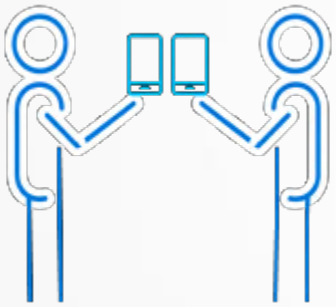
TiD 2017
质量竞争力大会

Microsoft 大数据平台和AI的应用

大数据推动变革

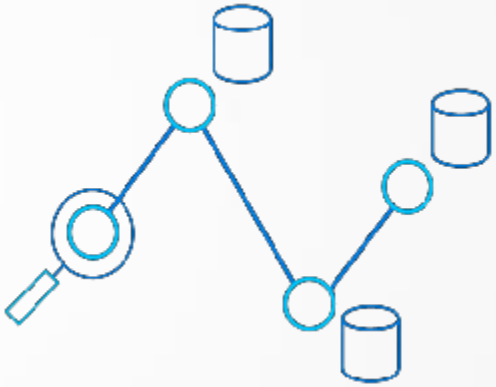
	Traditional	Big Data
Data characteristics	Relational data with highly modeled schema	All data with schema agility
Costs	Specialized HW	Commodity HW
Culture	Operational reporting Focus on rear-view analysis	Experimentation leading to intelligent action With machine learning, graph, a/b testing

大数据面临的挑战



Obtaining skills and capabilities

Determining how to get value



Integrating with existing IT investments

*Gartner Survey Analysis– Hadoop Adoption Drivers and Challenges (Stamford, CT: Gartner, 2015)

微软的大数据历程

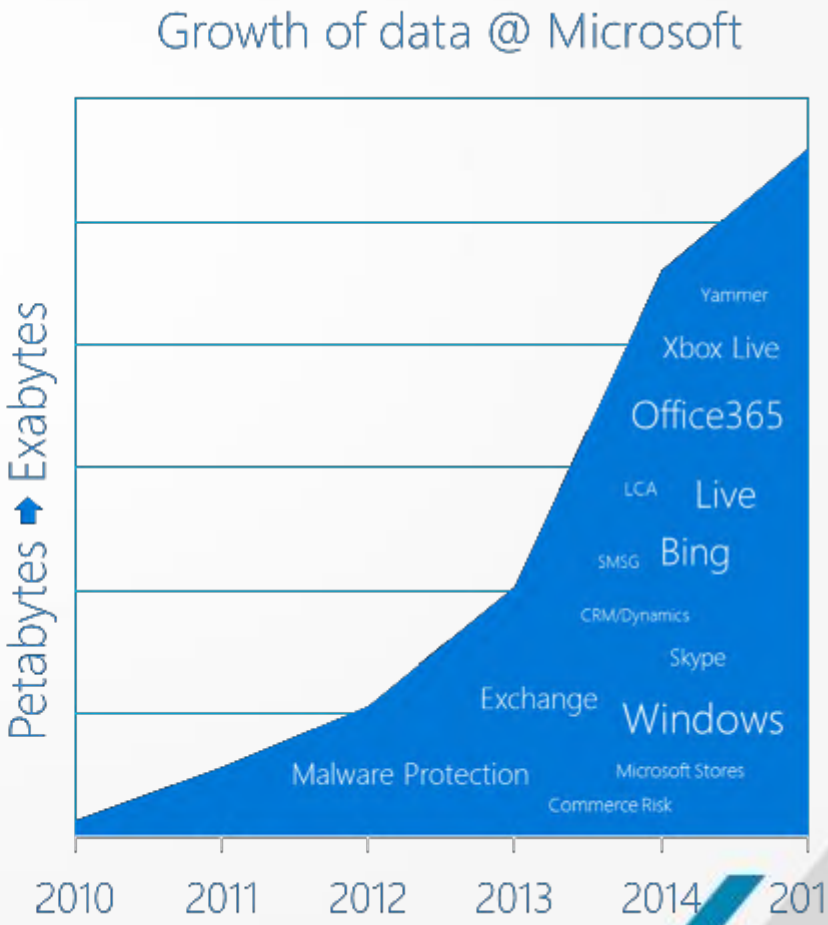
We needed to better leverage data and analytics to do more experimentation

So we:

- Designed a data lake for everyone to put their data into
- Built tools approachable by any developer
- Created machine learning tools for collaborating across large experiment models

Result:

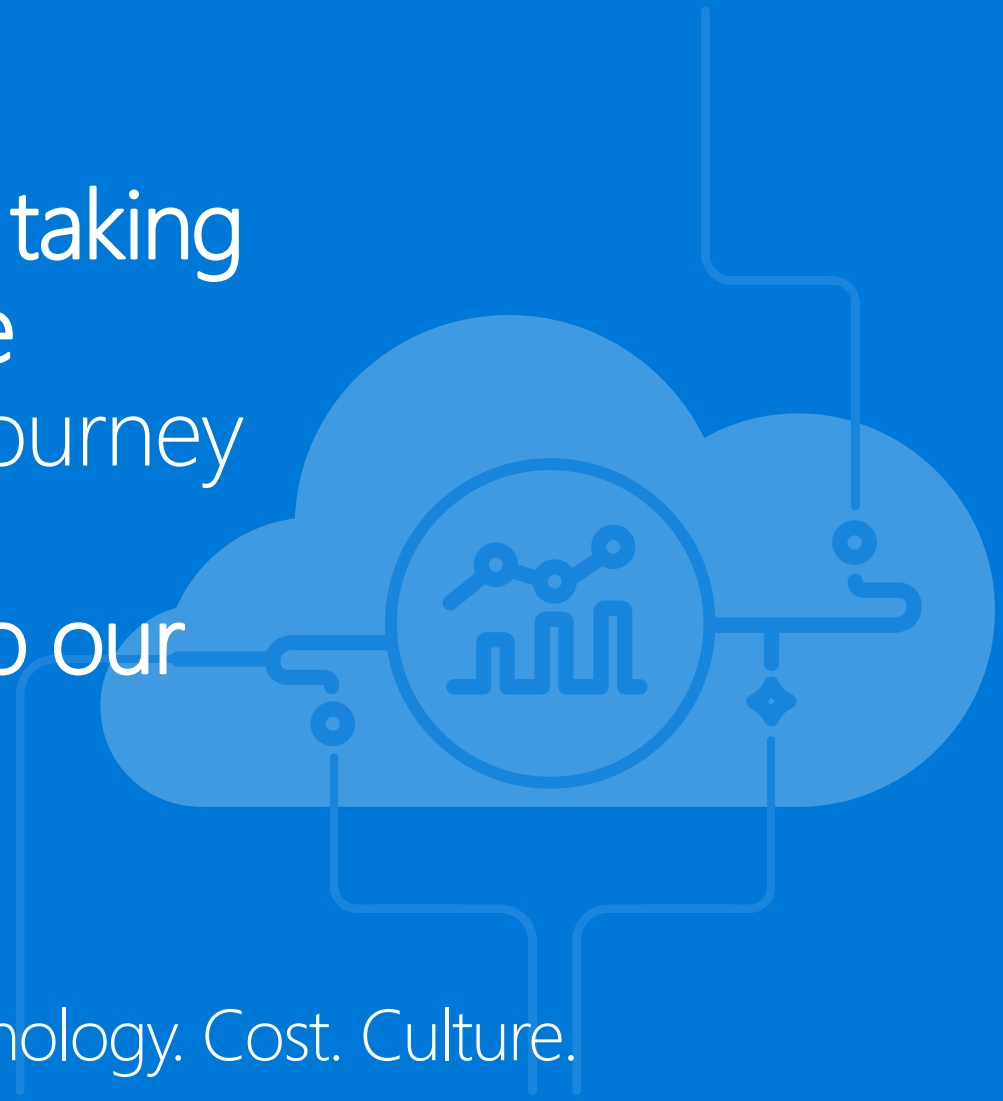
- Across Microsoft, ten thousand developers doing experimentation leading to better insights
- Leading to growth in our Microsoft businesses:
  - Office productivity revenue (45%YoY)*
  - Intelligent Cloud (100% YoY)*
  - Bing search share doubles



* Microsoft FY16 Q4 Results. URL: http://www.microsoft.com/en-us/Investor/earnings/FY-2016-Q4/press-release-webcast

