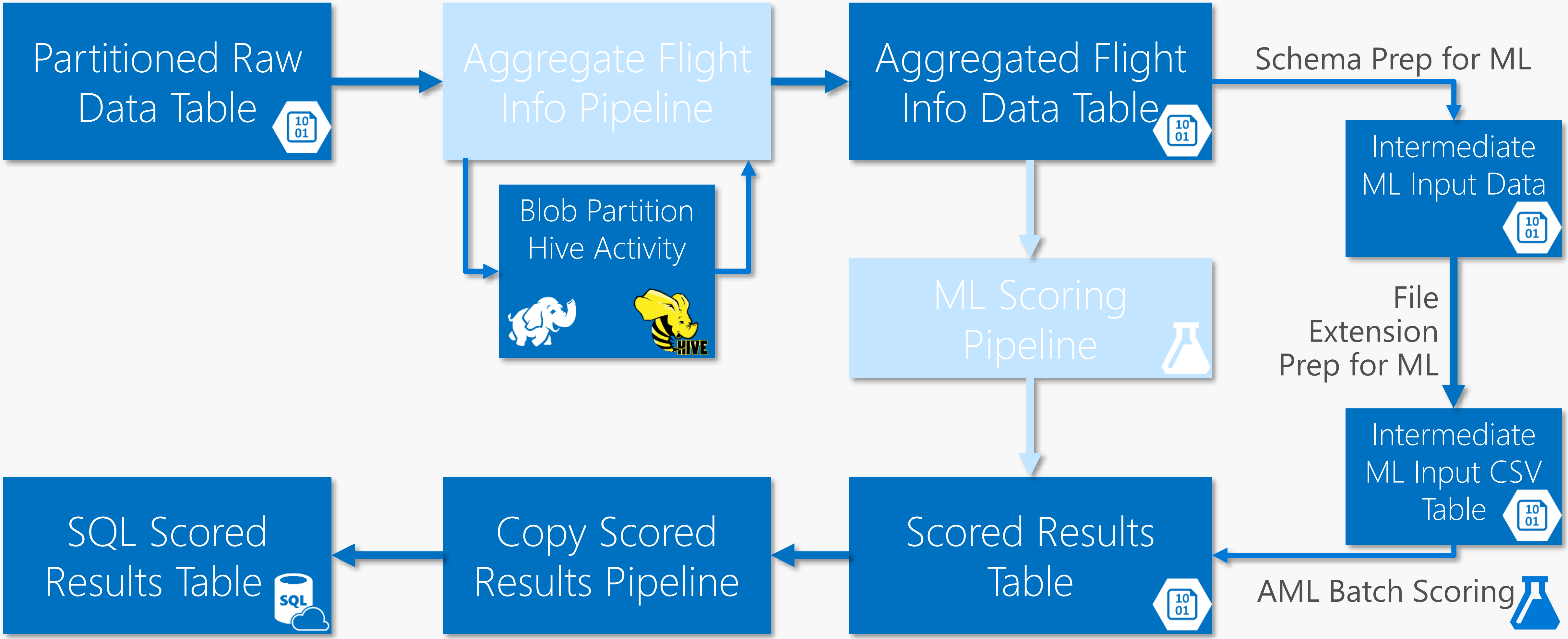
Pre-Processing in Predictive Maintenance



