BUDAPEST DATA

Gépi tanulás és fejlett analitikai lehetőségek az Azure-ban

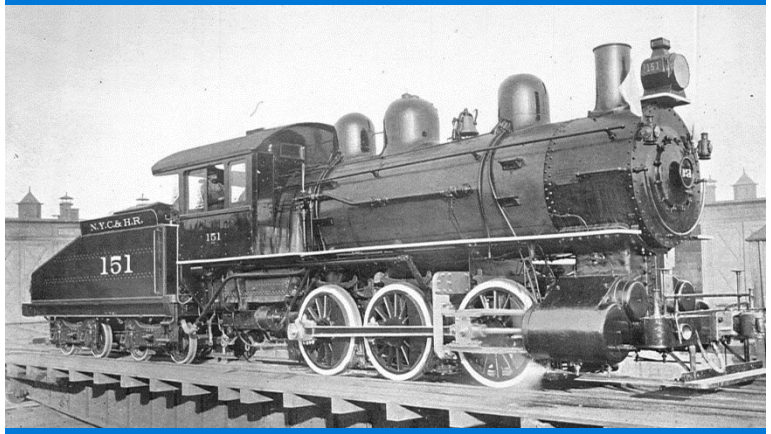
Dudás Viktor – Microsoft

Budapest, 2016. június 16.

Microsoft

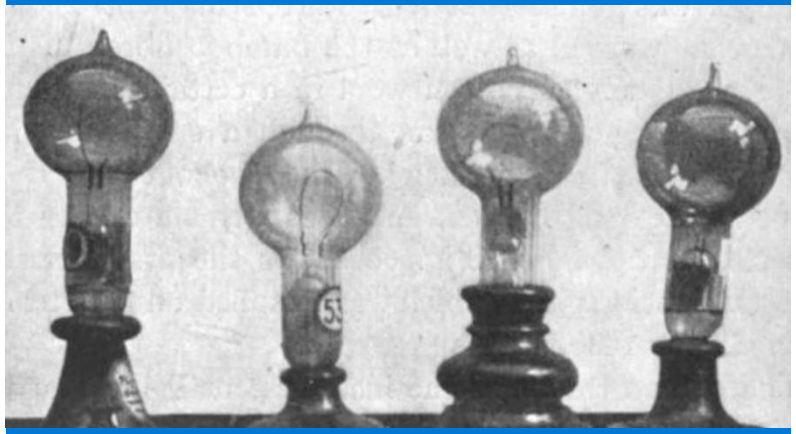
Néhány újítás alapjaiban zavarta meg a gazdasági helyzetet

1. Ipari forradalom



1760

2. Ipari forradalom

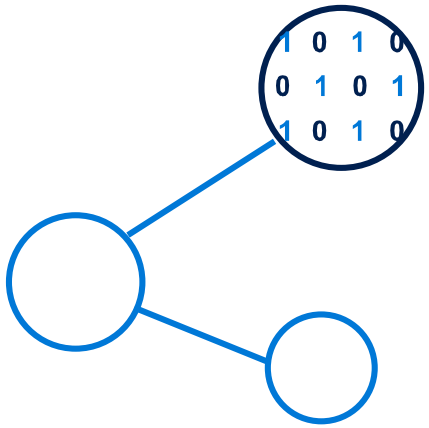


1870

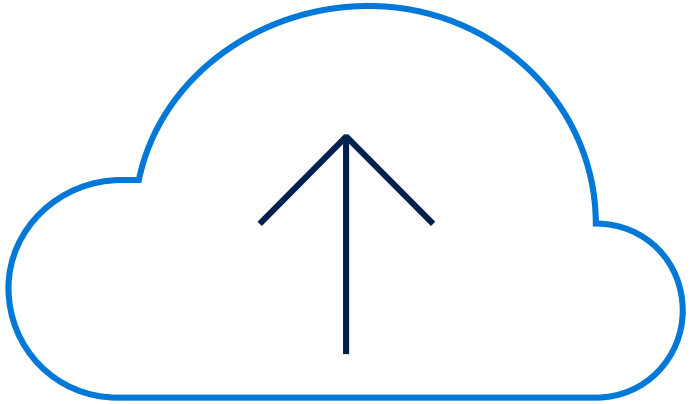
3. Ipari forradalom

1950

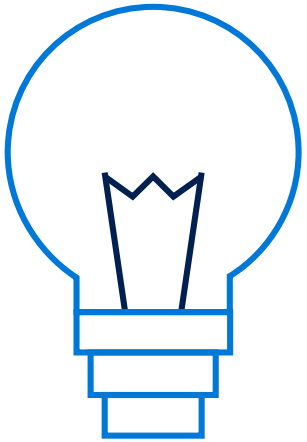
Napjainkban zajlik a 4. Ipari forradalom – a Digitalizáció



Big Data



Felhő



Intelligencia

Az adat a kulcs

$1.6T

Additional business value captured by companies that are leaders in using data assets to their advantage

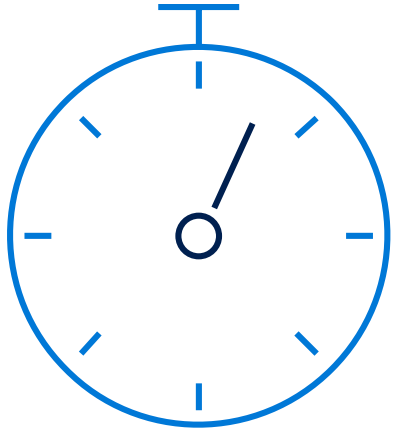
Source: IDC, 2014

10%

Percent of organizations expected to have a highly profitable business unit specifically for productizing and commercializing their data by 2020

Source: Gartner, 2016

Hatalmas tárolási és számítási kapacitás bárki számára az Azure-ban

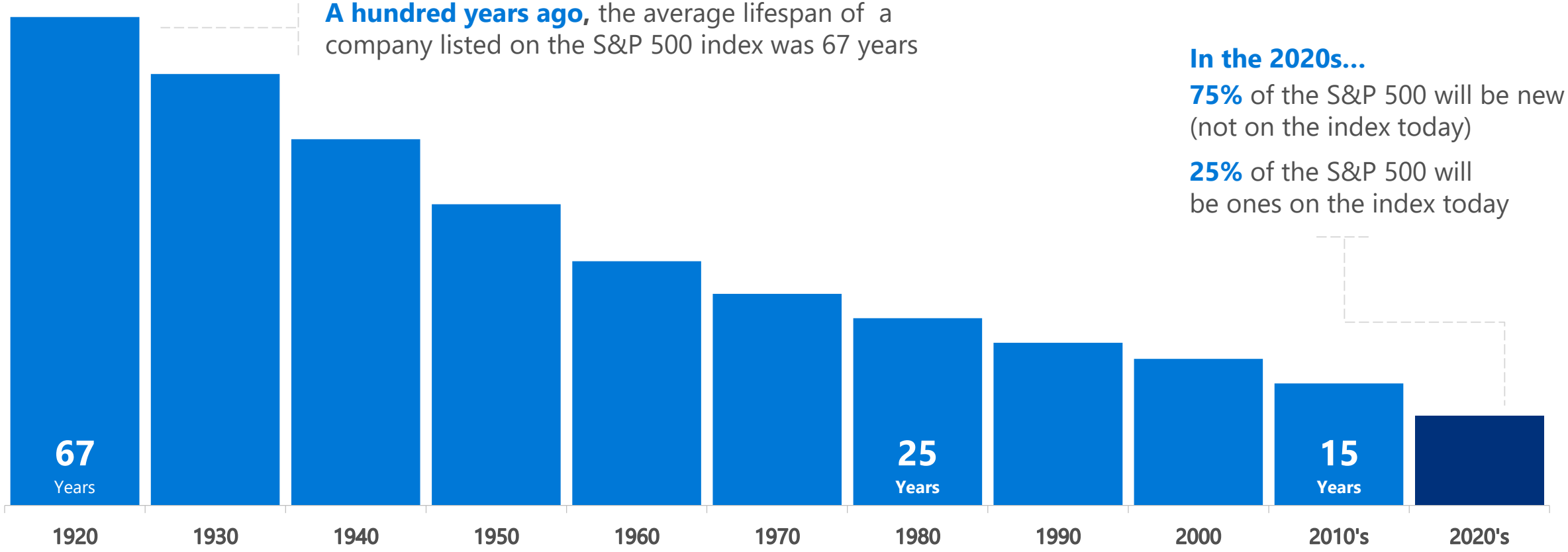


Speed

Scale

Economics

A cégek számára az alkalmazkodási idő lerövidült

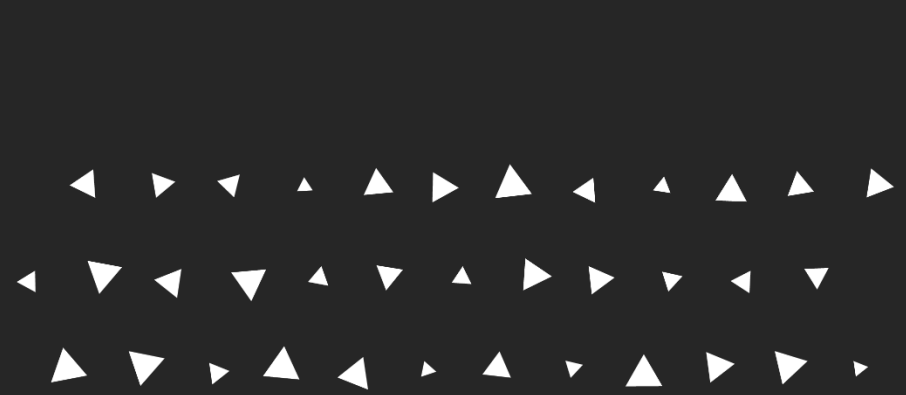


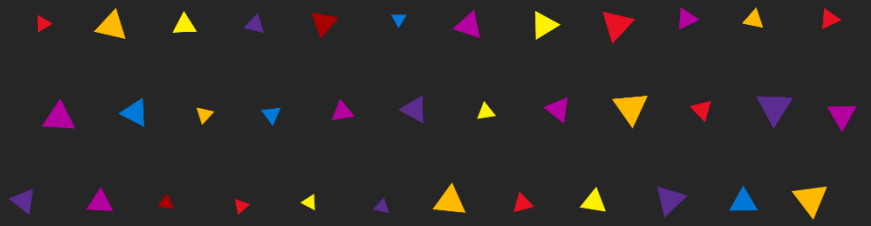
Source: BBC

Gondolatolvasás

"Connected brains" - Joseph Sirosh (Strata + Hadoop 2016)

https://www.youtube.com/watch?v=VLEpd52hr3Q



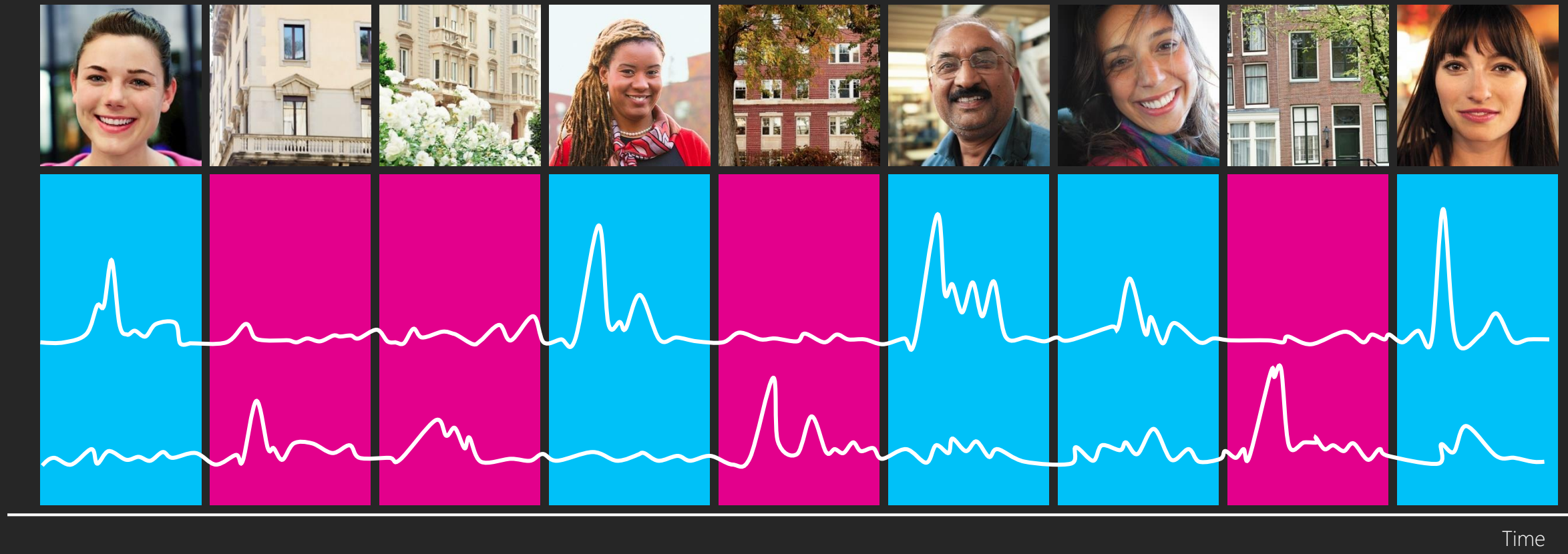


Képek

elektrokortikográfia (ECoG)