Microsoft is now **taking everything we've learned** on this journey

and **bringing it to our customers**



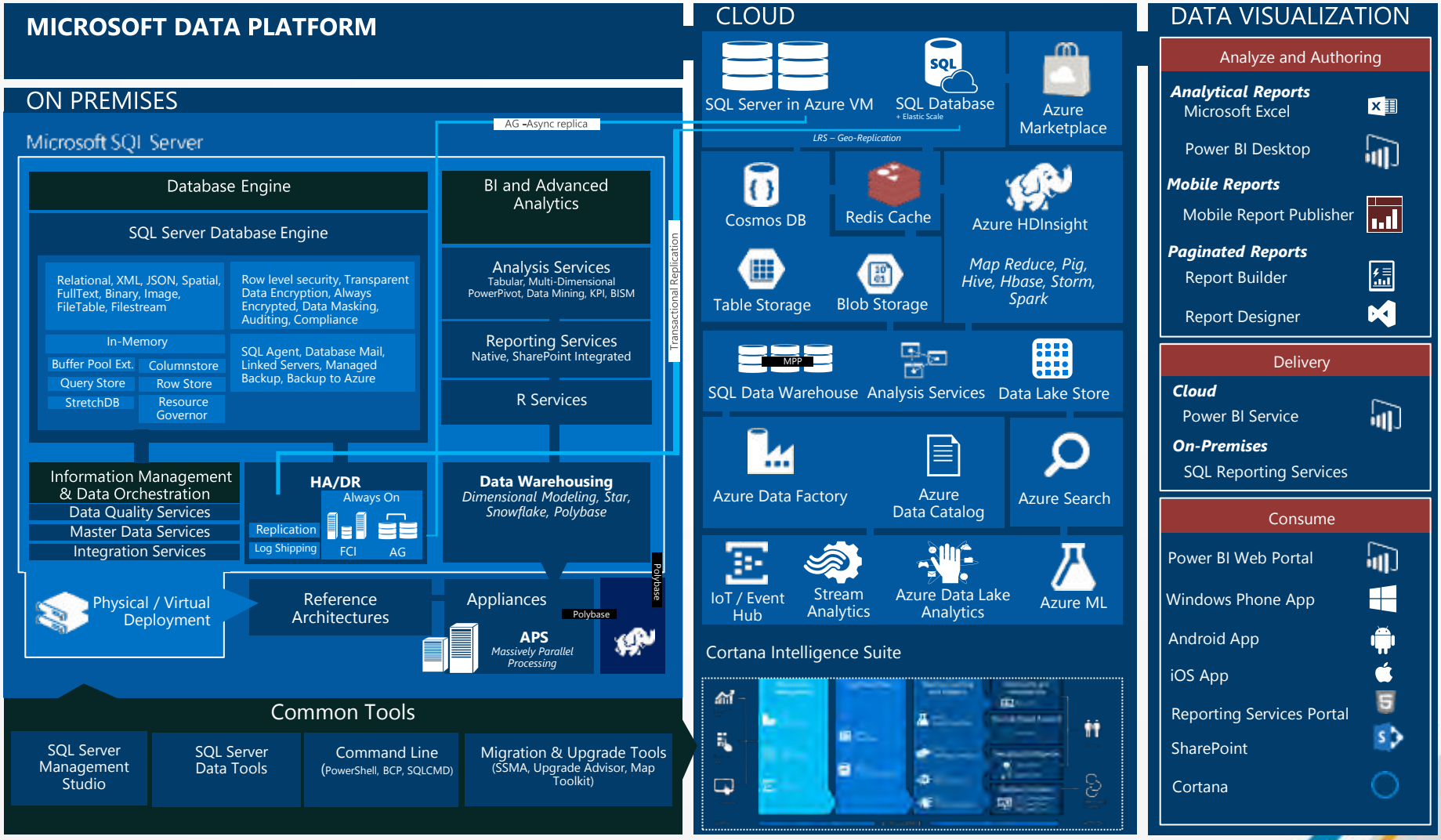
Technology. Cost. Culture.

TiD2017

微软大数据解决方案

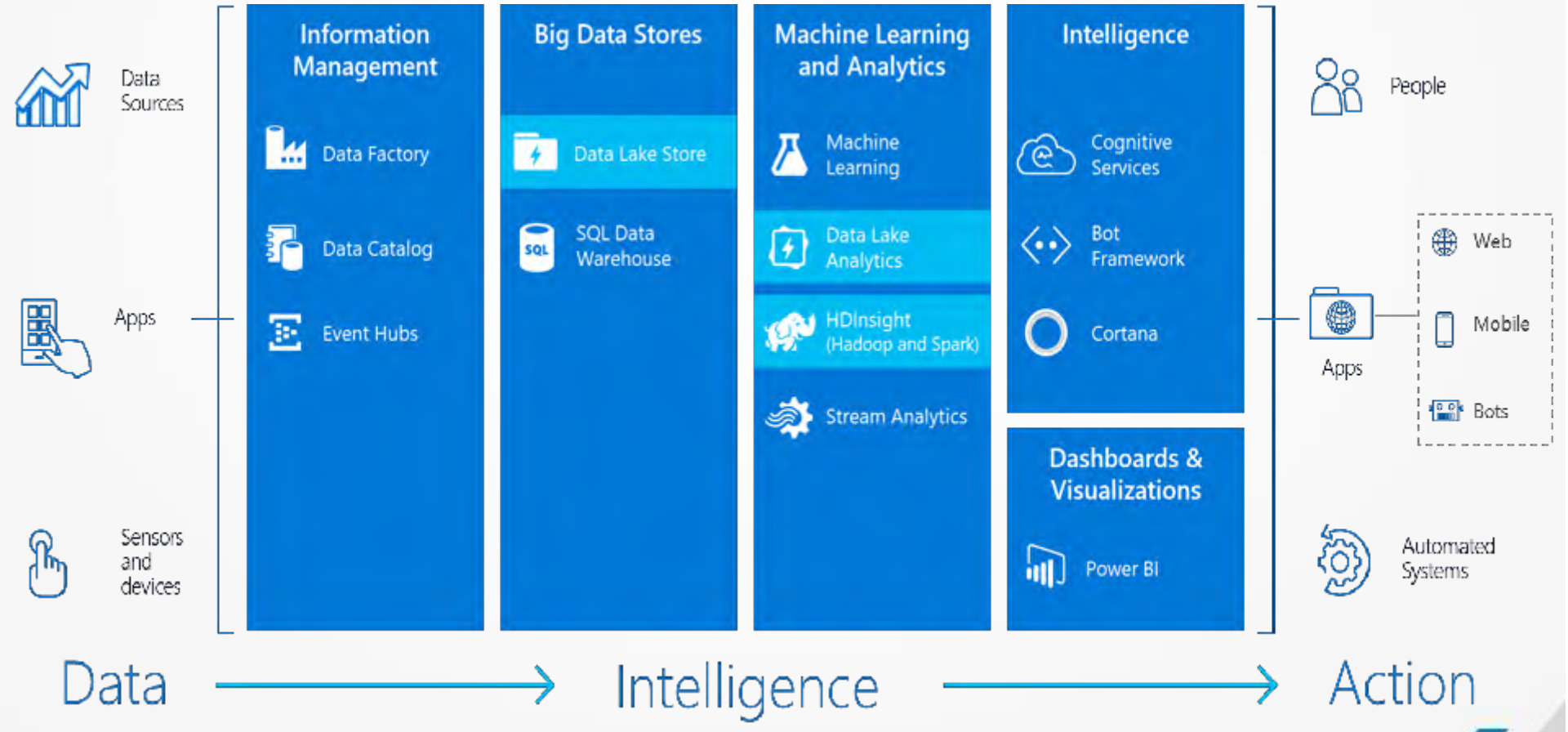
TiD2017

微软数据平台

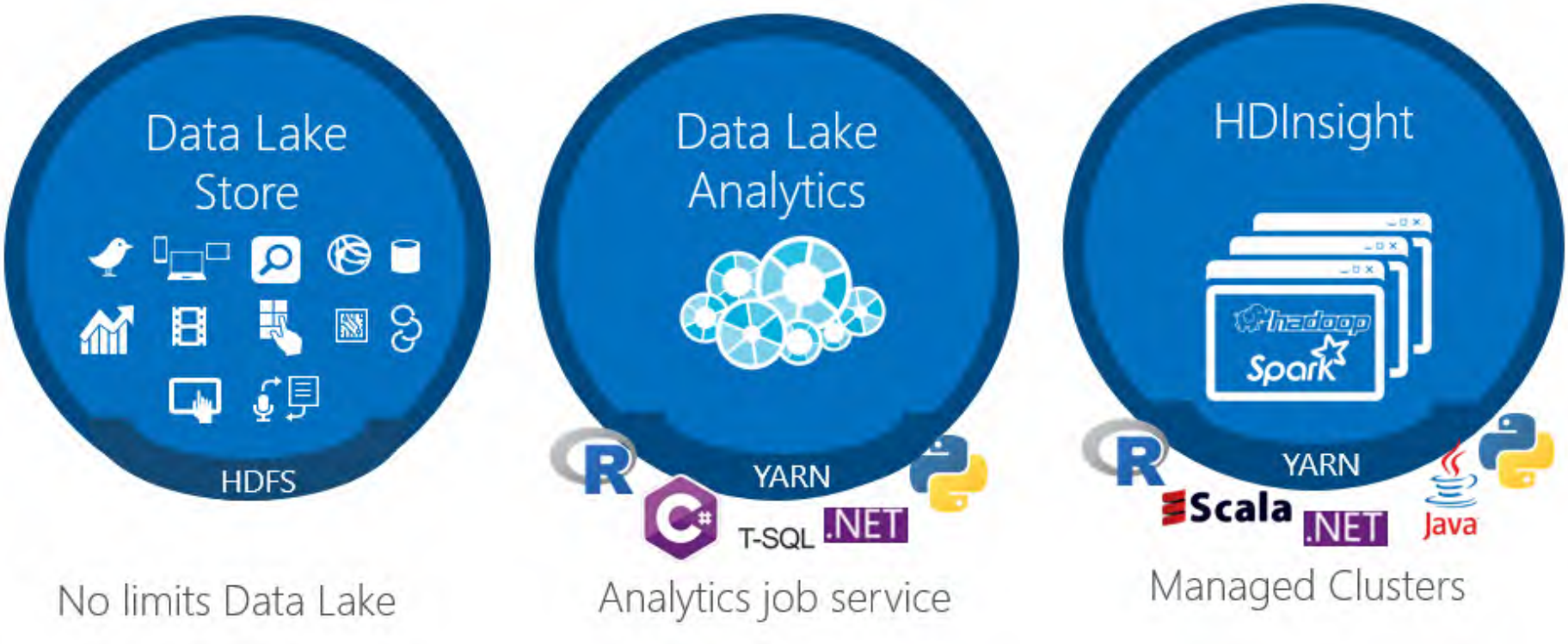


TiD2017

大数据作为Cortana智能套件的基石



Azure 数据湖泊介绍



Azure 数据湖泊

Big Data made easy

Analytics on any data, any size	Easier and more productive for all users	Enterprise-ready
		