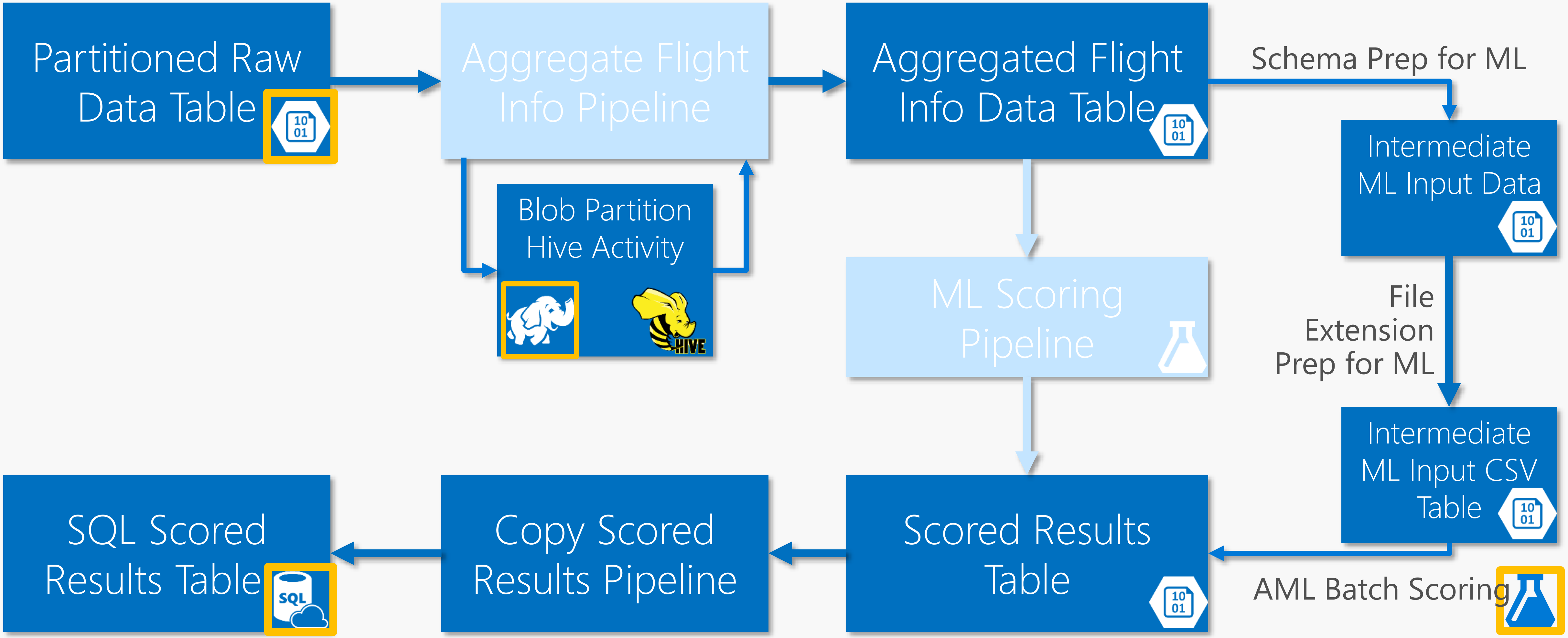
Pre-Processing in Predictive Maintenance