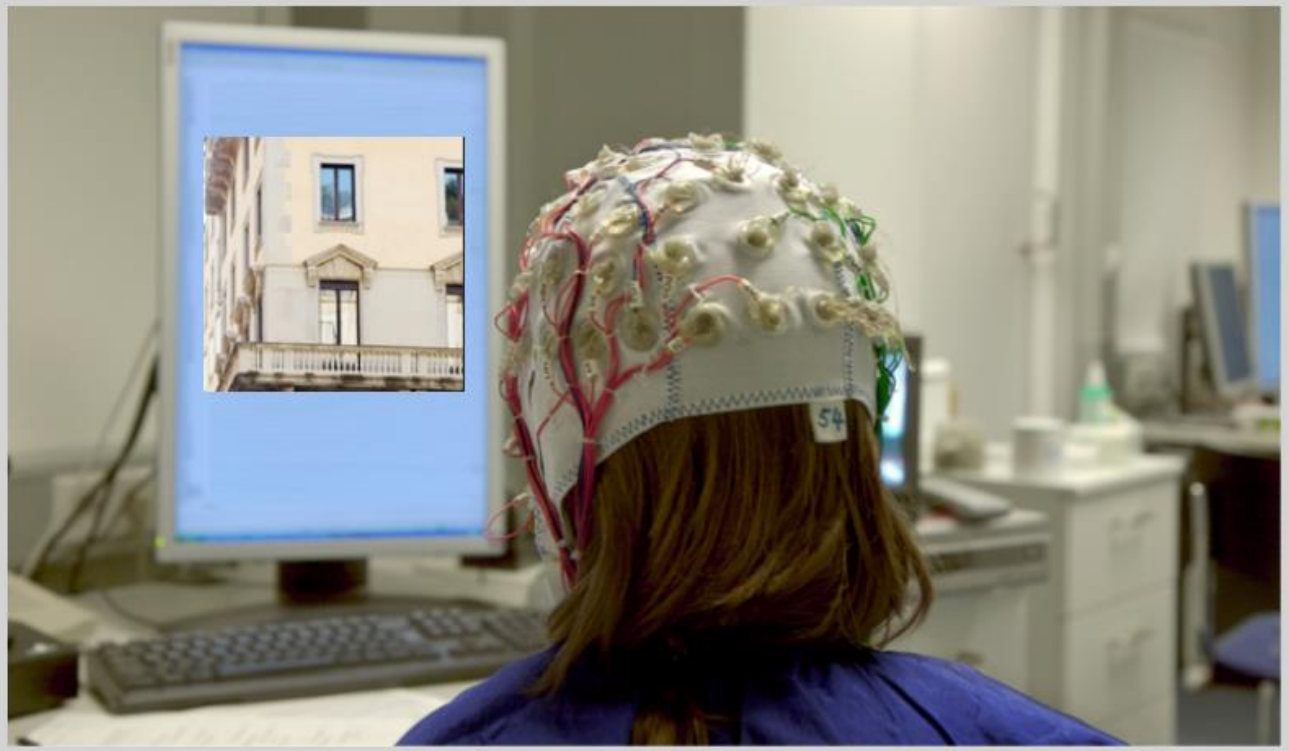
Agyhullámok


54


Time

Cortana Intelligence

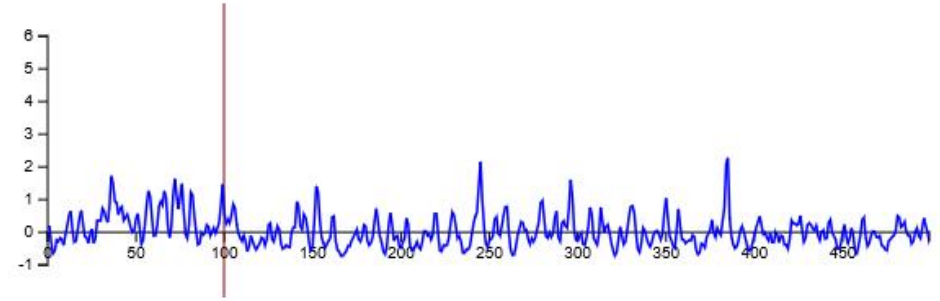
Patient CA



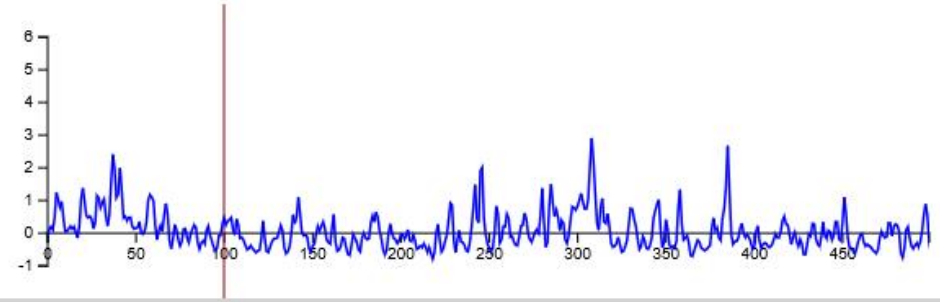
Decoding Brain Signal

House

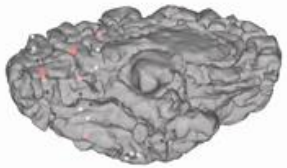
Select Electrodes 1



Select Electrodes 2



Brain Activity



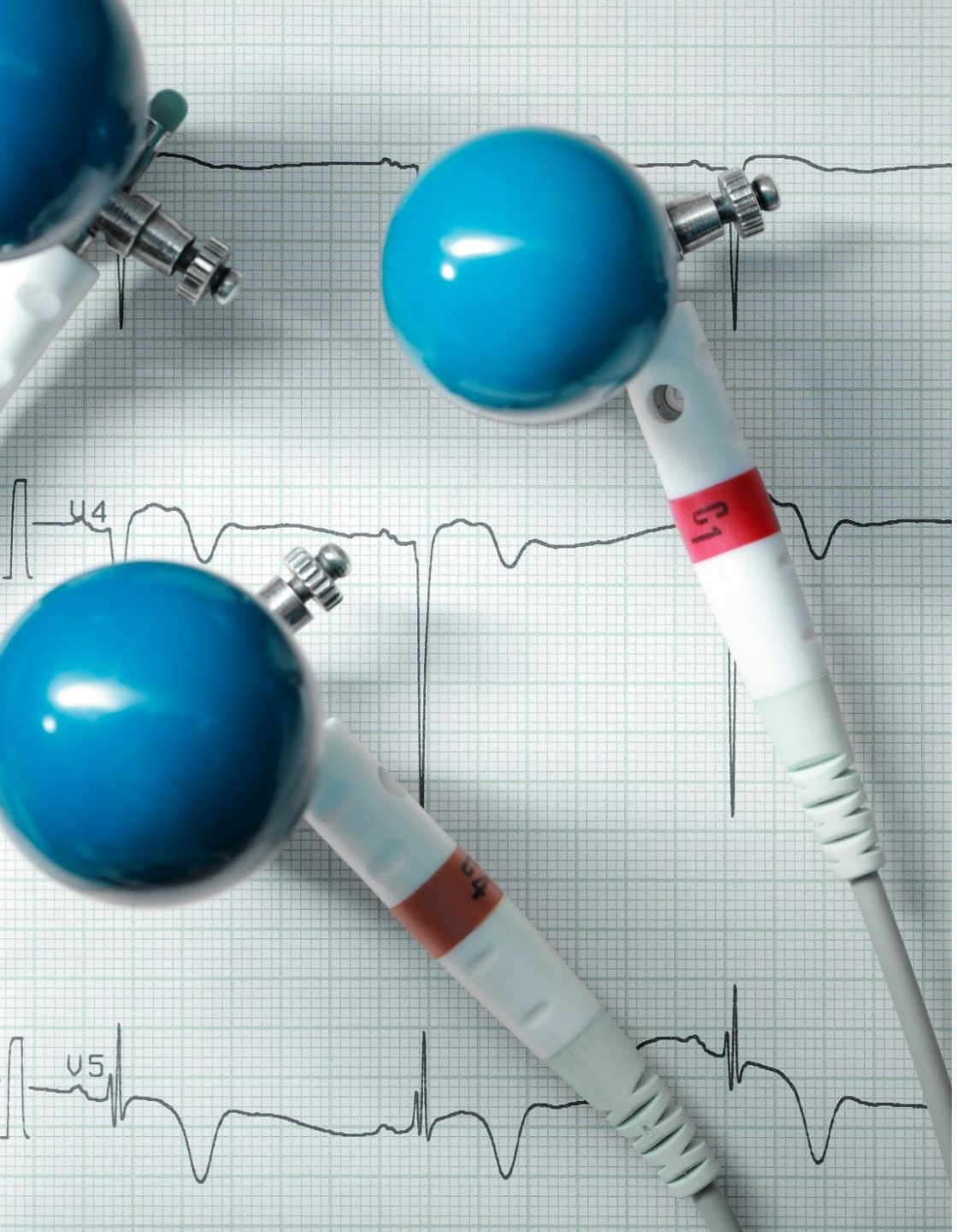
ECoG Signal

Brain Signal Voltage (Current)

0.831

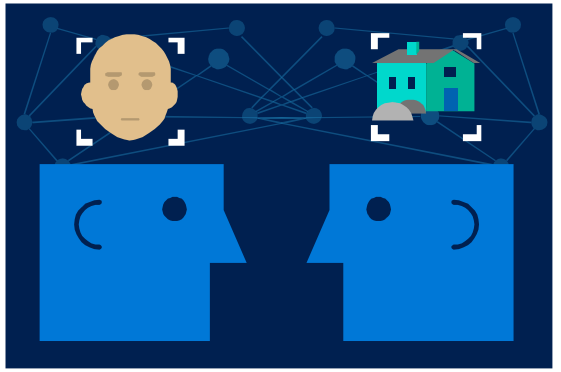
Duration

00:00:21



In-augural Cortana Intelligence Competition

Neural Signal Analysis



Timeline:
29 Mar – 28 Jun (3 Months)