Azure数据湖泊存储

BIG DATA IN THE CLOUD

Azure Data Lake Store

A hyper-scale repository for Big Data analytics workloads

Hadoop File System (HDFS) for the cloud

No limits to scale

Store **any data** in its native format

Enterprise-grade access control, encryption at rest

Optimized for analytic workload **performance**

Azure数据湖泊分析

Distributed analytics service built on Apache YARN

Elastic scale per query lets users focus on business goals—not configuring hardware

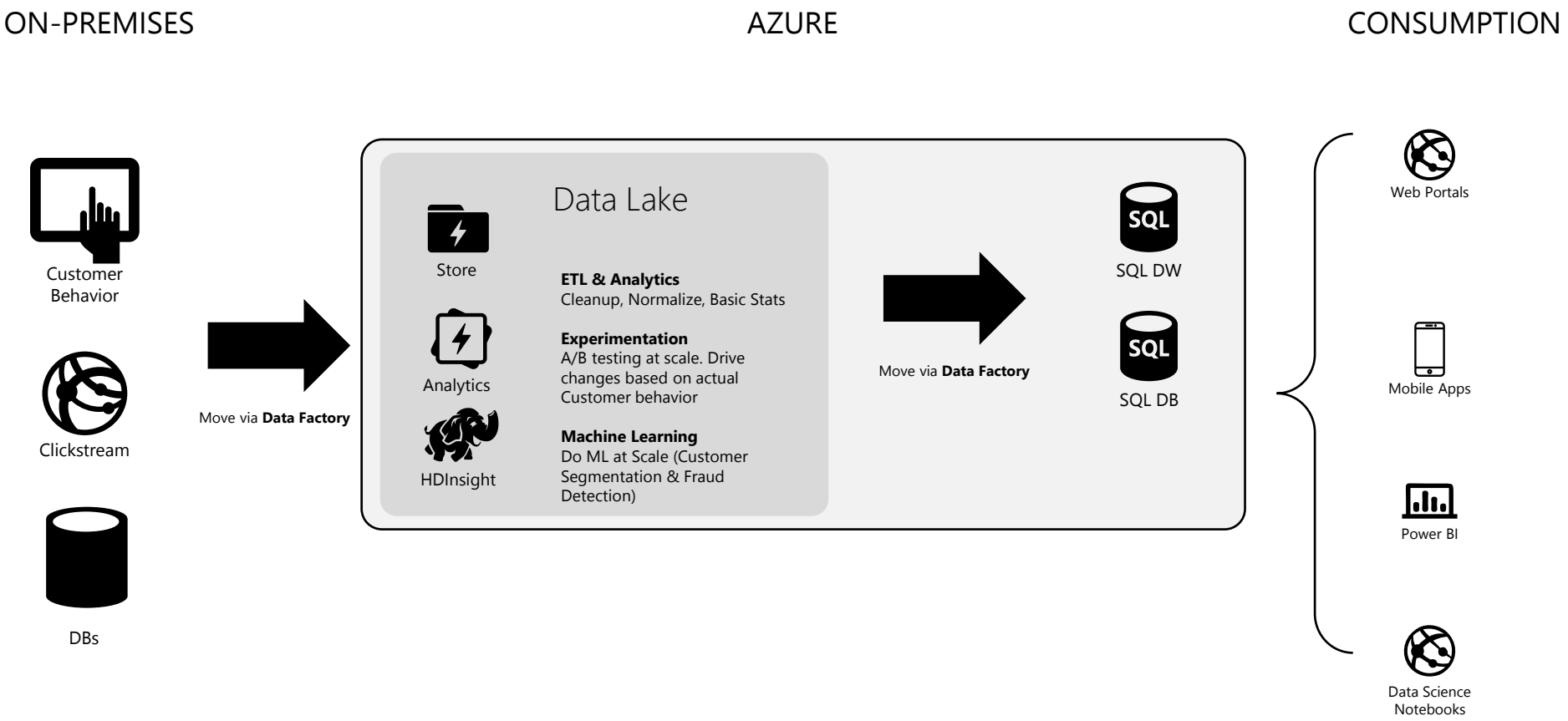
Includes U-SQL—a language that unifies the **benefits of SQL with the expressive power of C#**

Integrates with Visual Studio to develop, debug, and tune code faster

Federated query across Azure data sources

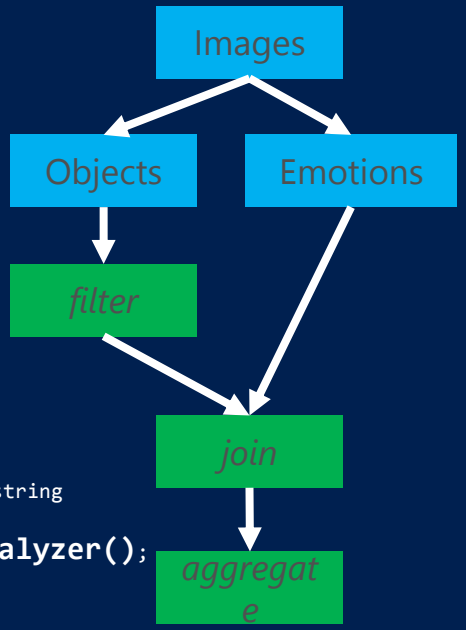
Enterprise-grade **role based access control**

数据湖泊分析场景


ON-PREMISES
AZURE
CONSUMPTION
Customer Behavior
Clickstream
DBs
Move via Data Factory
Data Lake
Store
Analytics
HDInsight
ETL & Analytics
Cleanup, Normalize, Basic Stats
Experimentation
A/B testing at scale. Drive changes based on actual Customer behavior
Machine Learning
Do ML at Scale (Customer Segmentation & Fraud Detection)
Move via Data Factory
SQL DW
SQL DB
Web Portals
Mobile Apps
Power BI
Data Science Notebooks

U-SQL/Cognitive Example

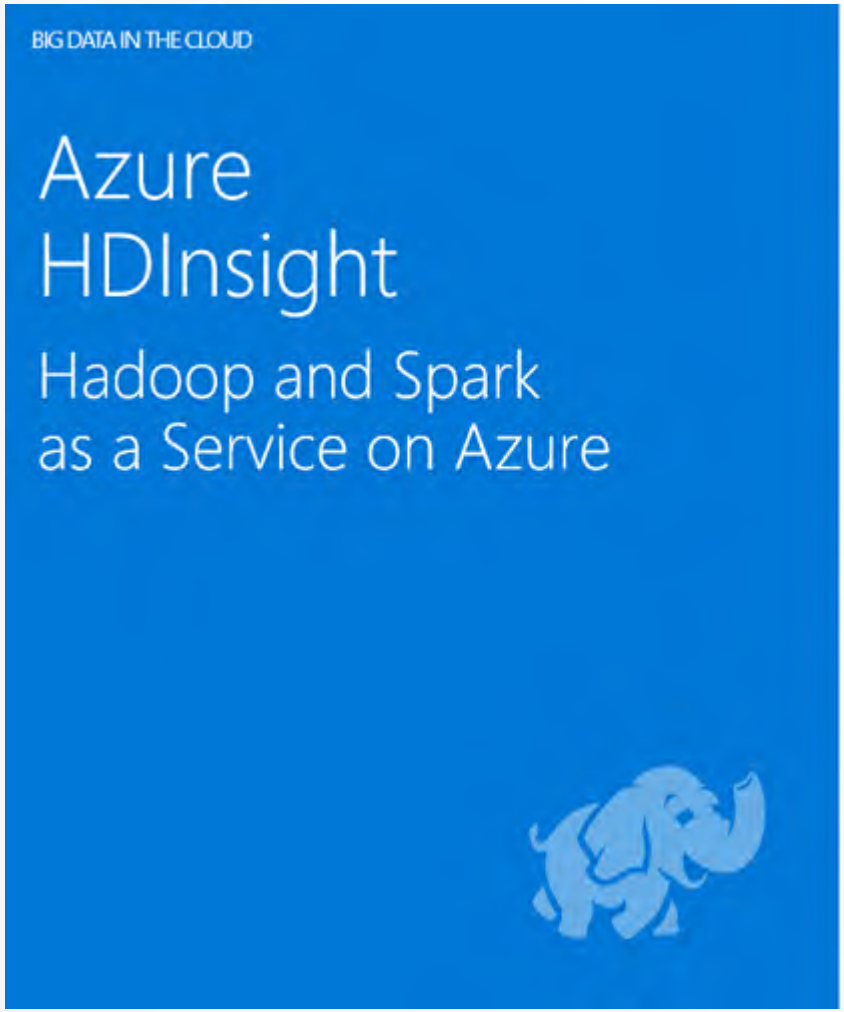
- Identify objects in images (tags)
- Identify faces and emotions and images
- Join datasets – find out which tags are associated with happiness




```
REFERENCE ASSEMBLY ImageCommon;
REFERENCE ASSEMBLY FaceSdk;
REFERENCE ASSEMBLY ImageEmotion;
REFERENCE ASSEMBLY ImageTagging;

@objects =
    PROCESS MegaFaceView
    PRODUCE FileName, NumObjects int, Tags string
    READONLY FileName
    USING new Cognition.Vision.ImageTagger();

@tags =
    SELECT FileName, T.Tag
    FROM @objects
         CROSS APPLY
             EXPLODE(SqlArray.Create(Tags.Split(';')))
                AS T(Tag)
    WHERE T.Tag.ToString().Contains("dog") OR
          T.Tag.ToString().Contains("cat");
```

```
@emotion_raw =
    PROCESS MegaFaceView
    PRODUCE FileName string, NumFaces int, Emotion string
    READONLY FileName
    USING new Cognition.Vision.EmotionAnalyzer();

@emotion =
    SELECT FileName, T.Emotion
    FROM @emotion_raw
         CROSS APPLY
             EXPLODE(SqlArray.Create(Emotion.Split(';')))
                AS T(Emotion);

@correlation =
    SELECT T.FileName, Emotion, Tag
    FROM @emotion AS E
         INNER JOIN
             @tags AS T
         ON E.FileName == T.FileName;
```