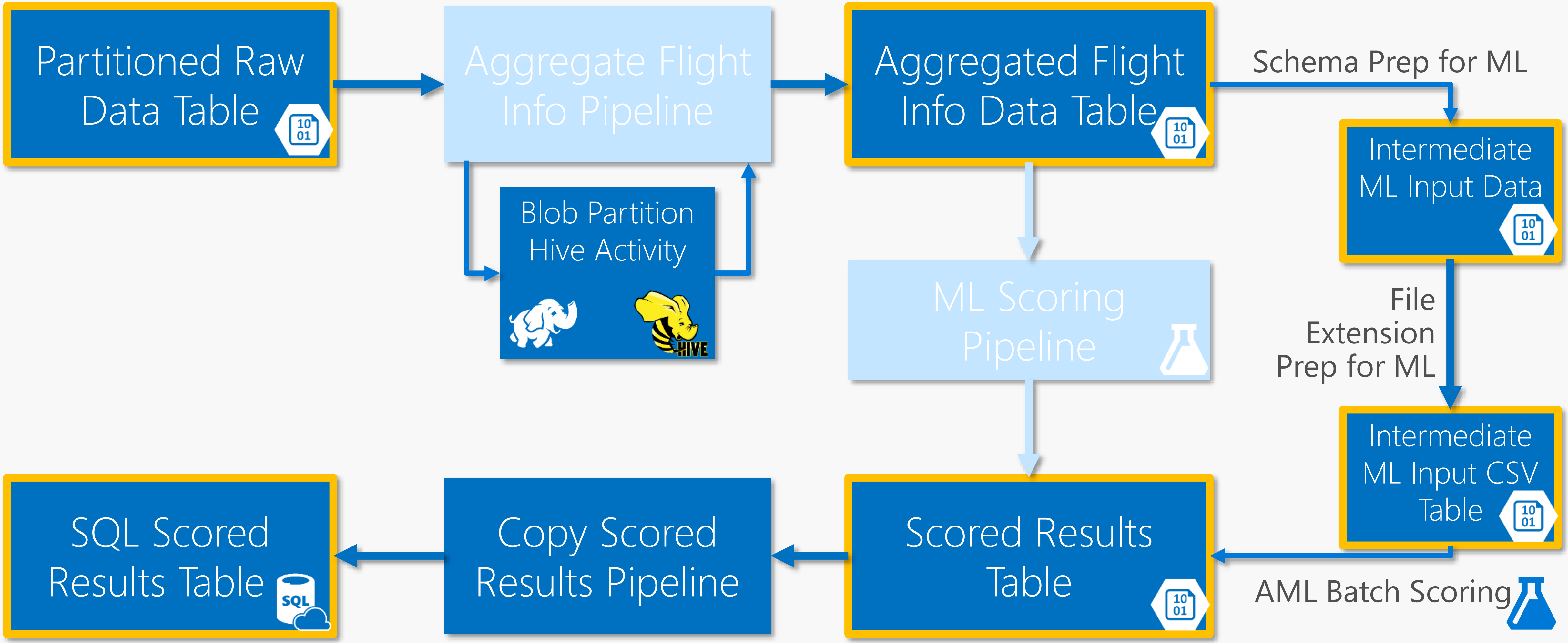
ADF in Predictive Maintenance



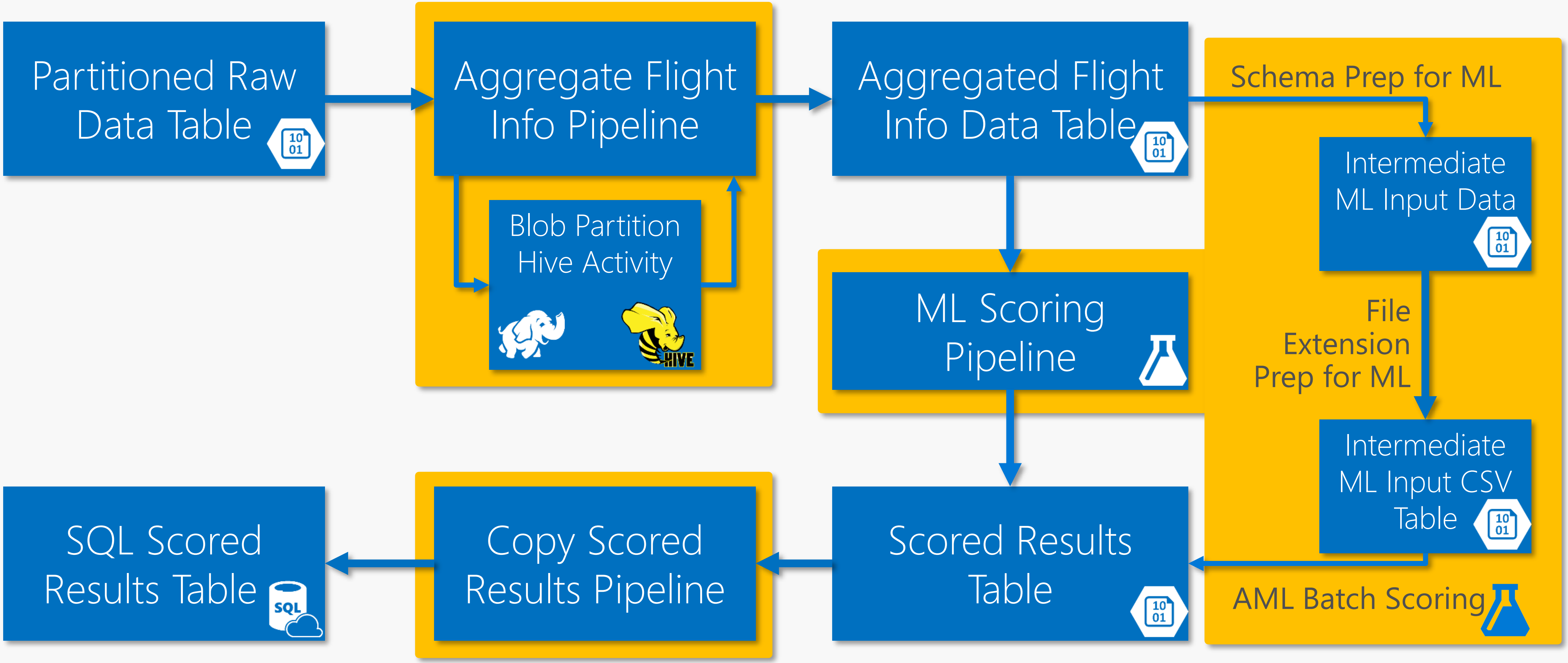
ADF – Linked Services



ADF – Datasets



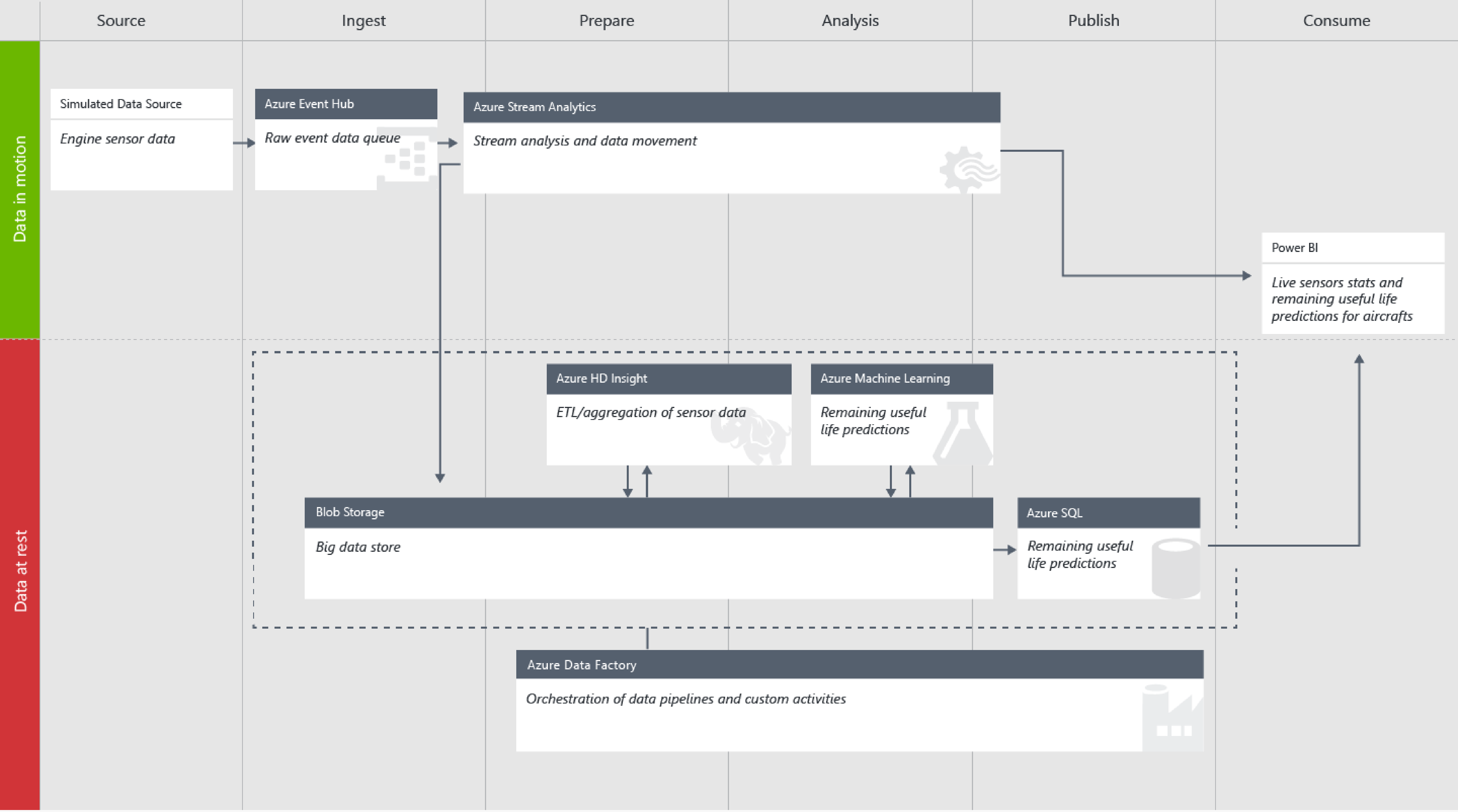
ADF – Pipelines

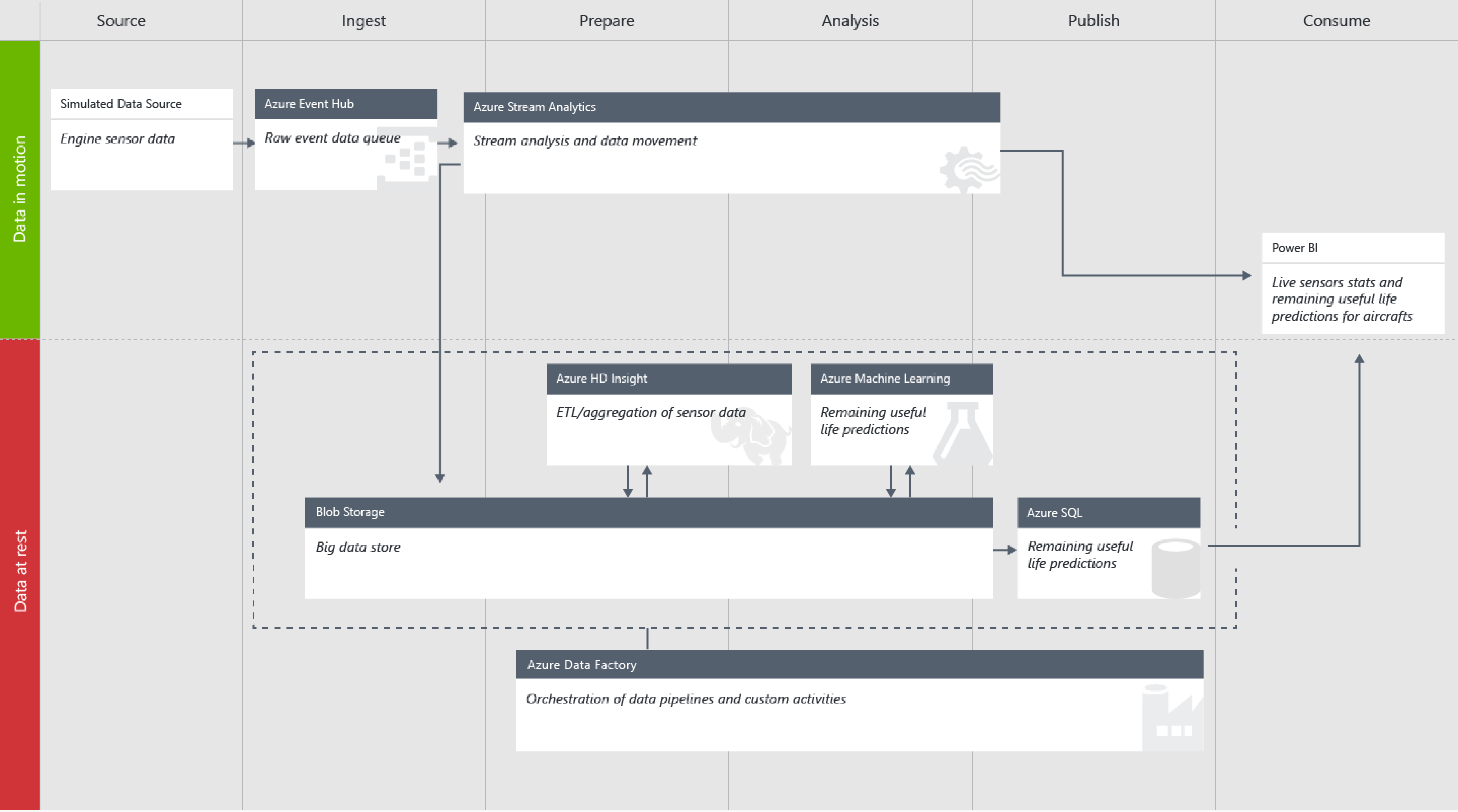


HOL 5: Pre-Processing and Orchestration

Step 8 (GitHub): Azure Data Factory

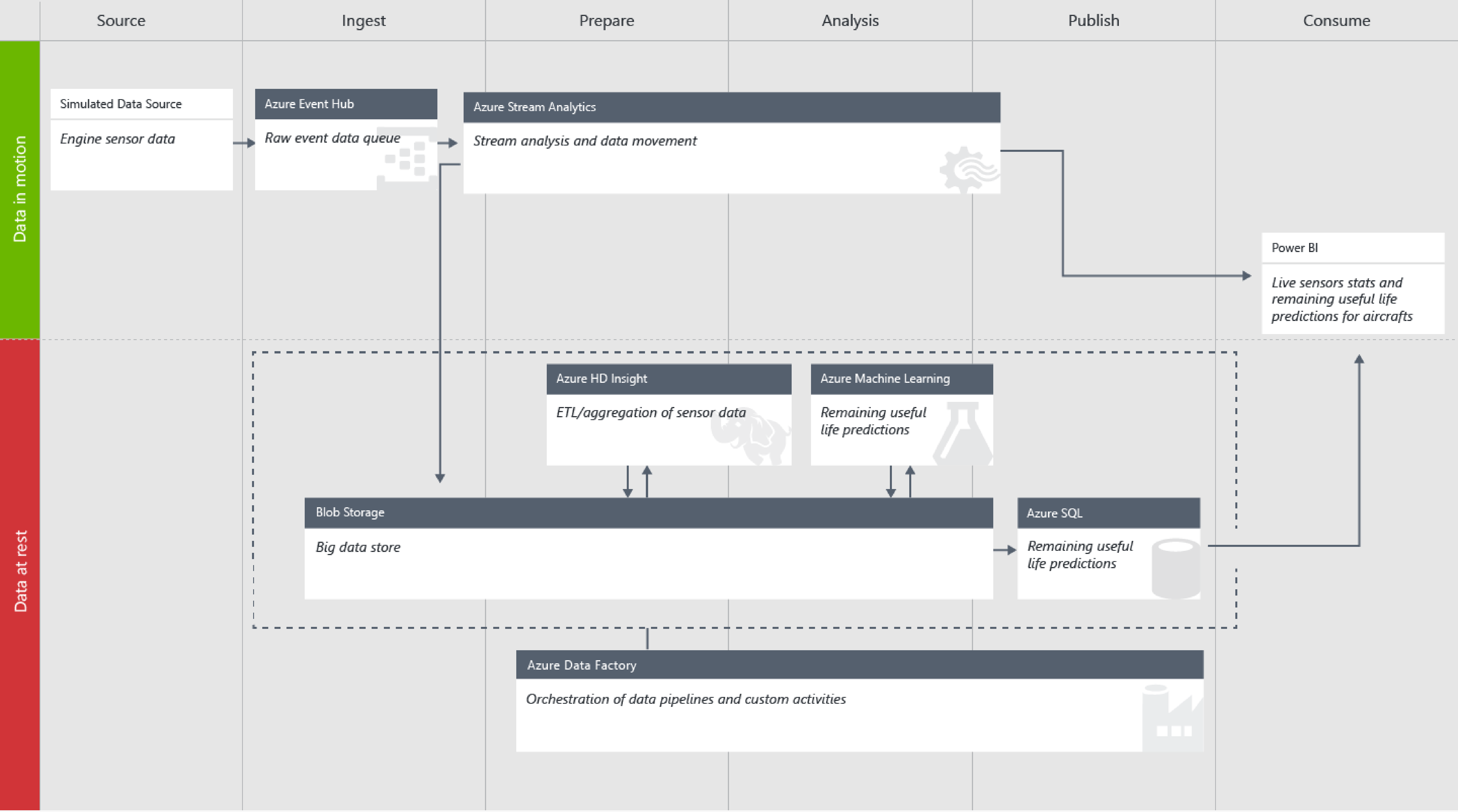
Visualisation


Source
Ingest
Prepare
Analysis
Publish
Consume
Data in motion
Data at rest
Simulated Data Source
Engine sensor data
Azure Event Hub
Raw event data queue
Azure Stream Analytics
Stream analysis and data movement
Power BI
Live sensors stats and remaining useful life predictions for aircrafts
Azure HD Insight
ETL/aggregation of sensor data
Azure Machine Learning
Remaining useful life predictions