Prizes Total:
$5,000

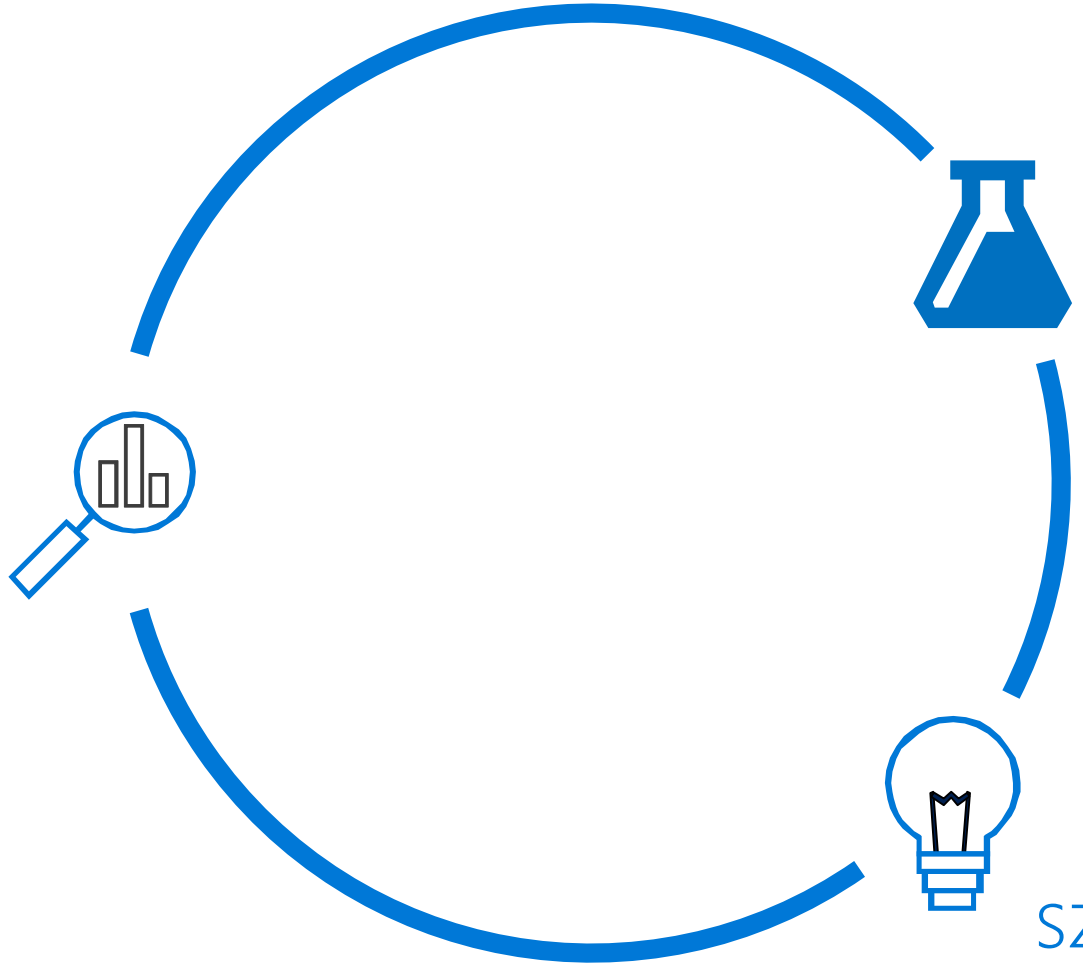
http://aka.ms/mindreader

- Analyze neural signals data collected from the sub-temporal cortical surface, when provided a visual stimuli
- Grand Prize : $3,000

Mi szükséges ilyen elemzésekhez?