Azure HDInsight



Fully-managed Hadoop and Spark for the cloud

100% Open Source Hortonworks data platform

Clusters up and **running in minutes**

Managed, monitored and supported by Microsoft with the **industry's best SLA**

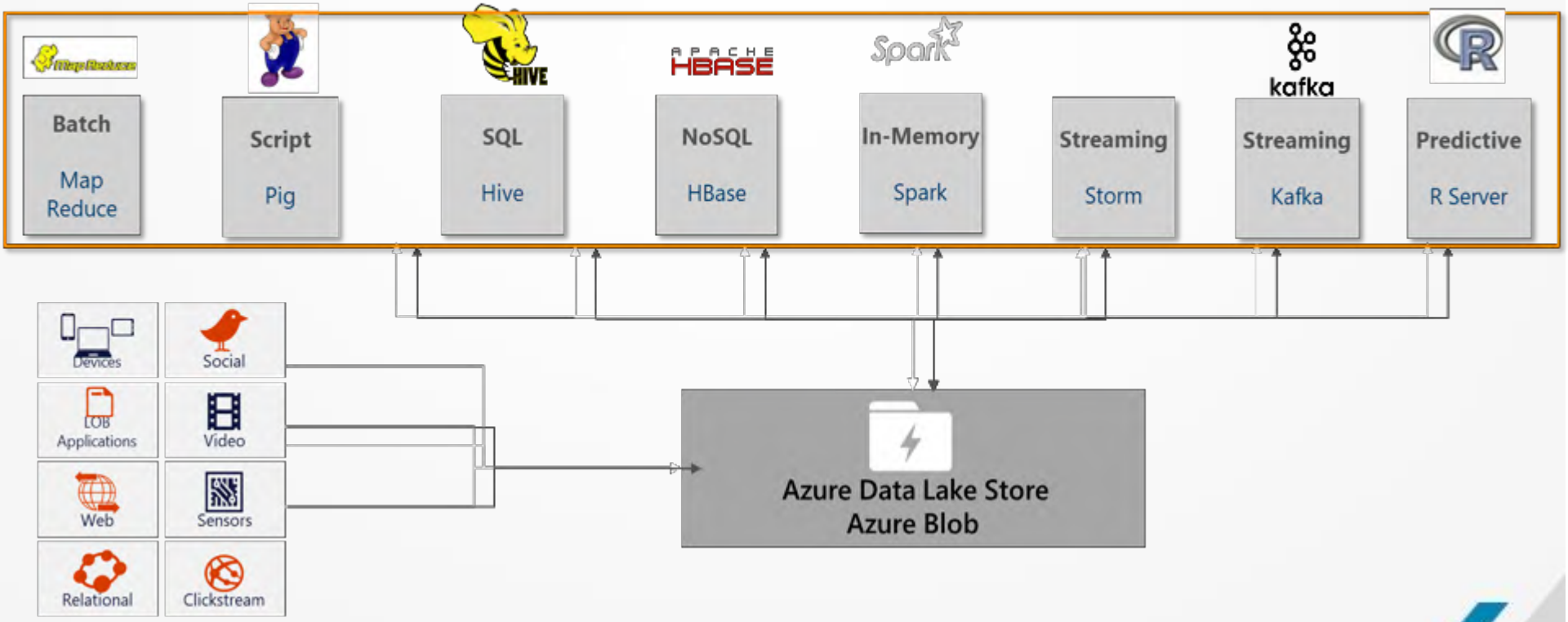
Familiar **BI tools for analysis**, or open source notebooks for **interactive data science**

63% lower TCO than deploy your own Hadoop on-premises*

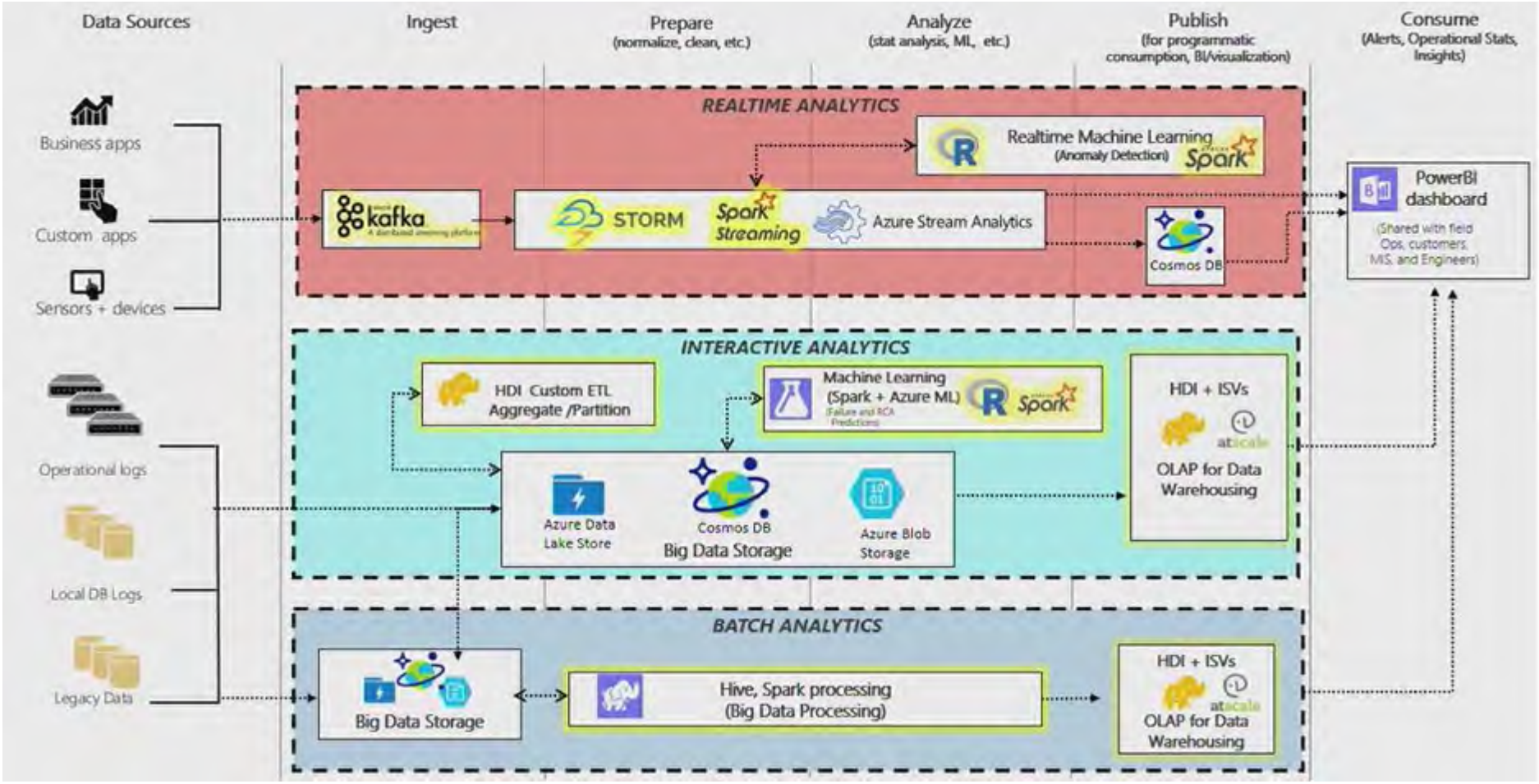
*IDC study "The Business Value and TCO Advantage of Apache Hadoop in the Cloud with Microsoft Azure HDInsight"

HDInsight - 大数据分析的工作负载

一个为大数据分析工作专业定制的可高度扩展，分布式和并行处理的云数据平台



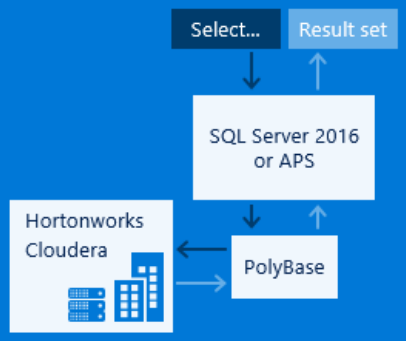
Reference Big Data Analytics Pipeline



PolyBase in SQL Server

BIG DATA ON-PREMISES

PolyBase in SQL Server 2016 and APS



For on-premises Big Data, partner with Hortonworks, Cloudera, or other 3rd party providers

Use **PolyBase to query** Hadoop and DW together

PolyBase is **built-in** to both SQL Server 2016 and APS with no additional costs