Blob Storage
Big data store
Azure SQL
Remaining useful life predictions
Azure Data Factory
Orchestration of data pipelines and custom activities



Source
Ingest
Prepare
Analysis
Publish
Consume
Data in motion
Data at rest
Simulated Data Source
Engine sensor data
Azure Event Hub
Raw event data queue
Azure Stream Analytics
Stream analysis and data movement
Power BI
Live sensors stats and remaining useful life predictions for aircrafts
Azure HD Insight
ETL/aggregation of sensor data
Azure Machine Learning
Remaining useful life predictions
Blob Storage
Big data store
Azure SQL
Remaining useful life predictions
Azure Data Factory
Orchestration of data pipelines and custom activities

Wrapping Up…


Source
Ingest
Prepare
Analysis
Publish
Consume
Data in motion
Data at rest
Simulated Data Source
Engine sensor data
Azure Event Hub
Raw event data queue
Azure Stream Analytics
Stream analysis and data movement
Power BI
Live sensors stats and remaining useful life predictions for aircrafts
Azure HD Insight
ETL/aggregation of sensor data
Azure Machine Learning
Remaining useful life predictions
Blob Storage
Big data store
Azure SQL
Remaining useful life predictions
Azure Data Factory
Orchestration of data pipelines and custom activities

gallery.cortanaanalytics.com/SolutionTemplate/Predictive-Maintenance-for-Aerospace-1

Cortana Analytics Gallery

Sign in

Browse all | Solution Templates | Experiments | Machine Learning APIs | Notebooks | Tutorials | Collections

SOLUTION TEMPLATE

Predictive Maintenance for Aerospace

Microsoft • published on February 16, 2016

Deploy

Preview

+ Add to Collection

1475 views

506 deployments

Tweet | Share

SERVICES USED

- Azure Event Hubs
- Azure Stream Analytics
- Azure Machine Learning
- Azure Data Factory
- Azure HDInsight
- Azure Blob Storage
- Azure SQL
- PowerBI

Summary

Air travel is central to modern life, however, aircraft engines are expensive and keeping them up and running requires frequent maintenance by highly skilled technicians. Modern aircraft engines are equipped with highly sophisticated sensors to track the functioning of these machines. By combining the data from these sensors with advanced analytics it's possible to both monitor the aircraft in real time, as well as predict the remaining useful life of an engine component so that maintenance can be scheduled in a way to prevent mechanical failures.