Microsoft

Adatokból intelligens döntéshozatal



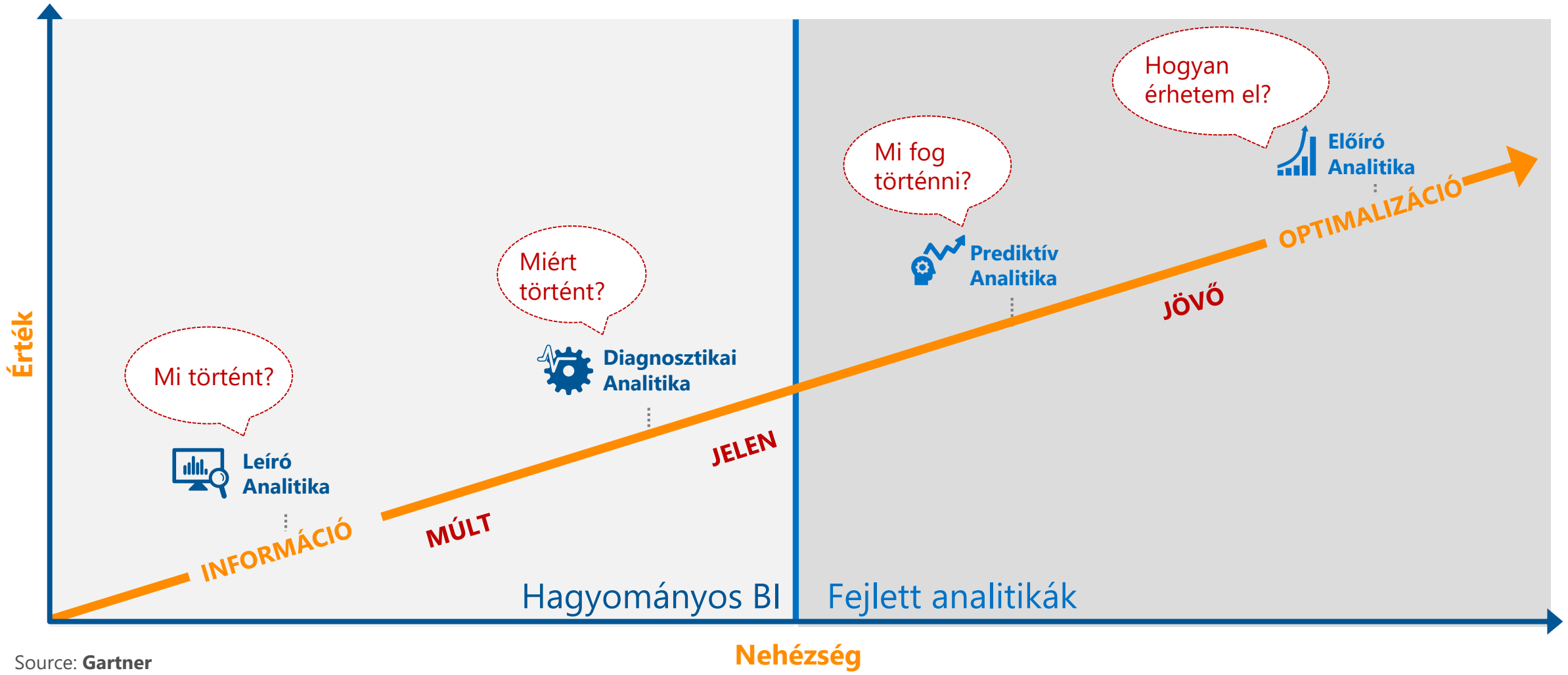
Adatok vizualizációja

Gépi Tanulás
Fejlett Analitika

Kognitív szolgáltatások

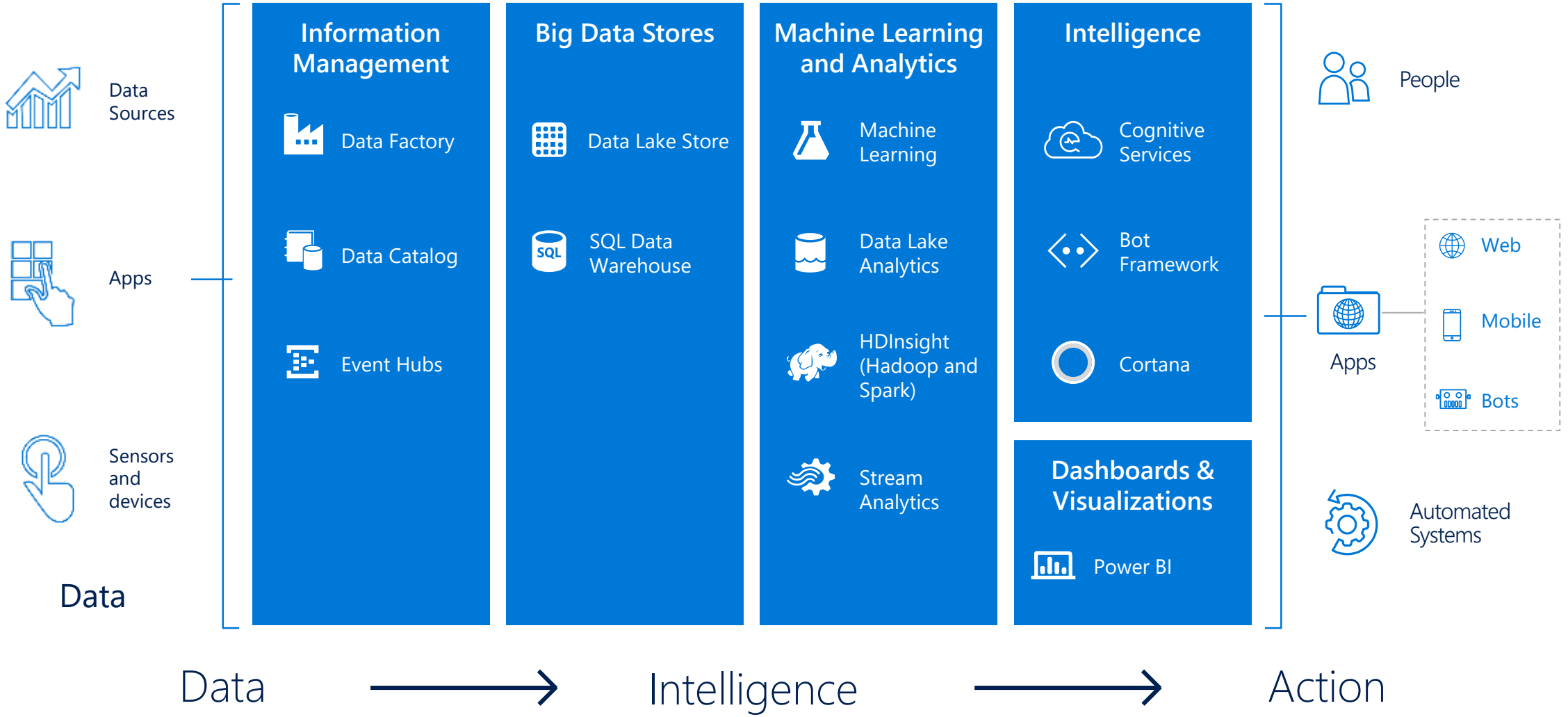
Fejlett analitika

Az üzleti intelligencián túl

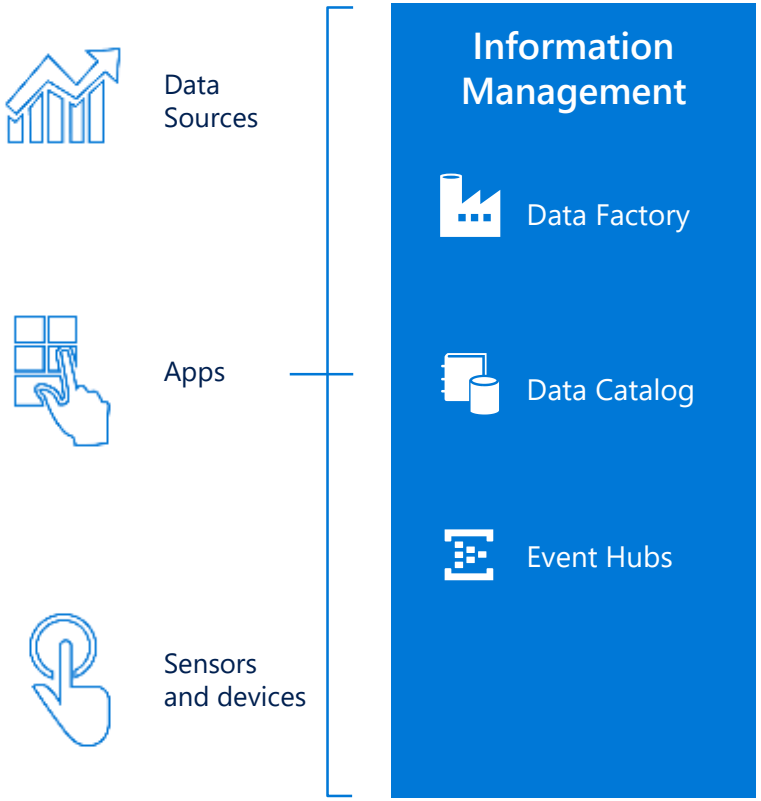


Source: **Gartner**

Adatalapú döntéshozatal a Cortana Intelligence segítségével



Information Management

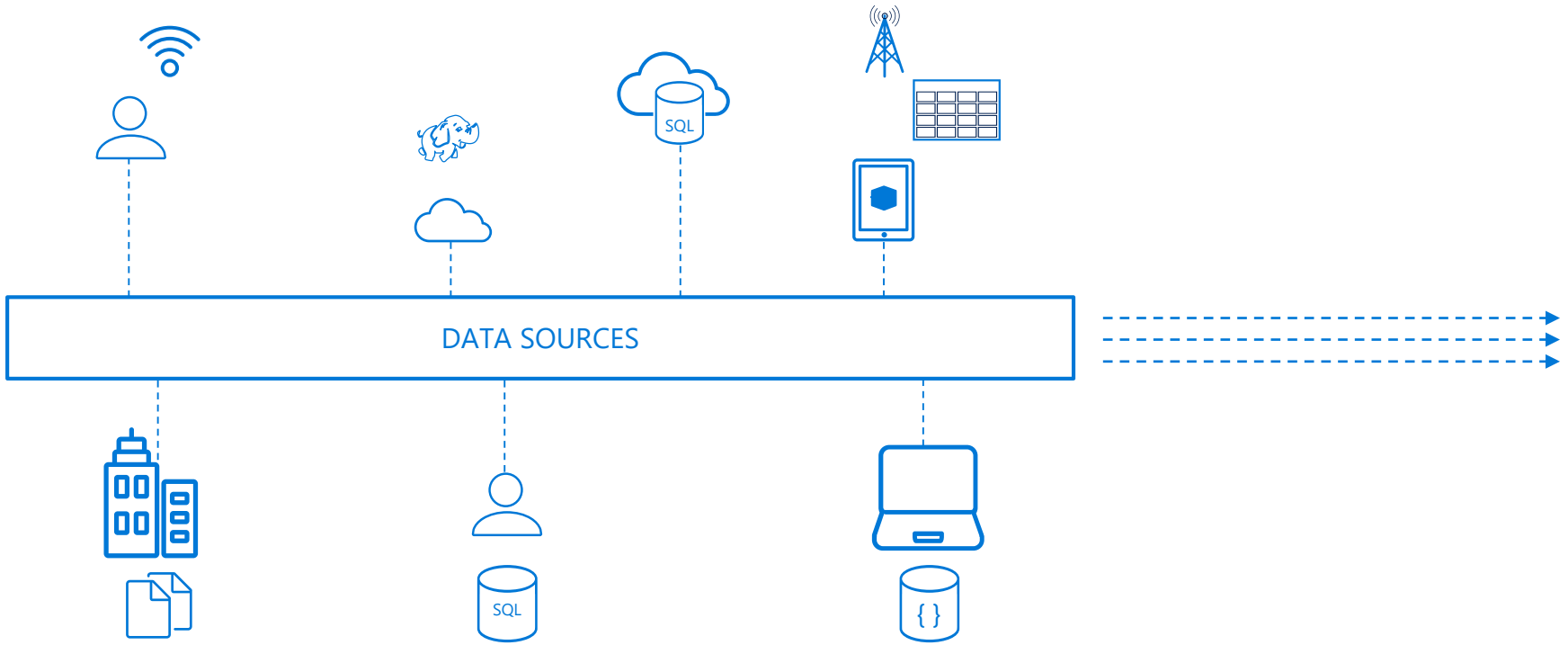


Data

Compose and orchestrate data services at scale

Information Management

- Data Factory
- Data Catalog
- Event Hubs

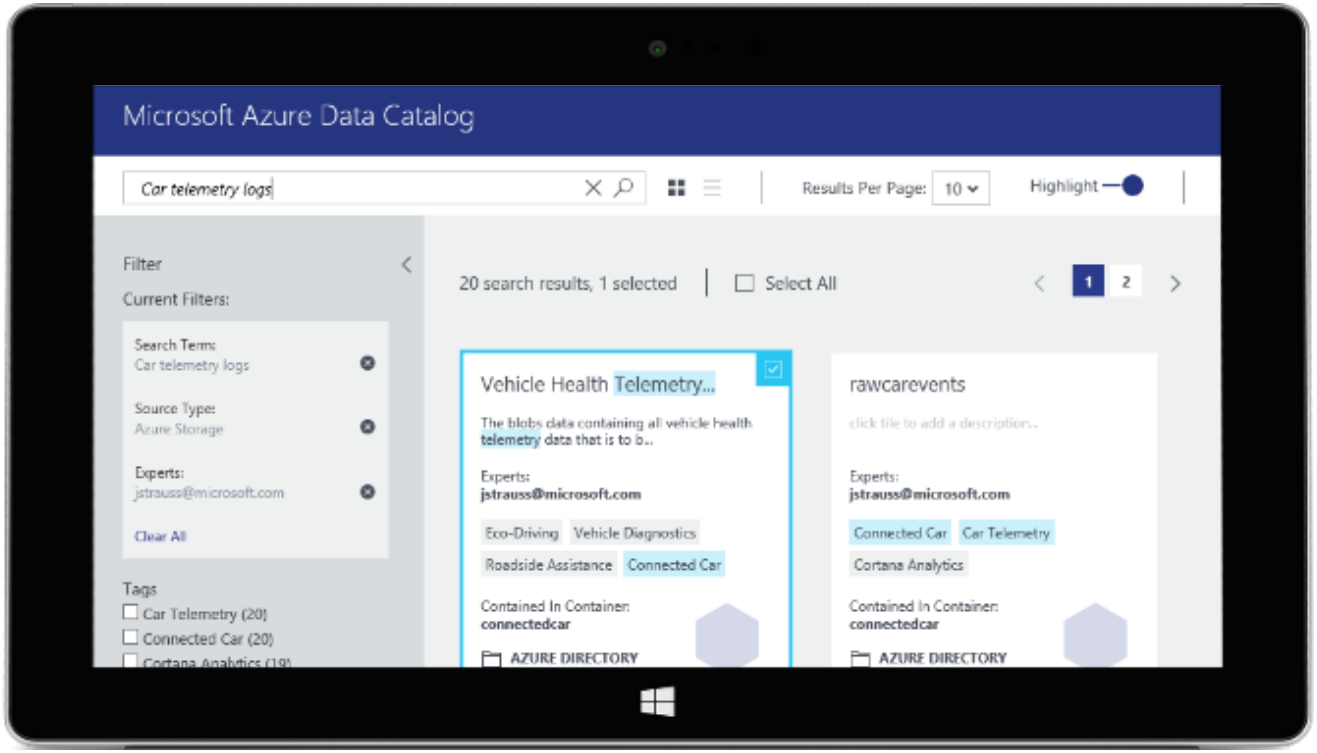


- Create, schedule, orchestrate, and manage data pipelines
- Visualize data lineage
- Connect to on-premises and cloud data sources
- Monitor data pipeline health
- Automate cloud resource management
- Move relational data for Hadoop processing
- Transform with Hive, Pig, or custom code

Get more value from your enterprise data assets

Information Management

- Data Factory
- Data Catalog
- Event Hubs

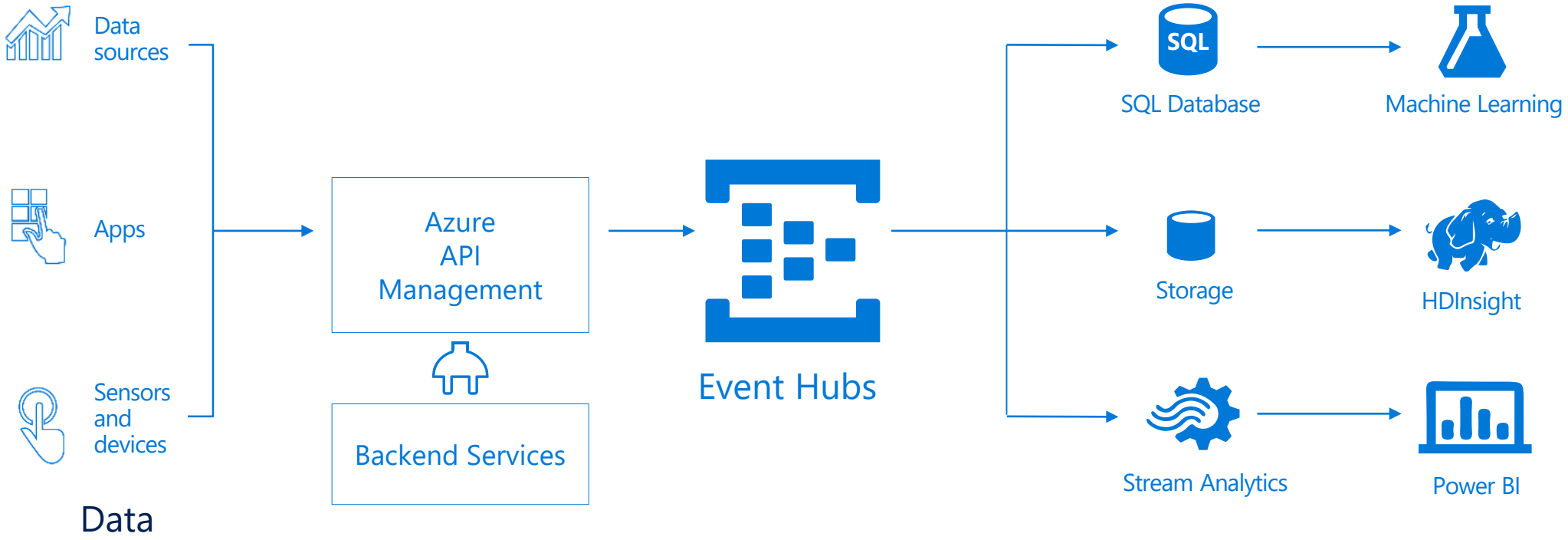


- Spend less time looking for data, and more time getting value from it
- Register enterprise data sources, discover data assets and unlock their potential, and capture tribal knowledge to make data understandable
- Bridge the gap between IT and the business, allowing everyone to contribute their insights, 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

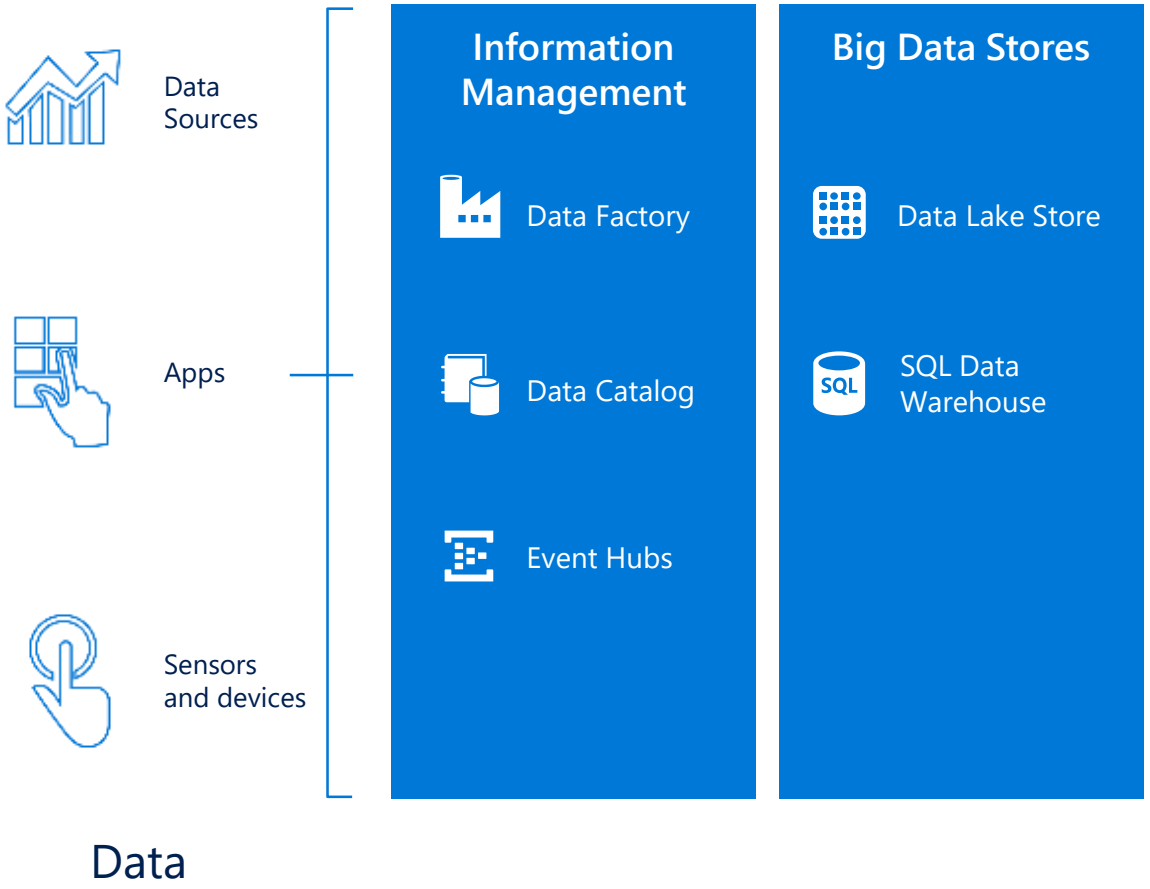
- Data Factory
- Data Catalog
- Event Hubs