TiD2017

微软大数据平台的应用

大数据高级分析或物联网场景

Data type usage by vertical

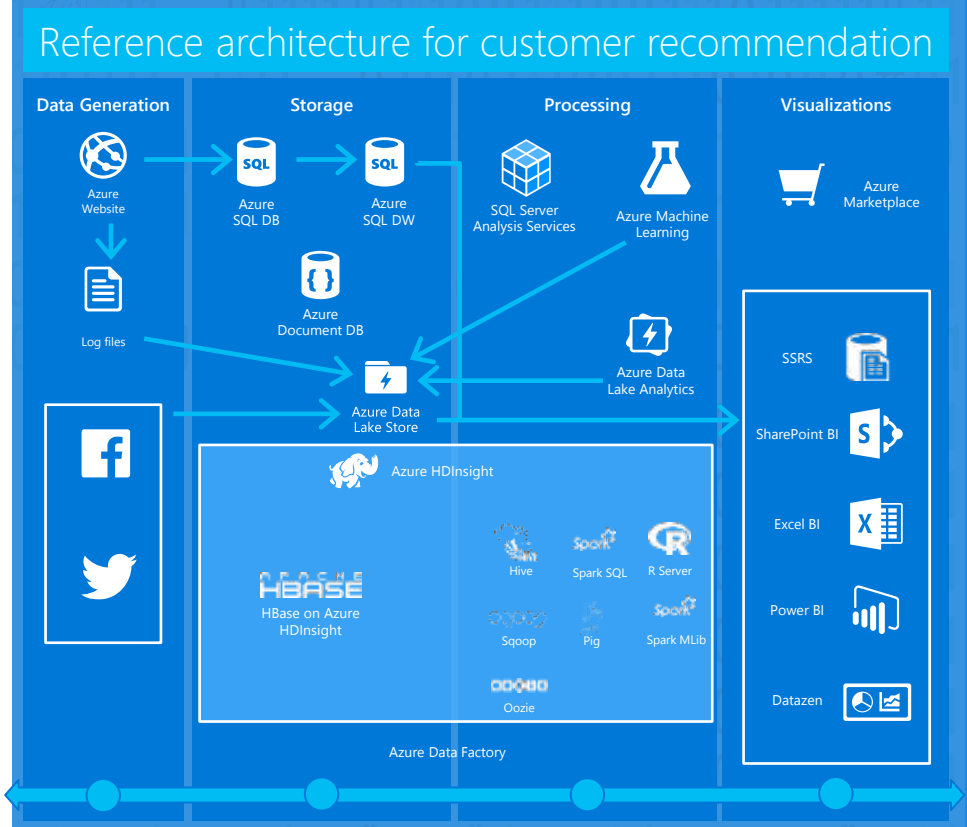
	Retail	Manufacturing	Financial Services	Health	Government
Solutions	• Real-time offers and personalized services • Demand forecasting • Sentiment analysis	• Manufacturing ops • Connected cars	• Customer experience • Risk assessment	• Remote health monitoring • Population health management	• Smart buildings • Transit and traffic optimization
New types of data	• Clickstream and behavior • Point of sales • Server logs • Sentiment and web	• Machine and sensor • Structured and unstructured	• Sentiment and web • Server logs • Structured and unstructured	• Patient vitals • Genomic data • Server logs	• Sentiment and web • Server logs • Structured and unstructured

大数据高级分析或物联网场景

Example reference architecture
Customer recommendation

Use Cortana Intelligence:

- Track non-relational data: use clickstream data / log files from the ecommerce website (ADLS)
- Track relational data: use customer purchase data (SQL DB)
- Use big data (ADLA) to do data preparation of non-relational data
- Use Machine Learning to model and predict customer behavior (AML) on both relational and non-relational data
- Have website display dynamic content from predictions to recommend products

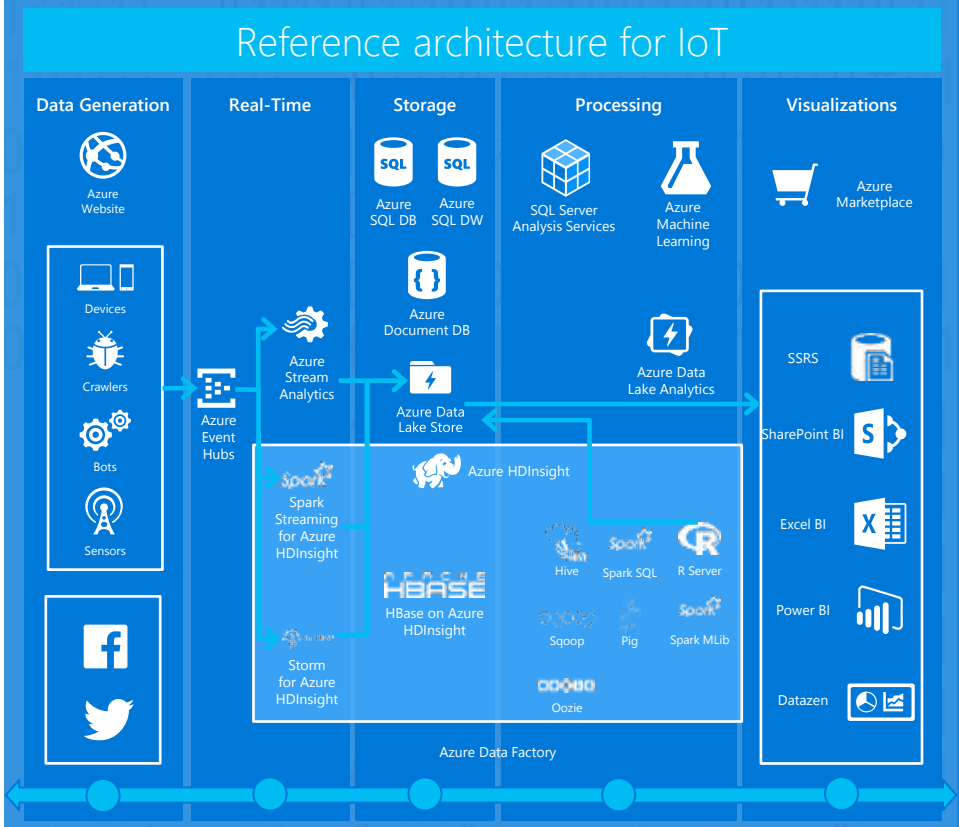


大数据高级分析或物联网场景

Example reference architecture
Internet of Things

Use Cortana Intelligence:

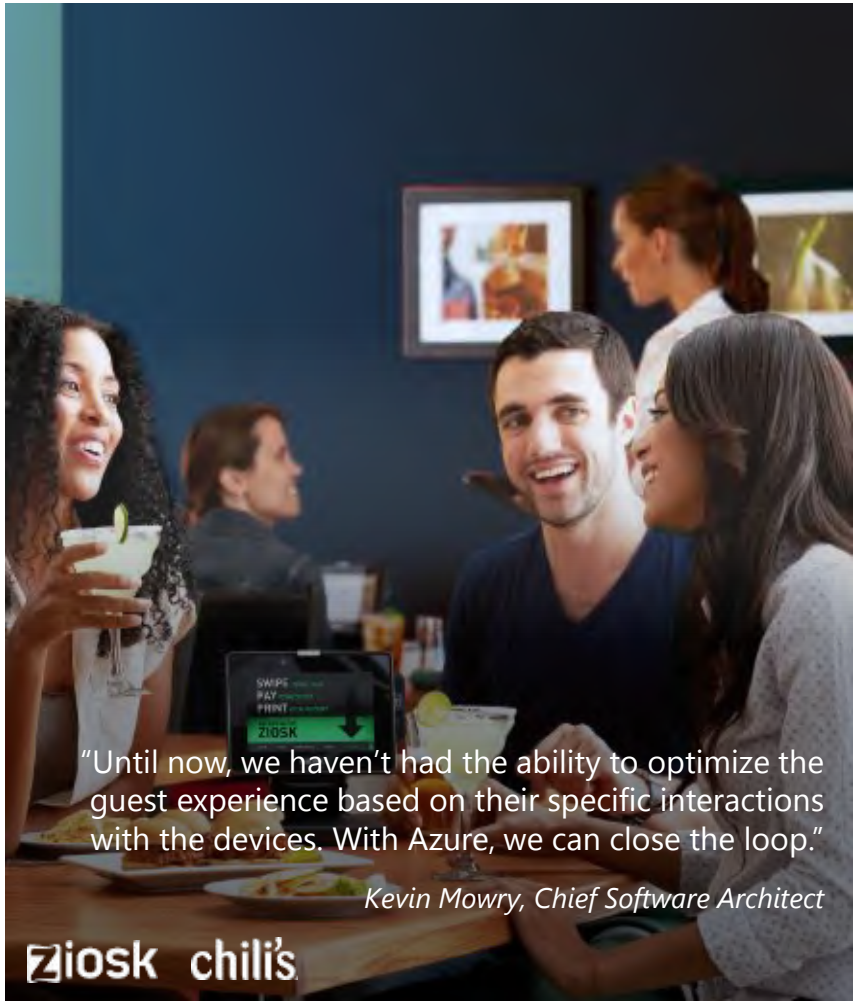
- Track non-relational data: collect devices or machine generated data in real-time (Event Hubs + Spark Streaming)
- Output real-time data to Power BI
- Land real-time data into permanent store (ADLS)
- Run machine learning at scale on data to find patterns (HDInsight + R Server)