The Cortana Analytics Predictive Maintenance for Aerospace Solution Template monitors aircraft and predicts the remaining useful life of aircraft engine components.

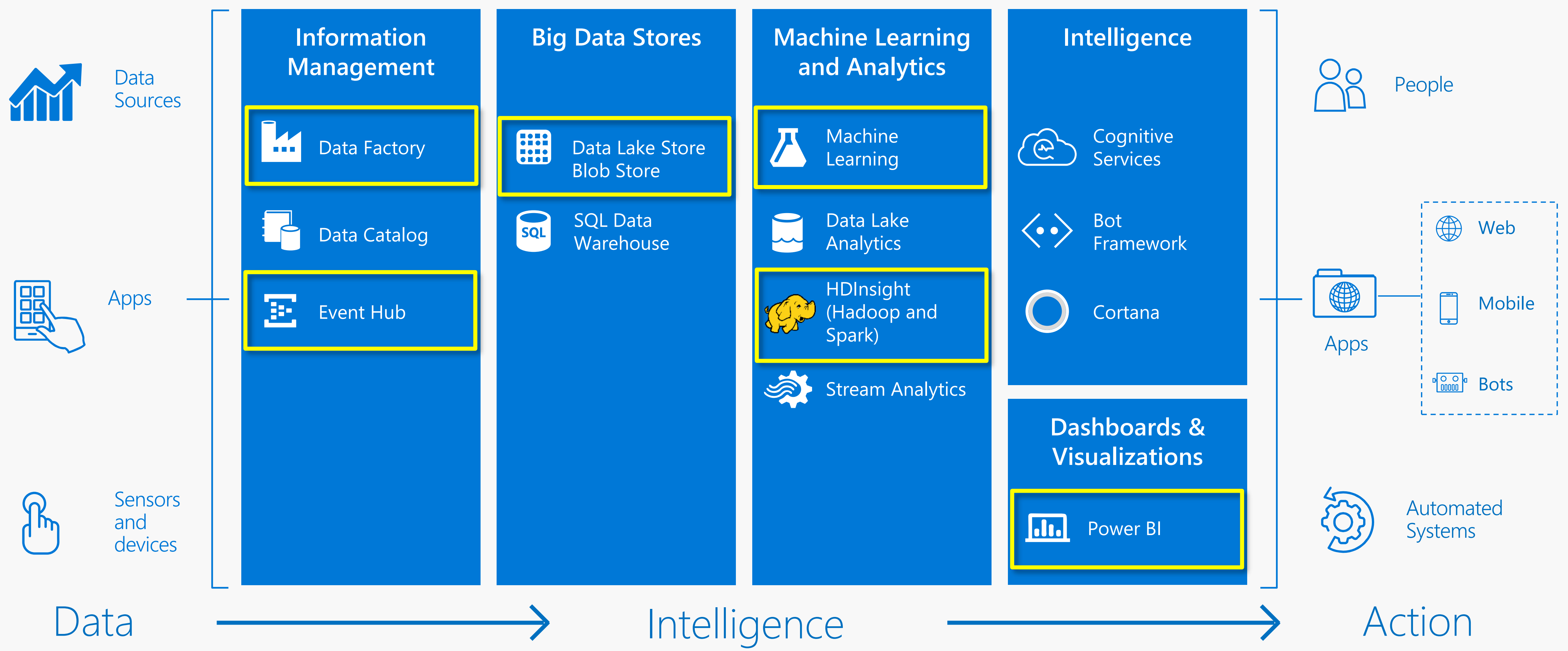
Description

Cortana Analytics opens new possibilities in the predictive maintenance space, including data ingestion, data storage, data processing and advanced analytics components—all the essential elements for building a predictive maintenance solution. While this solution is customised for aircraft monitoring, it can very easily be generalised for other predictive maintenance scenarios.