- 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

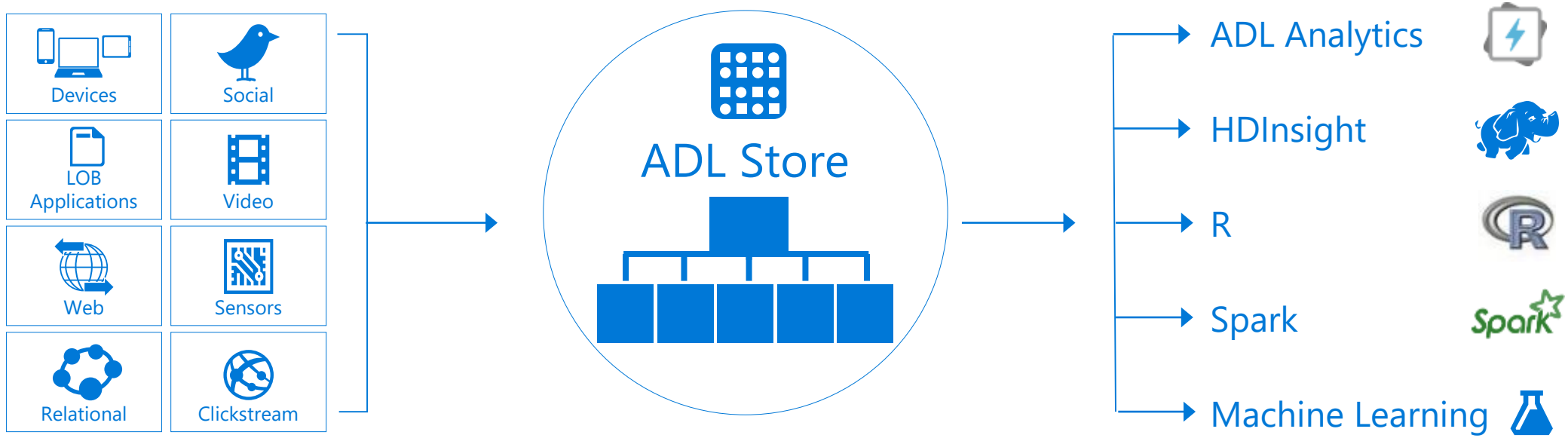


Data

A hyper-scale repository for big data analytics workloads

Big Data Stores

- Data Lake Store
- SQL Data Warehouse



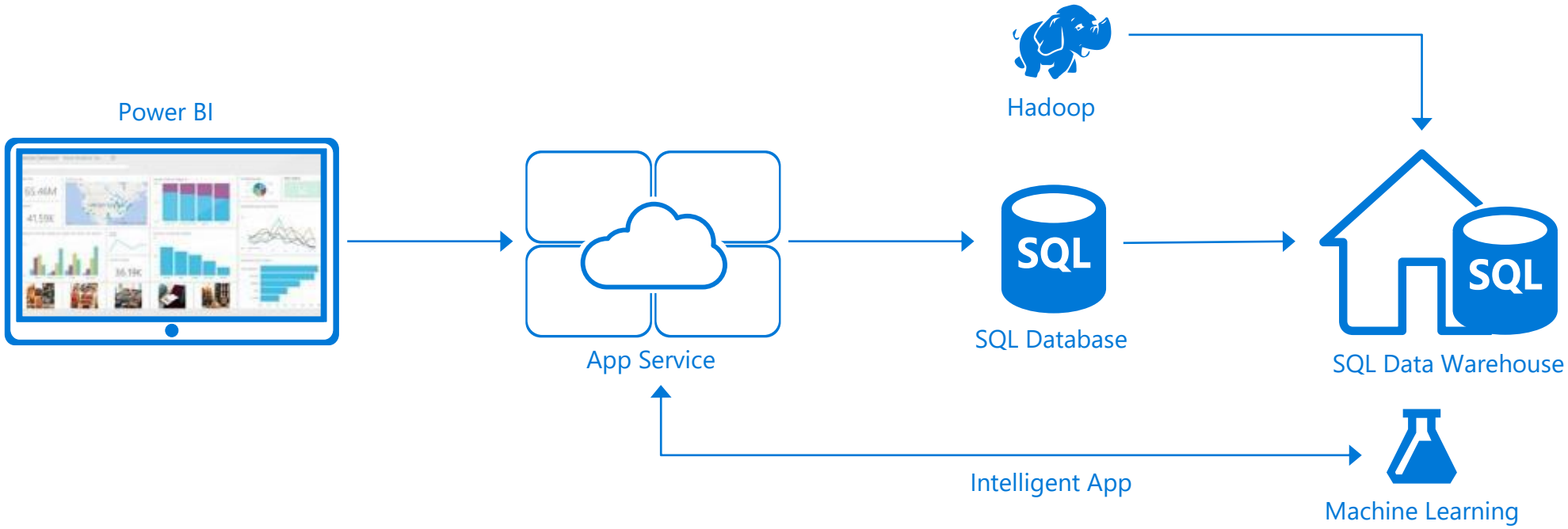
- A Hadoop Distributed File System for the cloud
- No fixed limits on file size
- No fixed limits on account size
- Unstructured and structured data in their native format
- Massive throughput to increase analytic performance
- High durability, availability, and reliability
- Azure Active Directory access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