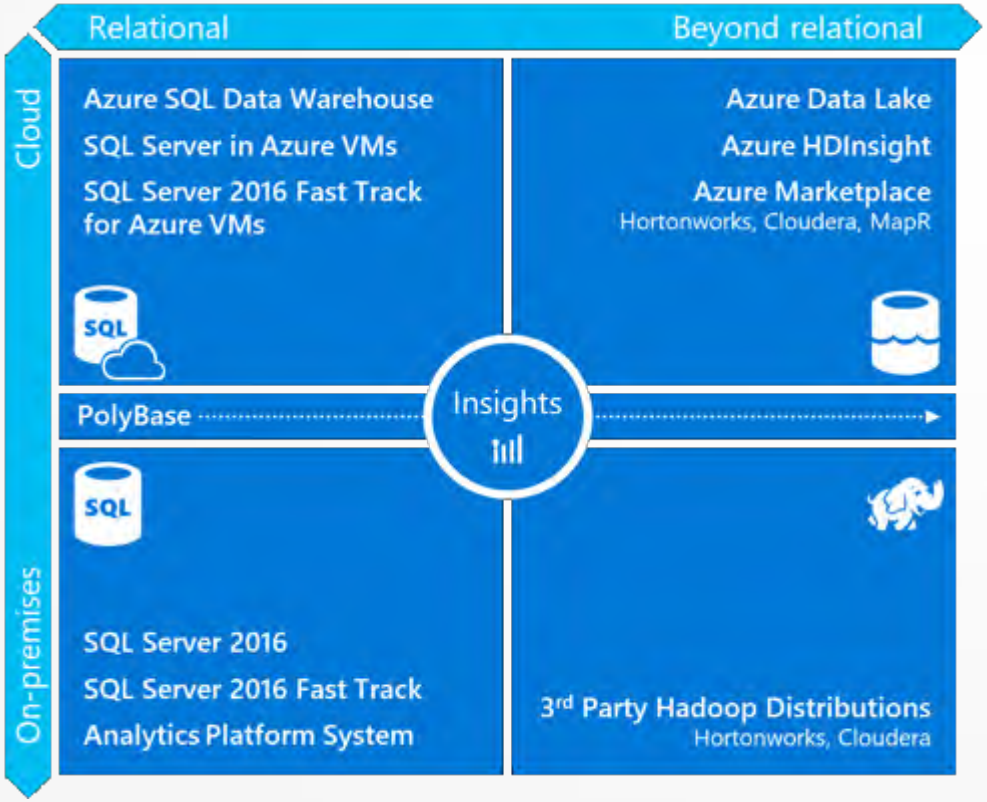
大数据高级分析或物联网场景

Customer example: Chili's Restaurants using Ziosk Tablets for IoT scenario

Scenario	Ziosk Tablets on every restaurant | Wanted to improve guest satisfaction, customer insights, and restaurant efficiency
Solution	Azure HDInsight (Hadoop-as-a-service), Azure Machine Learning, Power BI to aggregate data in real-time and identify relationships between customer behavior and purchases
Result	• Optimize guest experience on tablets by delivering customized offers in real-time • Understand restaurant metrics such as customer wait times, wait staff efficiencies, restaurant sales, etc.



基于大数据的数据仓库



Fastest insights
Real-time insights with breakthrough query performance

Analytics built-in
Real-time insights with analytics built in

Choice of deployment
Leading solutions—on-premises and in the cloud

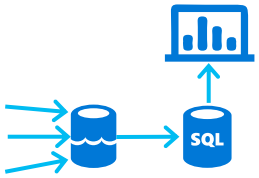
Layers of security
Least vulnerable database 6 years in a row

Any data, any scale
A hybrid solution that grows in step with customer needs

More for the price
Customers do more with industry-leading TCO

数据仓库的大数据现代化

Common technical scenarios for modernizing data warehouse with big data

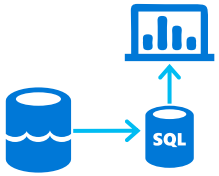


ETL Offload Data Prep

- Leverage cost effective big data to offload ETL and data preparation
- Transformation is done using big data scripts/code instead of through an ETL tool
- When data is prepared, load into warehouse for interactive BI

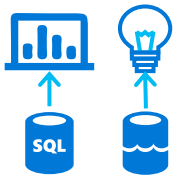
Single view (DW + Big Data)

- Customer 360 degree view now includes both relational (customer purchases, contacts) and non-relational (sentiment, clickstream, etc.)
- Data warehouse and Data Lake are side-by-side collecting data and queried using federated query technologies (PolyBase)



Active Archive (DW Hot storage / Big Data cold)

- Leverage cost effective big data to archive all data indefinitely as “cold” storage that can be processing at anytime with big data
- Pull in frequently used hot data for interactive BI in the data warehouse



Run data science / experimentation

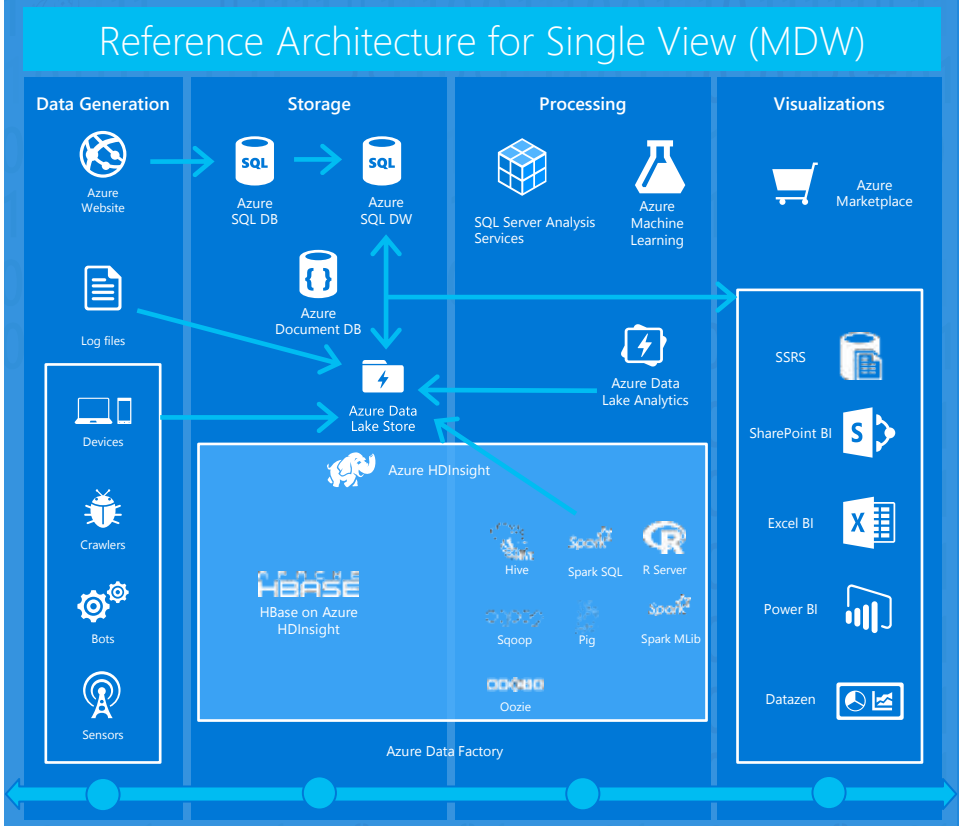
- Keep data warehouse for operational reporting and BI
- Leverage a separate big data implementation for running data science / experimentation with uncurated, non-relational data to find new patterns and new insights
- When patterns are found, operationalize into the data warehouse for reporting
- Use R Server for data science

数据仓库的大数据现代化

Example reference architecture
Modern Data Warehouse [Single View]

Use Cortana Intelligence:

- Track relational data: Move relational point of sales and purchase data through operational stores (SQL DB) into relational DW (SQL DW)
- Track non-relational data: Track non-relational data from telemetry, clickstream, social medial into big data store (ADLS)
- Run queries across joining the two through federated query (PolyBase or ADLA)
- Show results in BI tools (Power BI)



数据仓库的大数据现代化

Customer example: Tangerine instantly adapts to customer feedback

Scenario	Lack of insight for targeted campaigns Inability to support data growth
Solution	Azure HDInsight (Hadoop-as-a-service) and Microsoft Analytics Platform System (APS) to do instant analysis of social sentiment and customer feedback across digital, face-to-face and phone interactions.
Result	• Reduced time to customer insight • Ability to make changes to campaigns or adjust product rollouts based on real-time customer reactions • Ability to offer incentives and new services to retain—and grow—its customer base

通用的大数据解决方案

Target Audience

Technical Decision Maker (primary)

Who care about supporting the business with strategic IT investments

ITPros (influencer)

Example: Database Administrators, Operations team

Business Decision Maker (influencer)

LOB Decision Makers who care about solving business problems

Probing Questions

What do you currently do for big data today?

What new types of **non-relational** data do you need to capture (e.g. sentiment and web, clickstream, machine/sensor, geographic, server logs, etc.)?

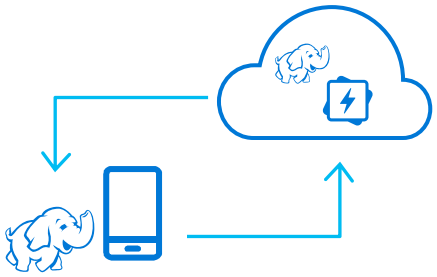
Do you currently have a **Hadoop deployment on-premises** and want to get out of the business of managing your own datacenter?

How do you collect Azure cloud data generated by other sources like Azure Website, IoT Suite, etc. and what kinds of analysis would you like performed on it?