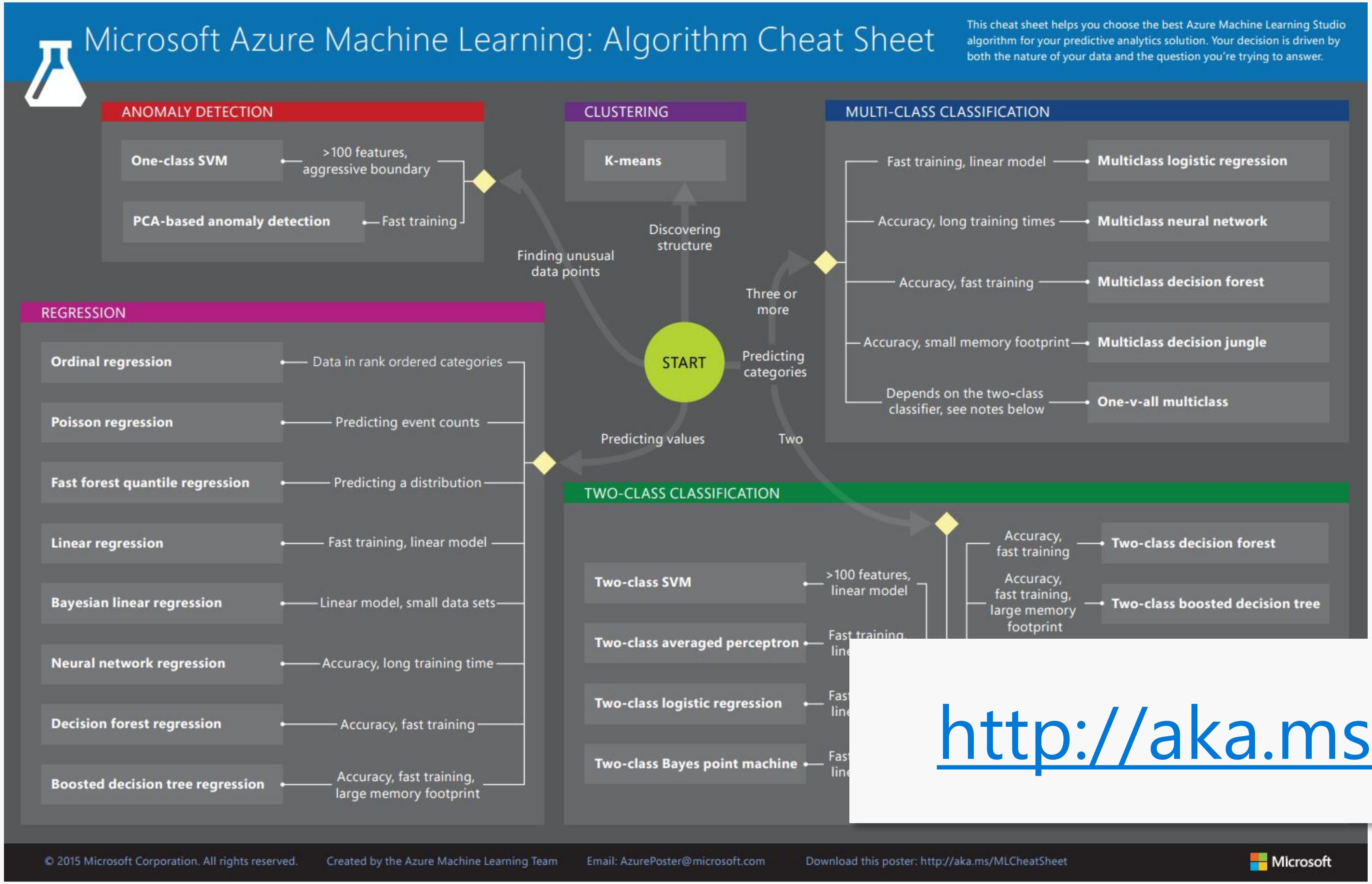
The solution template uses several Azure services, such as Event Hubs for ingesting aircraft sensor readings into Azure. Stream Analytics provides real-time insights on engine health and stores that data in long-term storage for more complex, compute-intensive batch analytics. HDInsight transforms the sensor

https://gallery.cortanaanalytics.com/SolutionTemplate/Predictive-Maintenance-for-Aerospace-1

Cortana Intelligence Suite



When which algorithm?



http://aka.ms/MLCheatSheet

http://blogs.technet.com/b/machinelearning/archive/2015/09/22/how-to-find-an-algorithm-that-fits.aspx

ML Algorithms – Explained!

http://aka.ms/MLexplained

Thanks!