Elastic data warehouse as a service with enterprise-class features

Big Data Stores

Data Lake Store

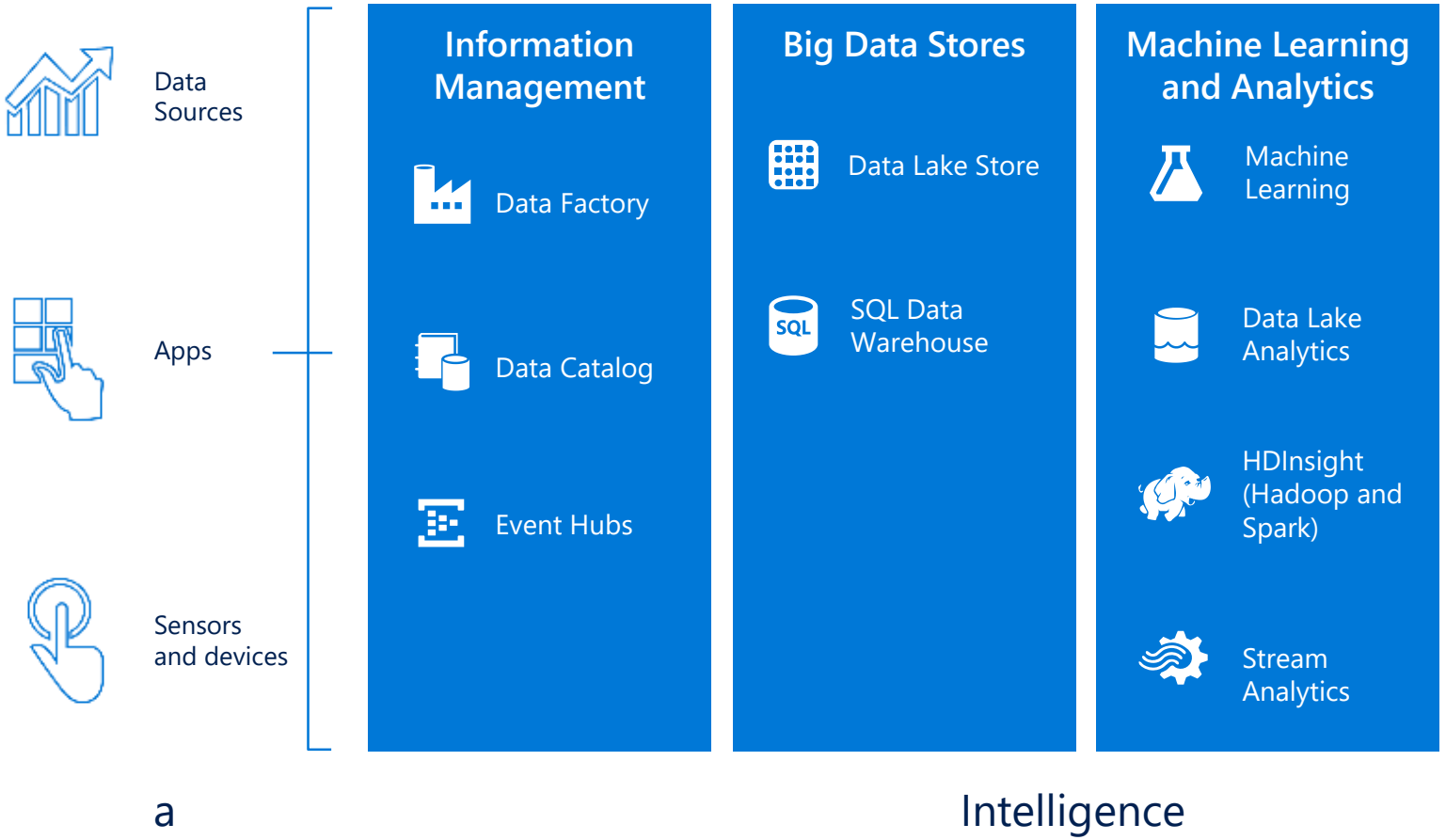
SQL Data Warehouse



- 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

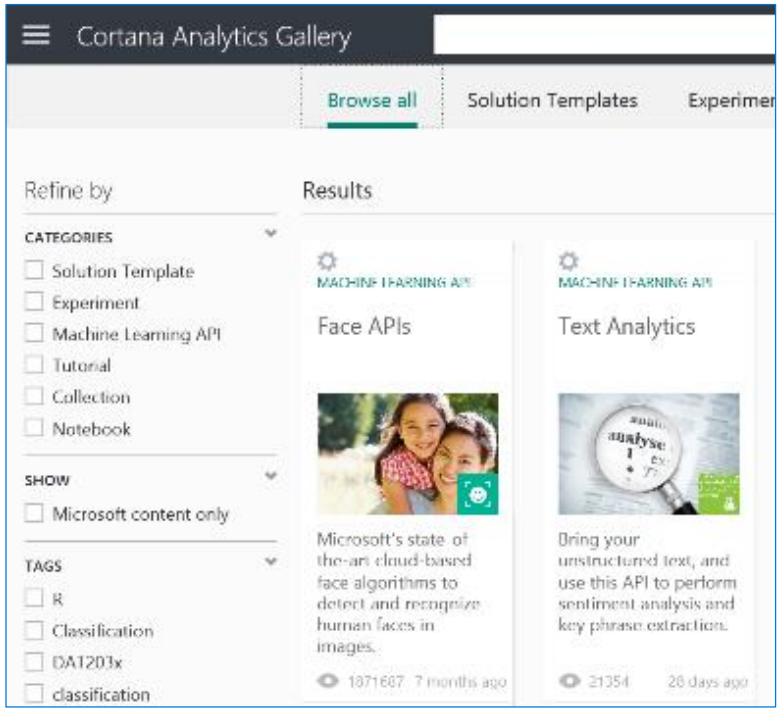
Machine Learning and Analytics

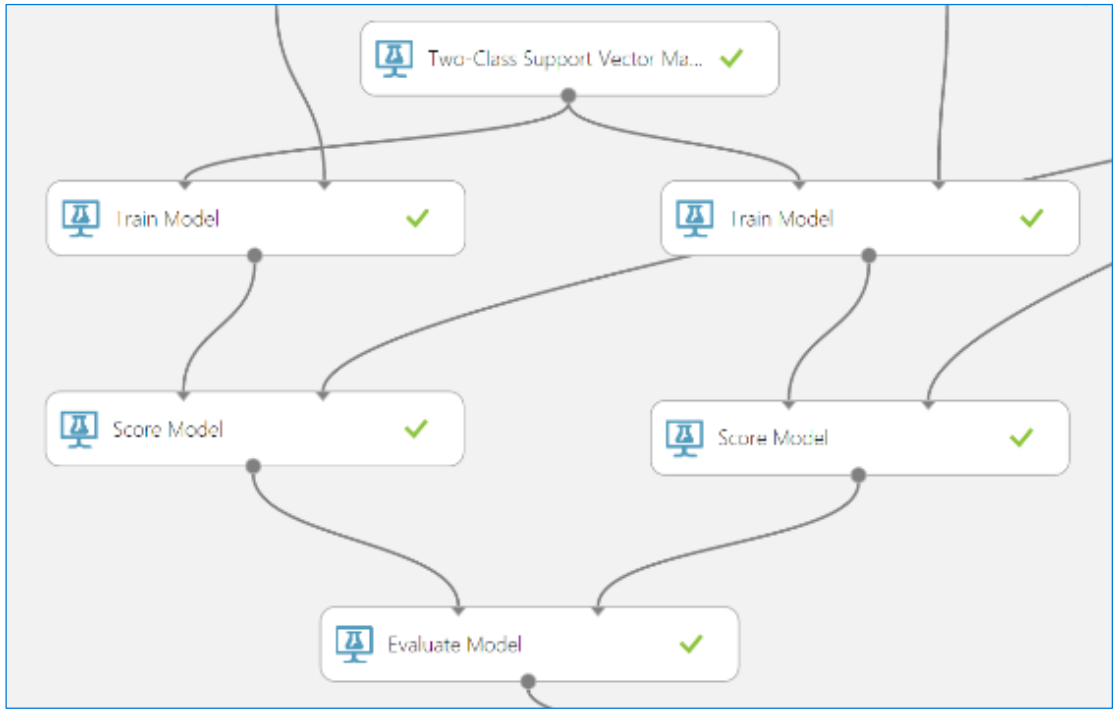


Easily build, deploy, and share predictive analytics solutions

Machine Learning and Analytics

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



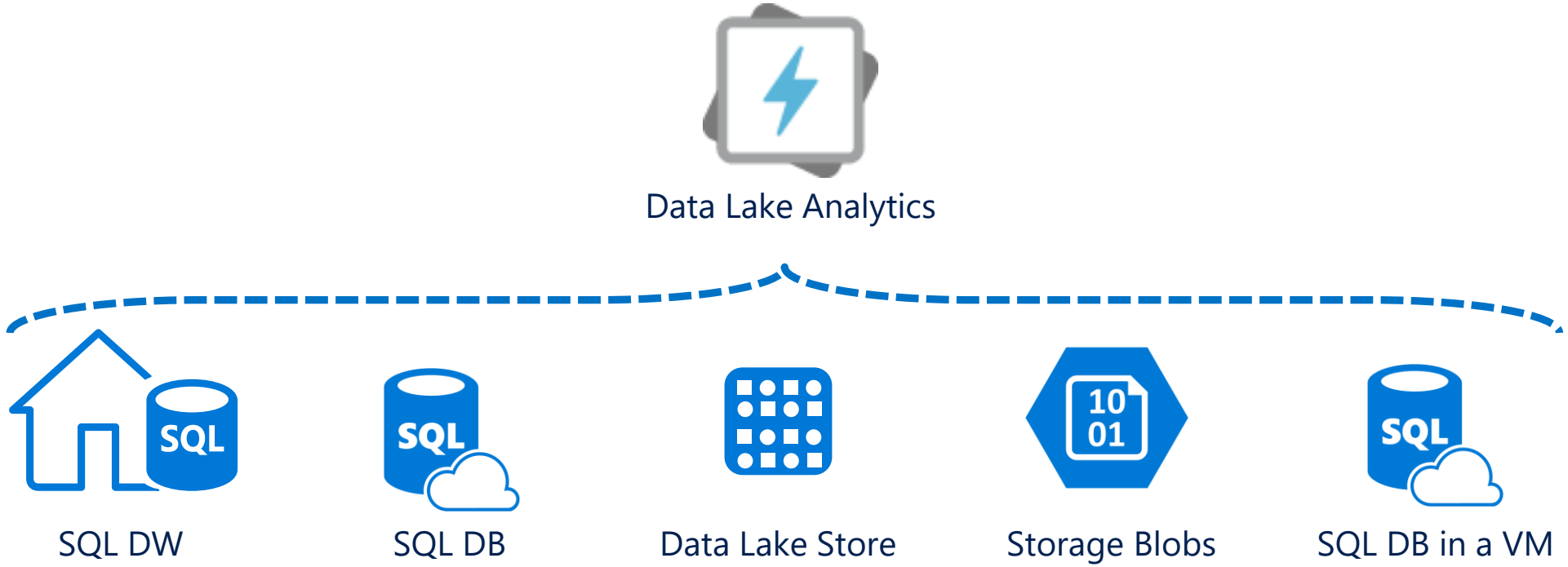


- Simple, scalable, cutting edge. A fully managed cloud service that enables you to easily build, deploy, and share predictive analytics solutions.
- Deploy in minutes. Azure Machine Learning means business. You can deploy your model into production as a web service that can be called from any device, 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

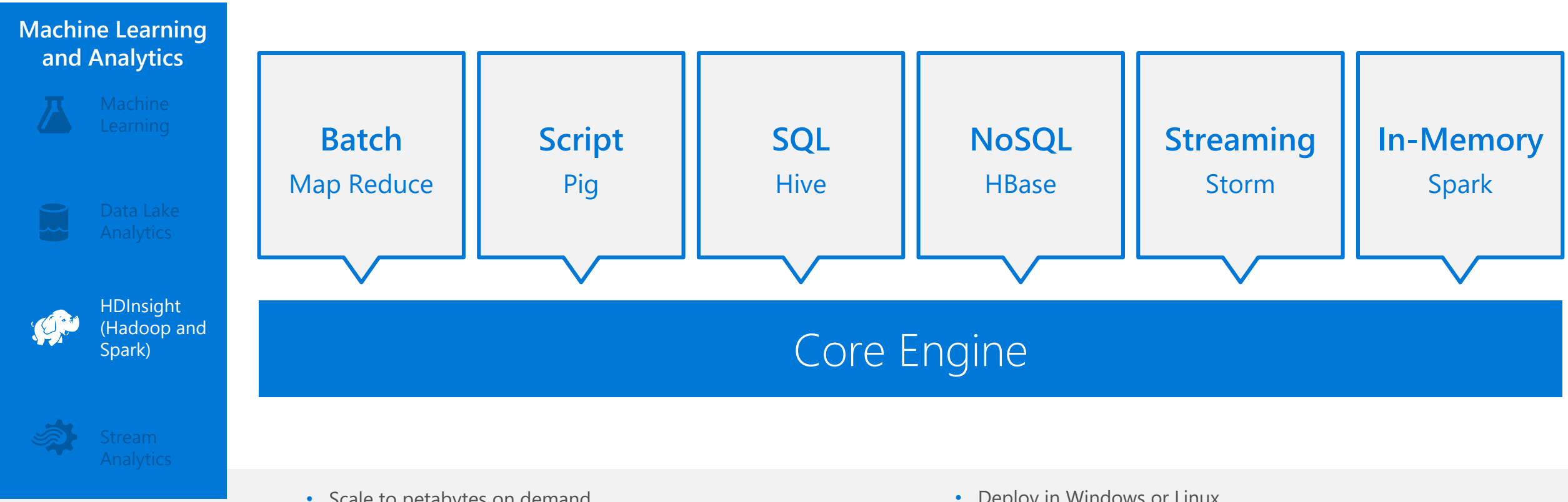
Machine Learning and Analytics

- 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



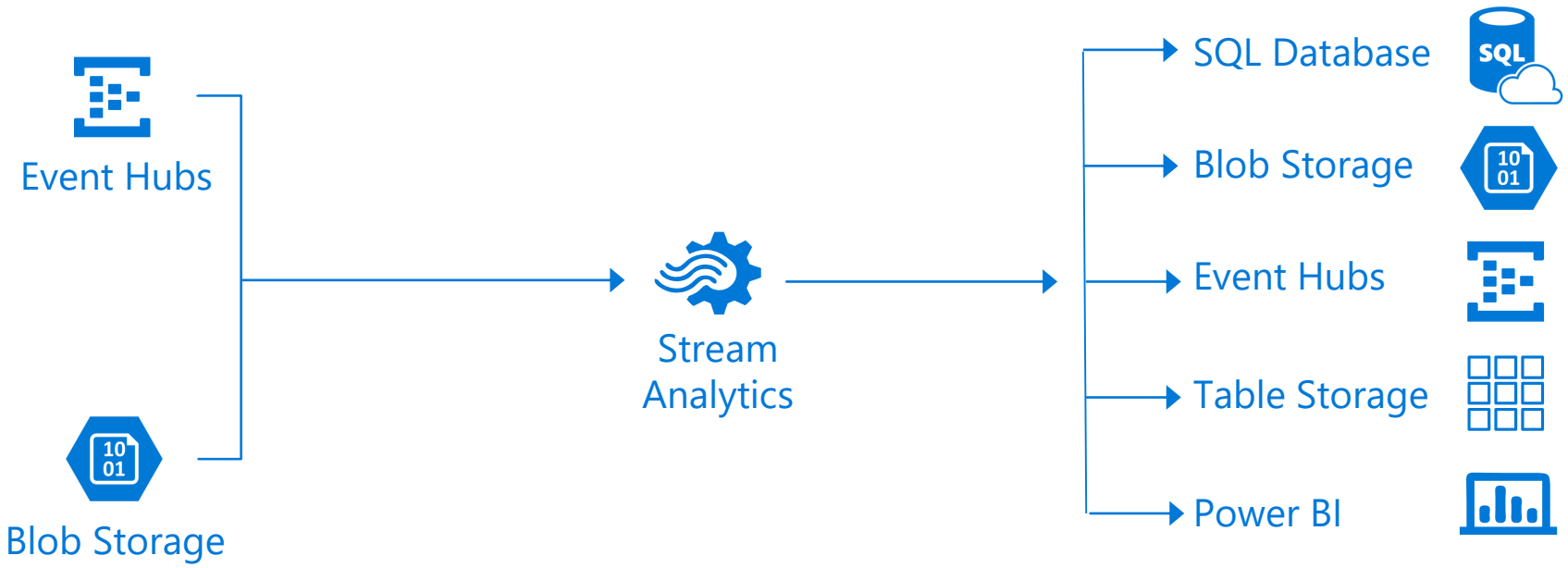
- Scale to petabytes on demand
- Process unstructured and semi-structured data
- Develop in Java, .NET, and more
- Skip buying and maintaining hardware