What experiments / data science do you want on **all of your data (curated or not)?**

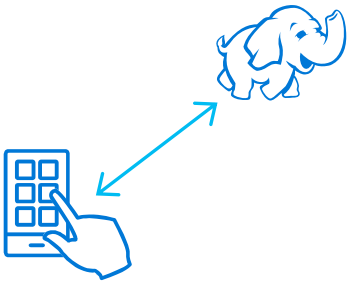
通用的大数据解决方案

Common technical scenarios for big data stand-alone



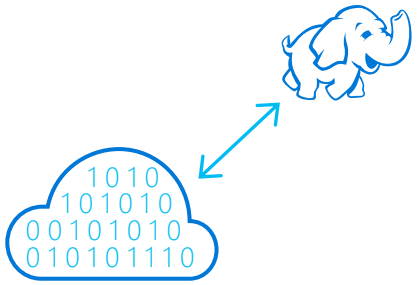
On-prem Hadoop migration to cloud/hybrid

- Leverage cost effective cloud to manage Hadoop and get out of managing own servers
- Hybrid Hadoop scenarios with on-prem (use cloud for dev/test, backup, or burst)
- Use R Server with on-prem Hadoop



Adding intelligence to applications

- Add big data to applications for the app to receive data in real-time and predict customer behavior



Analyze data stored in Azure from other services

- Azure services generating lots of data (e.g. Azure Web Apps, IOT Suite, Media Services, etc.). Use Data Lake Analytics or HDInsight to process data to find new insights

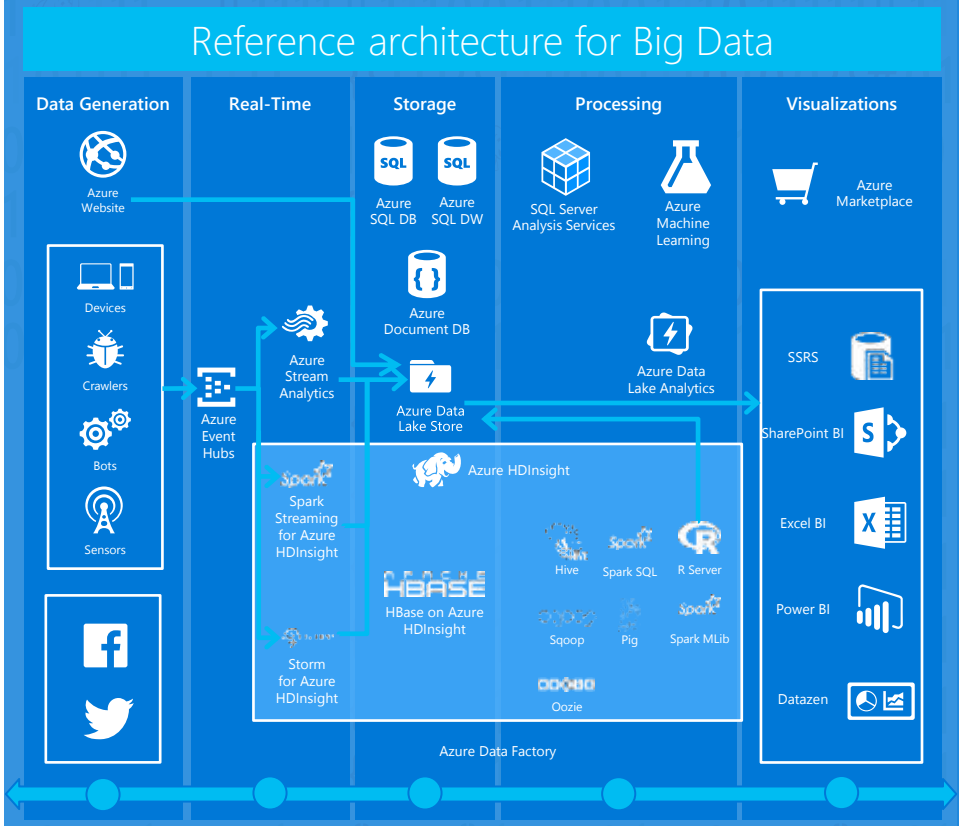
Run data science / experimentation

- Leverage a new big data implementation for running data science / experimentation with uncurated, non-relational data to find new patterns and new insights

通用的大数据解决方案

Example reference architecture
Analyze Data Stored in Azure

- Track realtime data from IOT Suite: collect data from IOT Suite in permanent store (ADLS)
- Track other Azure data (Azure Website generating web logs) and store in ADLS
- Run Machine Learning through R Server for HDInsight to find patterns in data
- Show results in BI tools (Power BI)



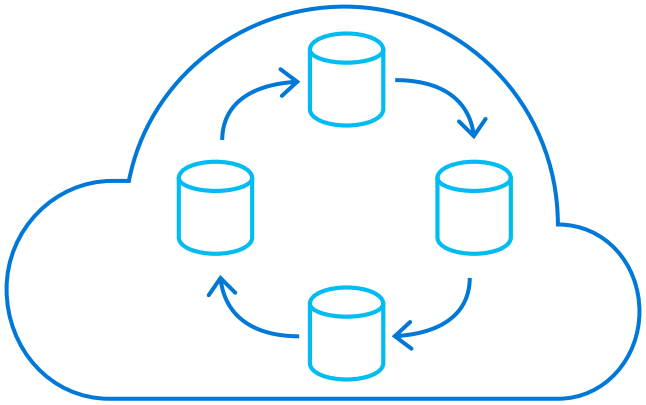
通用的大数据解决方案

Customer example: Virginia Tech crunch endless amounts of Genomic data

Scenario	DNA sequencers are generating 15PB of genomic data each year. Virginia Tech needed to process it to foster medical breakthroughs including new cancer treatments. They were evaluating creating a multimillion dollar supercomputer center, but wanted to find a different way to process the data.
Solution	Azure HDInsight (Hadoop-as-a-service) was chosen to process genome data resulting in significant cost savings as they only pay for what they need.
Result	• Significant cost savings with the cloud • Elastic scale that keeps up with huge data volumes • Powering the search for cancer treatments

微软大数据开发工具

Easy for administrators to spin up quickly



- Deploy big data projects in minutes
- No hardware to install, tune, configure or deploy
- No infrastructure or software to manage
- Scale to tens to thousands of machines instantly

Debug and Optimize your Big Data programs with ease

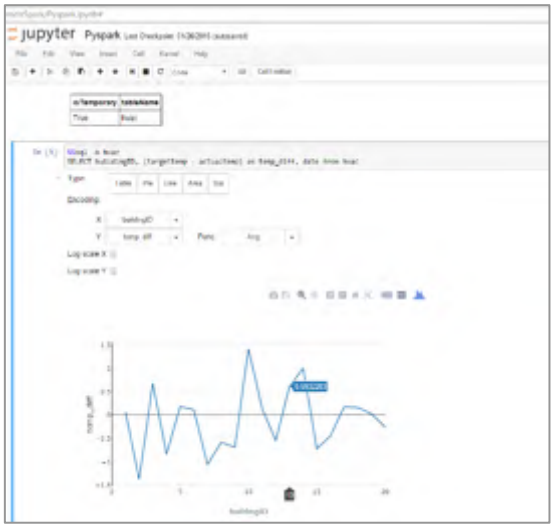
- Deep integration with Visual Studio, Visual Studio Code, Eclipse, & IntelliJ
- Easy for novices to write simple queries
- Integrated with U-SQL, Hive, Storm, and Spark
- Actively offers recommendations to improve performance and reduce cost
- Playback visually displays job run

Develop massively parallel programs with simplicity

```
@t = EXTRACT date string
           , time string
           , author string
           , tweet string
     FROM "/Input/MyTwitterHistory.csv"
     USING Extractors.Csv();

@res = SELECT author AS author
           , COUNT(*) AS tweetcount
       FROM @t
       GROUP BY author;

OUTPUT @res TO "/Output/MyTwitterAnalysis.csv"
ORDER BY tweetcount DESC
USING Outputters.Csv();
```

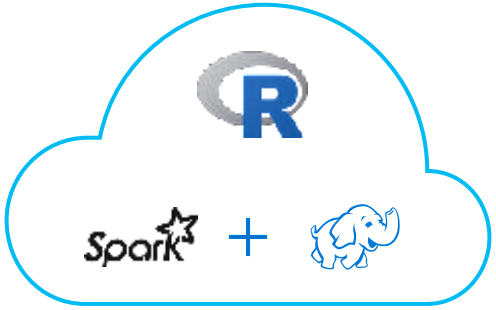
- U-SQL: a simple and powerful language that's familiar and easily extensible
- Unifies the declarative nature of SQL with expressive power of C#
- Leverage existing libraries in .NET languages, R and Python
- Massively parallelize code on diverse workloads (ETL, ML, image tagging, facial detection)

Easy notebook experience for data scientists



- Most popular notebooks, Jupyter and Zeppelin out-of-the-box
- Combine code, statistical equations and visualizations
- Worked w/ Jupyter community to enhance kernel to allow Spark execution through REST endpoint

Easy for data scientists with familiar R language

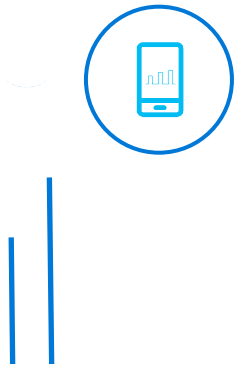


*Applies to HDInsight only

R Server for HDInsight

- Largest portable R parallel analytics library
- Terabyte-scale machine learning—1,000x larger than in open source R
- Up to 100x faster performance using Spark and optimized vector/math libraries
- Enterprise-grade security and support

Easy for business analysts with interactive reports over big data



- Interactive BI with big data
- Spark 2.1 integration
- Interactive Hive with LLAP- keeps data compressed running in-memory 25x faster
- ODBC driver to use Power BI or third party tools (Tableau, SAP, Qlik, etc.)