- 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

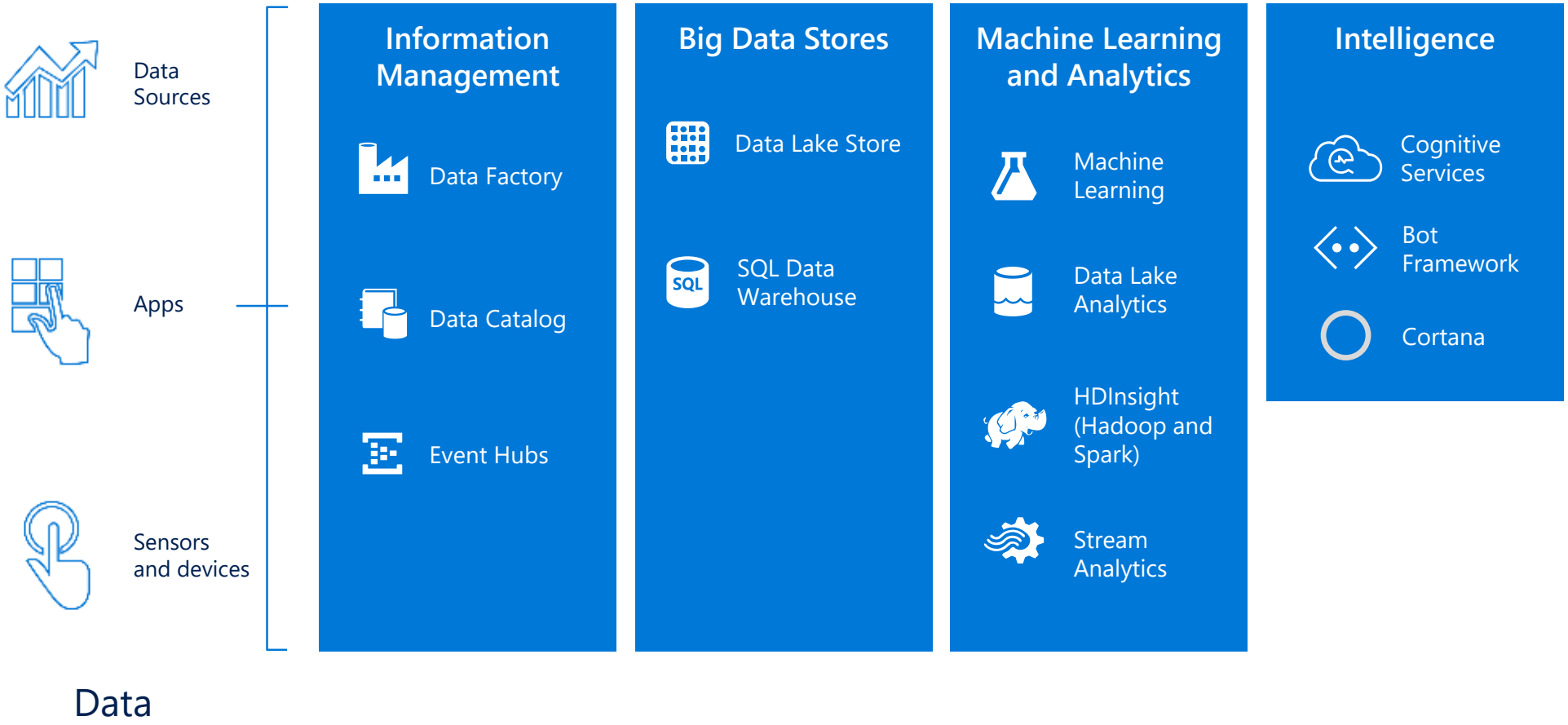
Machine Learning and Analytics

- 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

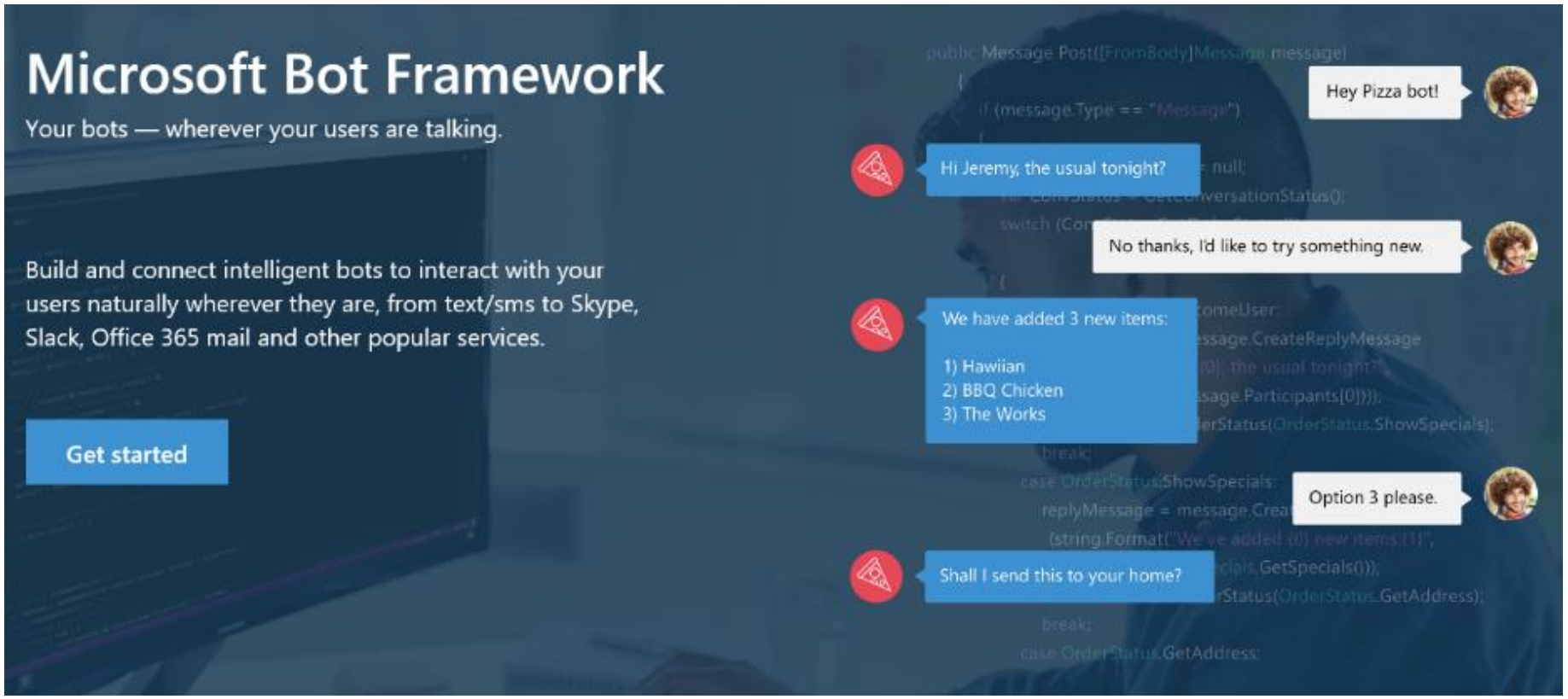
- Faces, images, emotion recognition and video intelligence
- Spoken language processing, speaker recognition, custom speech recognition
- Natural language processing, sentiment and topics analysis, spelling errors

- 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

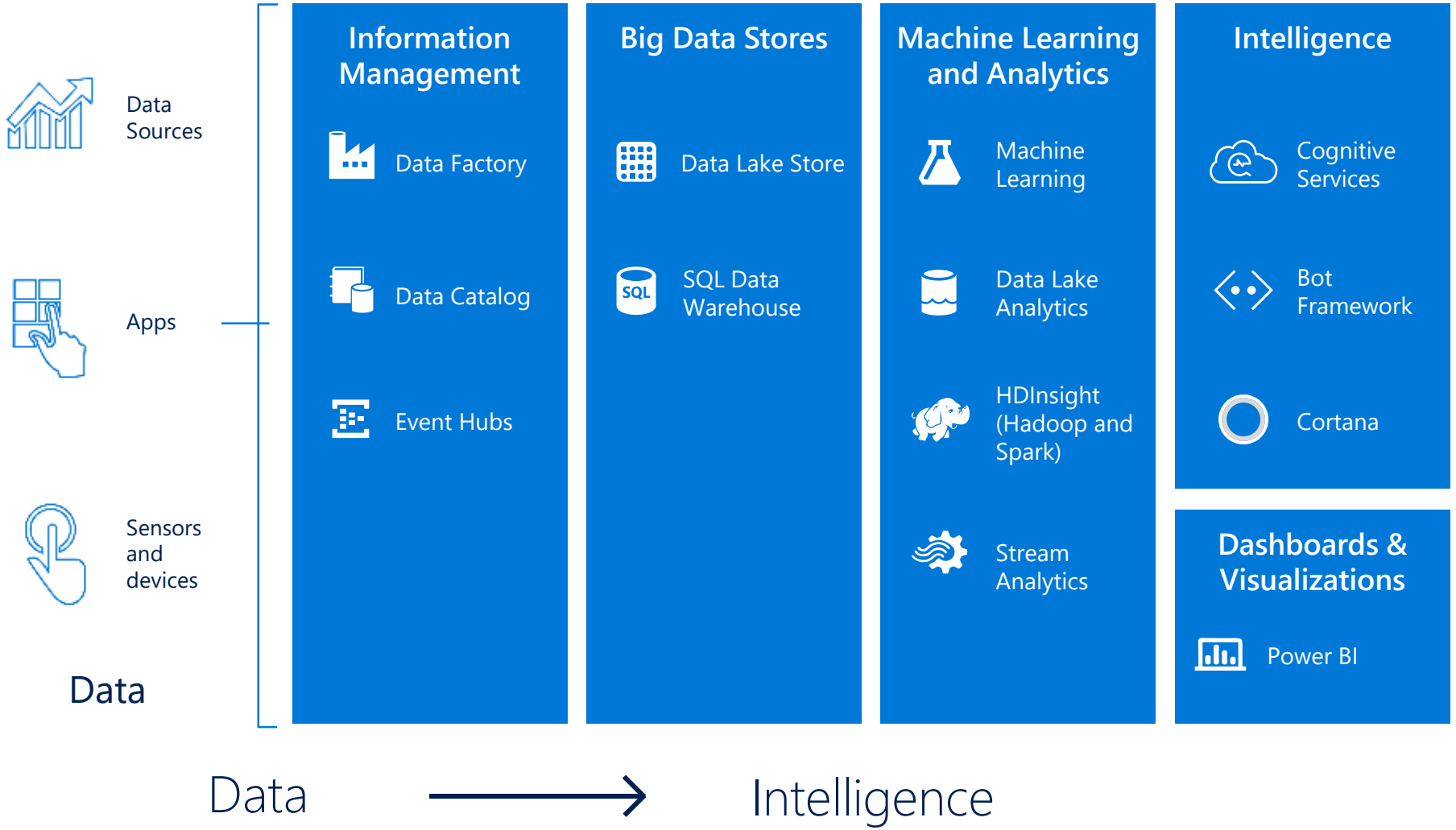

Cognitive Services


Bot Framework

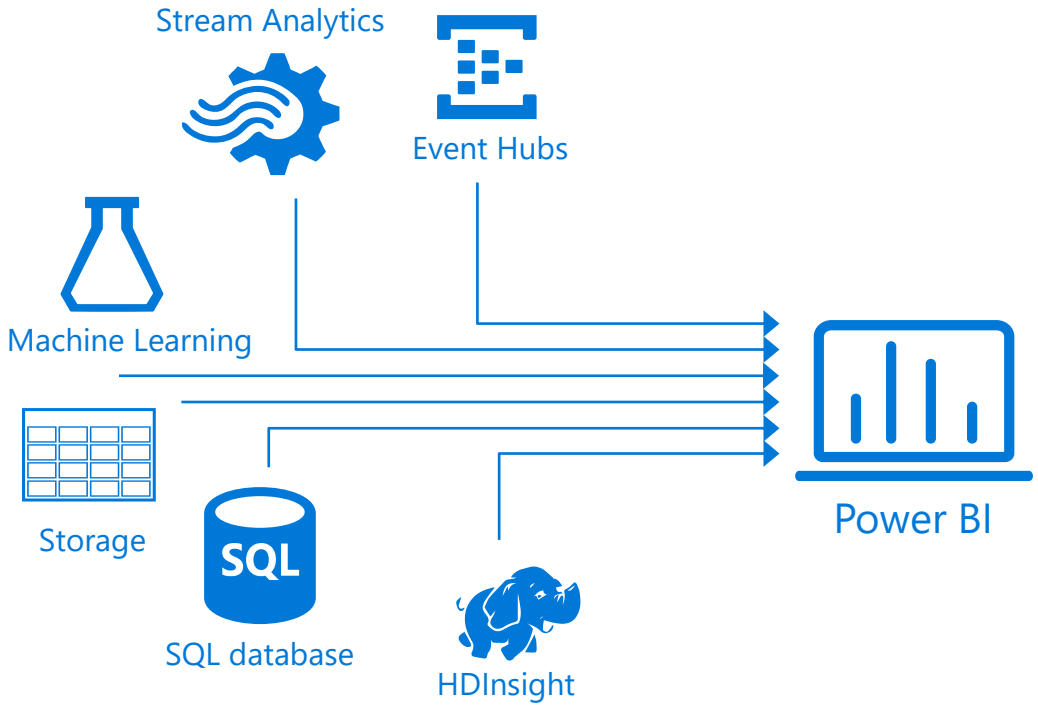

Cortana

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

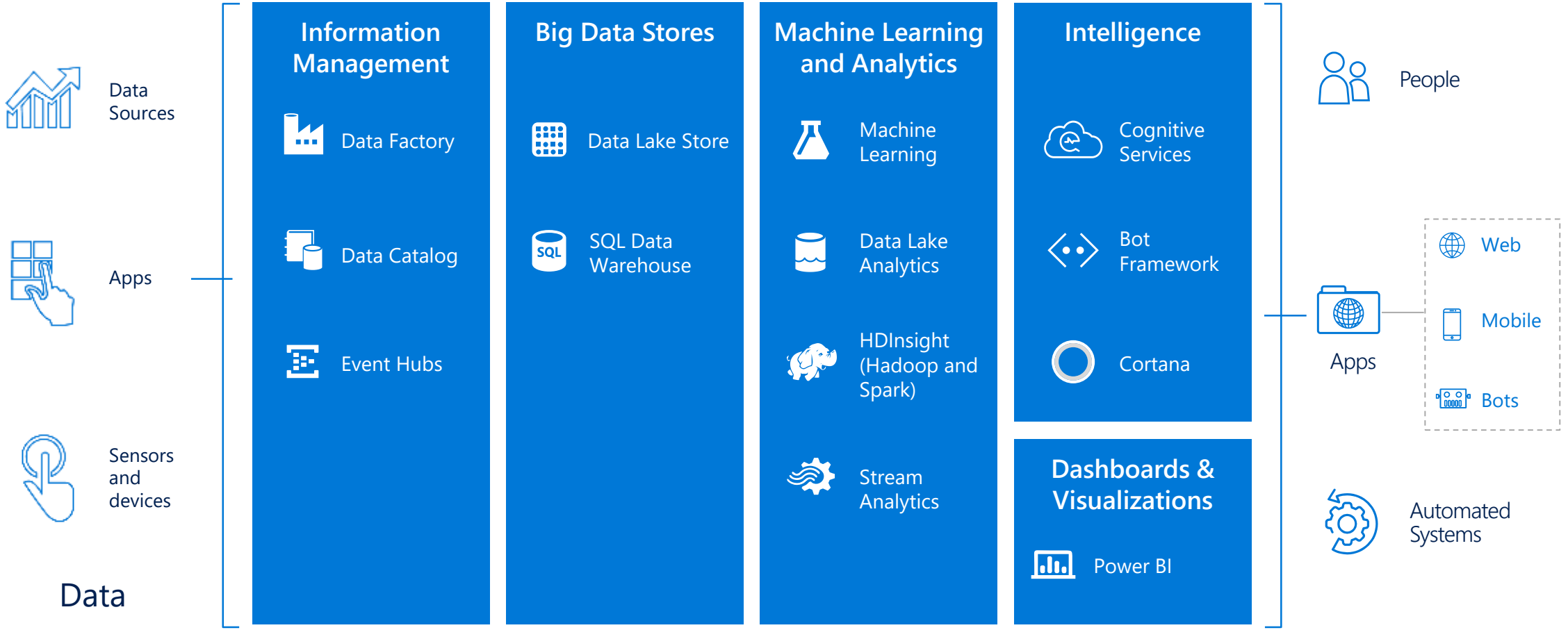


Dashboards & Visualizations

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



Video

Using Spark to Power the Office 365 Delve Organization Analytics

https://youtu.be/gVZtSukah0U

Microsoft Azure Machine Learning

Gépi tanulás a felhőben

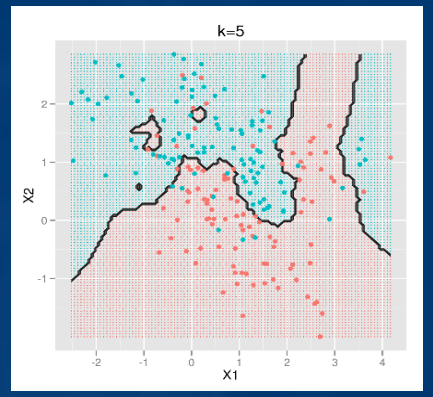
Video

Microsoft's Vision For Azure Machine Learning

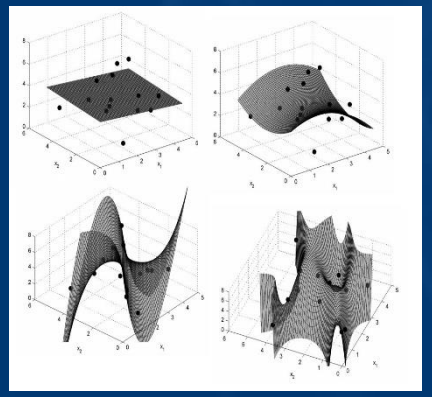
https://youtu.be/SJtNJepz-pM

Gyakori Gépi tanulási feladatok

Osztályozás

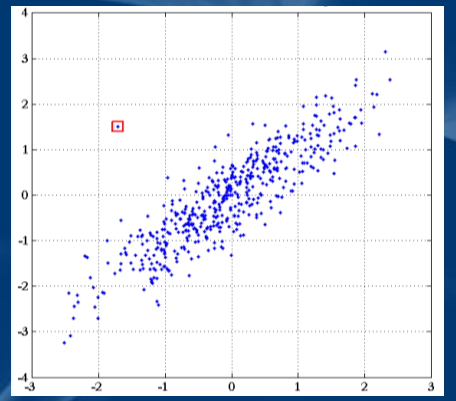


Regresszió

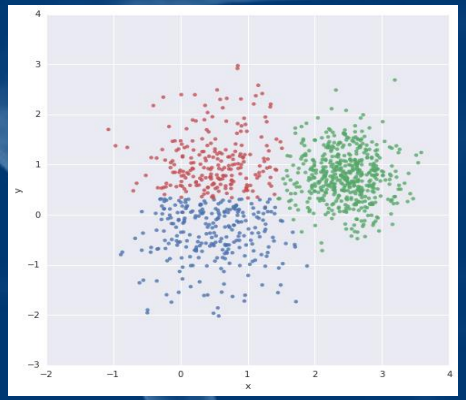


Ajánló motorok

Anomália felismerés



Szegmentálás, klaszterezés



DEMO
Prediktív Elemzés

Cél:

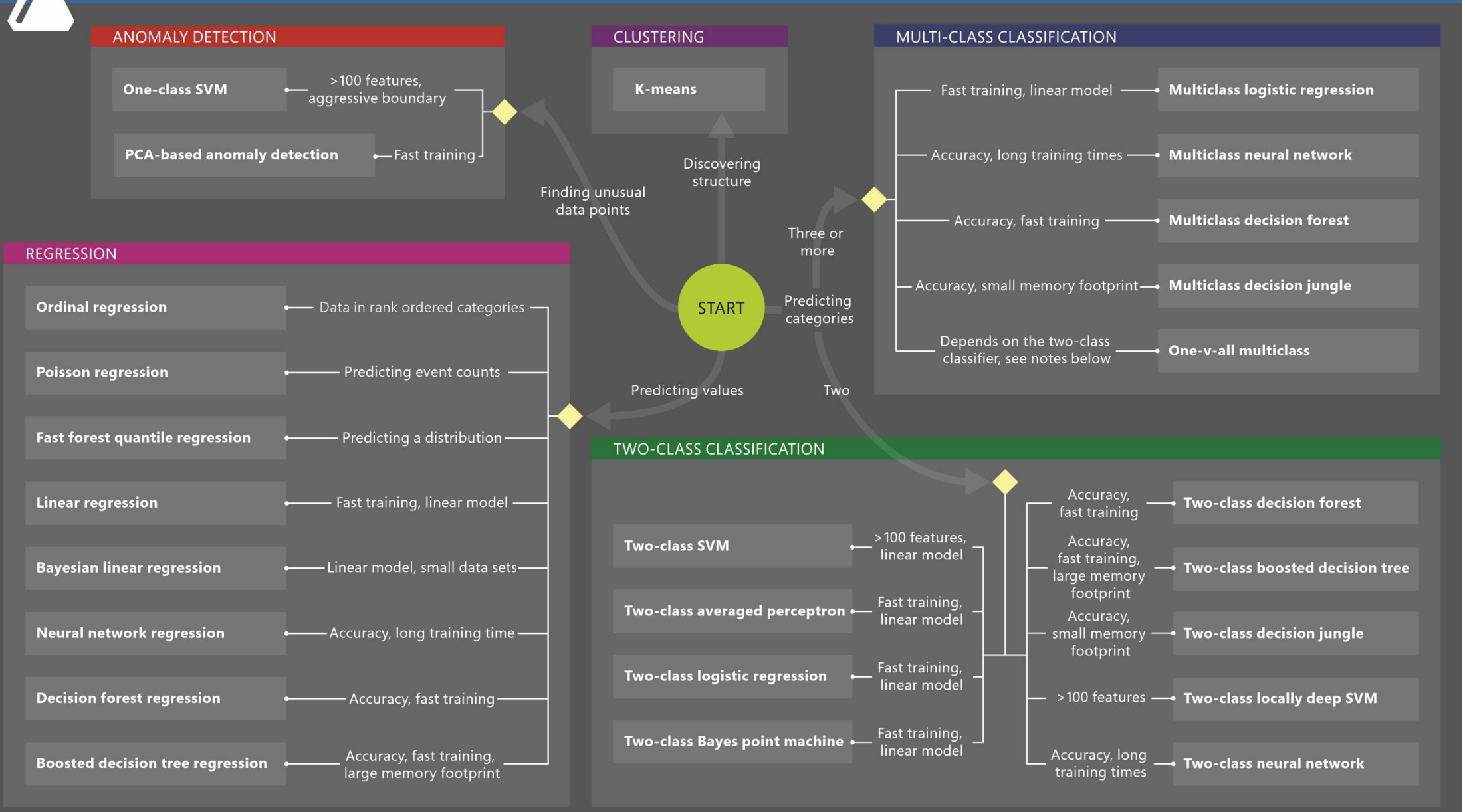
- Jövedelem becslés (>50K vagy<=50K) kor, nem, iskolai végzettség, családi állapot (felügyelt, bináris klasszifikáció)

Adatforrás:

- 1994-es amerikai népszámlálás
- http://archive.ics.uci.edu/ml/machine-learning-databases/adult/adult.names

Microsoft Azure Machine Learning: Algorithm Cheat Sheet

This cheat sheet helps you choose the best Azure Machine Learning Studio algorithm for your predictive analytics solution. Your decision is driven by both the nature of your data and the question you're trying to answer.



Created by the Azure Machine Learning Team

Email: AzurePoster@microsoft.com

Download this poster: http://aka.ms/MLCheatSheet

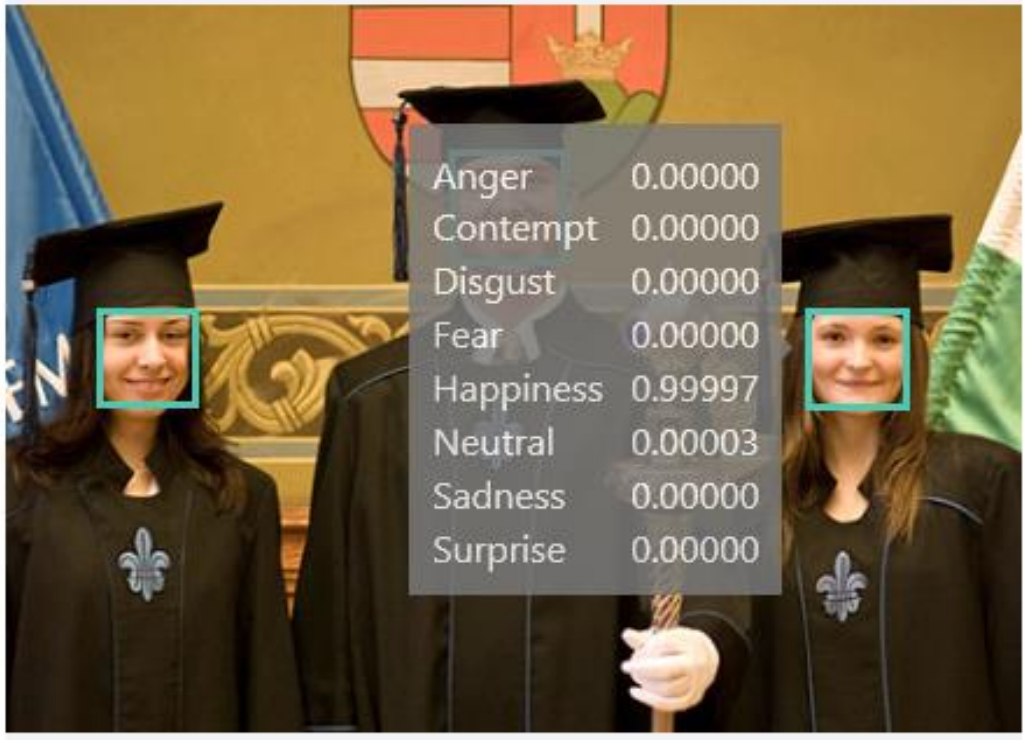
Microsoft

Microsoft Cognitive Ser

microsoft.com/cognitive-services/en-us/emotion-api

Home APIs Applications Developers Pricing My account

You can also click the open image button or drag-and-drop to upload your own images, or input a URL for a remote image. We don't keep your images for this demo.




```
Detection Result:
3 faces detected

JSON:
[
  {
    "faceRectangle": {
      "left": 201,
      "top": 65,
      "width": 53,
      "height": 53
    },
    "scores": {
      "anger": 5.096765e-8,
      "contempt": 1.27204885e-7,
      "disgust": 0.0000462597673,
      "fear": 8.077383e-13,
      "happiness": 0.999947965,
      "neutral": 0.0000055845685,
```

Video

Microsoft Cognitive Services: Introducing the Seeing AI app

https://youtu.be/R2mC-NUAmMk