央媒客户案例

全球精准传播系统

案例：全球精准传播系统

发现全球各国家、地区热点事件，了解热点事件的传播影响力

发现传播过程中的传播主体（意见领袖、活跃用户、搭桥者等角色）。分析其社交媒体影响力。

甄选传播主体，根据其画像，利用推荐引擎进行精准内容投放。
利用其影响力定向引导舆论

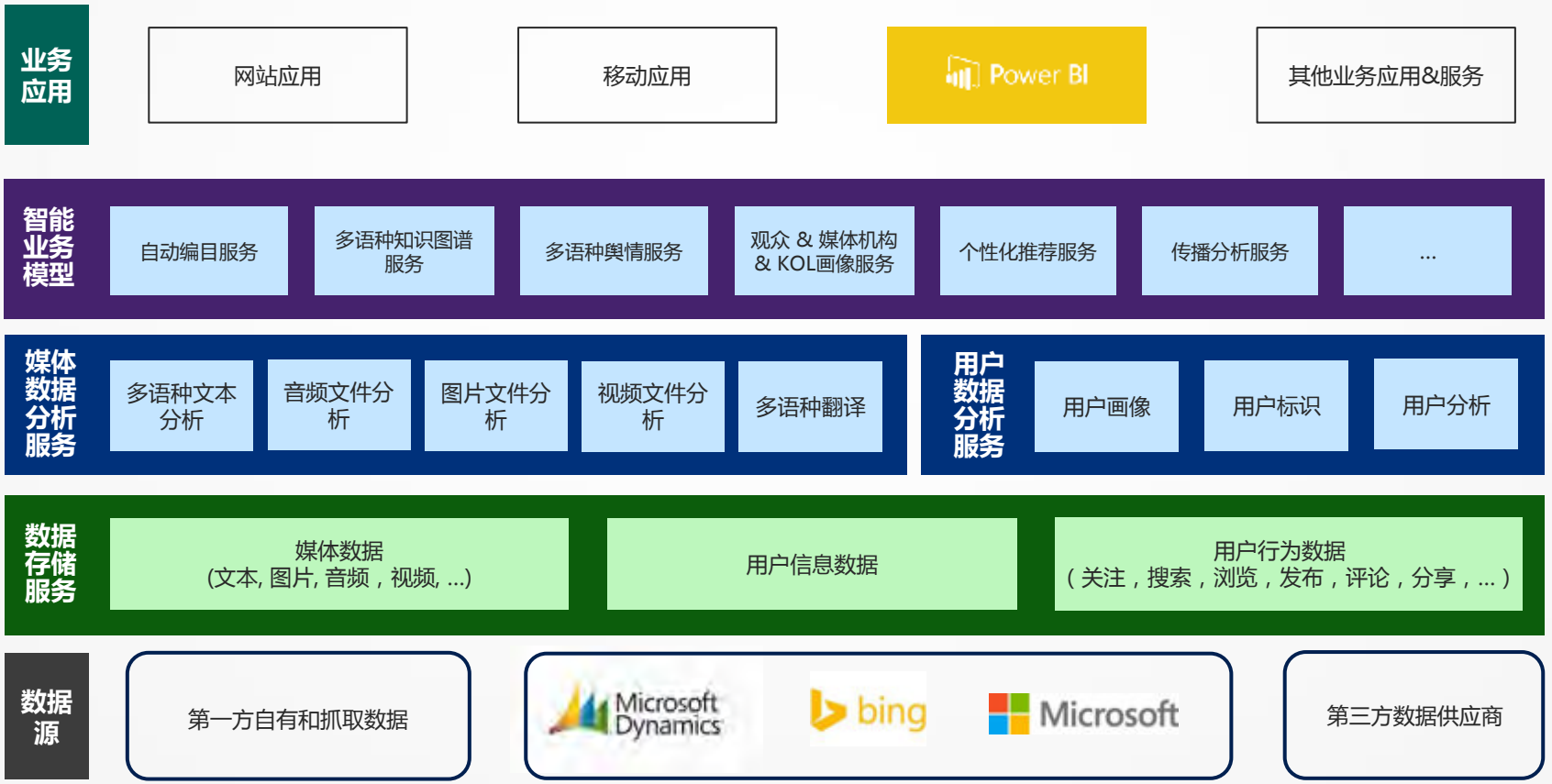
对传播主体进行全维度画像，深入了解传播主体属性。

典型业务场景

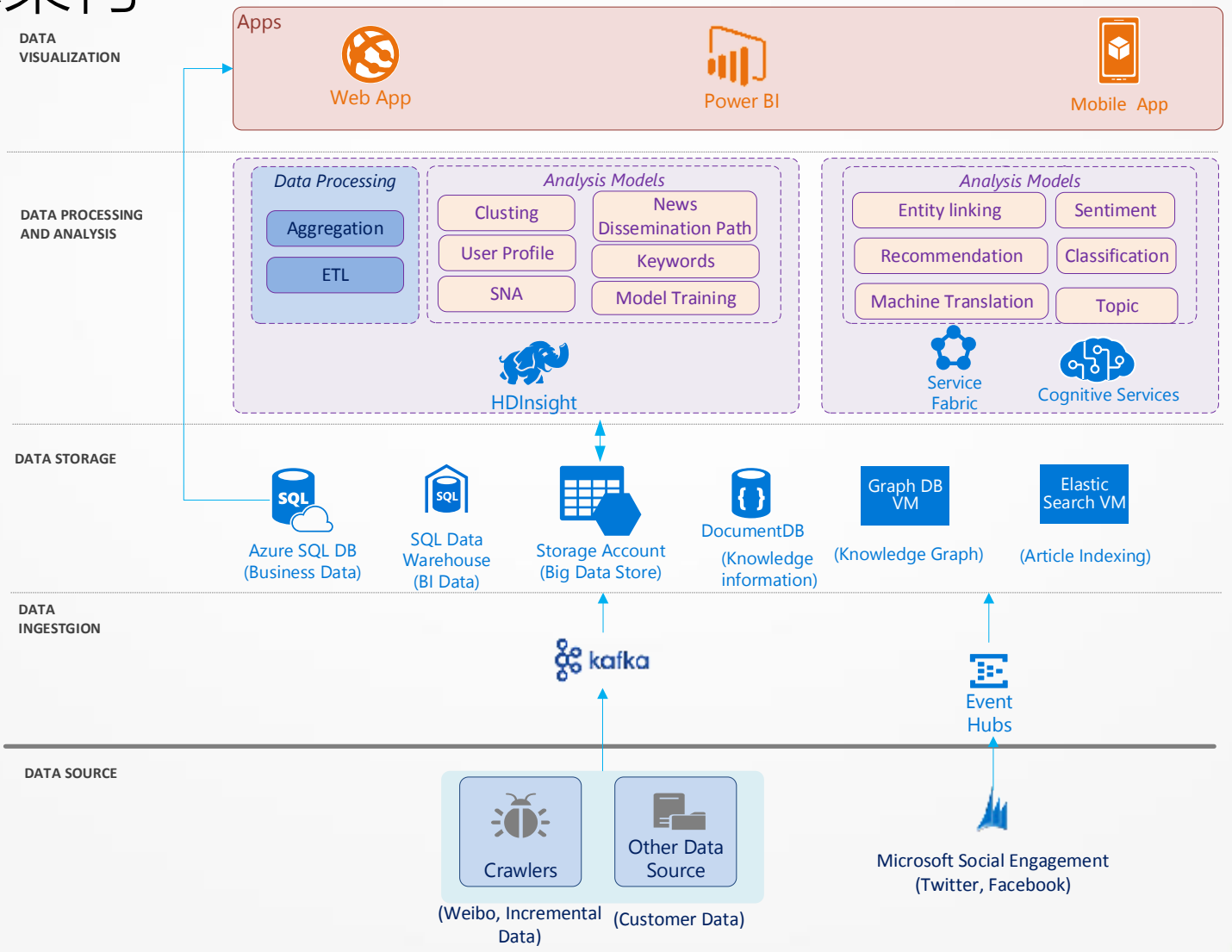
党政机构：
- 日常外宣（国家形象、旅游宣传、...）
- 高层出访造势
- 国际倡议（一带一路、亚投行、...）

央企/规模民企：
- 海外工程（高铁、水坝、...）
- 海外投资（矿山、机场、港口、...）
- 海外产品销售

功能架构


业务应用
网站应用
移动应用
Power BI
其他业务应用&服务
智能业务模型
自动编目服务
多语种知识图谱服务
多语种舆情服务
观众 & 媒体机构 & KOL画像服务
个性化推荐服务
传播分析服务
...
媒体数据分析服务
多语种文本分析
音频文件分析
图片文件分析
视频文件分析
多语种翻译
用户数据分析服务
用户画像
用户标识
用户分析
数据存储服务
媒体数据 (文本, 图片, 音频，视频, ...)
用户信息数据
用户行为数据 (关注，搜索，浏览，发布，评论，分享，...)
数据源
第一方自有和抓取数据
Microsoft Dynamics
bing
Microsoft
第三方数据供应商

技术架构


DATA VISUALIZATION
Apps
Web App
Power BI
Mobile App
DATA PROCESSING AND ANALYSIS
Data Processing
Aggregation
ETL
Analysis Models
Clusting
News Dissemination Path
User Profile
Keywords
SNA
Model Training
HDInsight
Analysis Models
Entity linking
Sentiment
Recommendation
Classification
Machine Translation
Topic
Service Fabric
Cognitive Services
DATA STORAGE
Azure SQL DB (Business Data)
SQL Data Warehouse (BI Data)
Storage Account (Big Data Store)
DocumentDB (Knowledge information)
Graph DB VM
(Knowledge Graph)
Elastic Search VM
(Article Indexing)
DATA INGESTGION
kafka
Event Hubs
DATA SOURCE
Crawlers
Other Data Source
(Weibo, Incremental Data)
(Customer Data)
Microsoft Social Engagement (Twitter, Facebook)

微软方案核心能力

多语种分析能力：
- 多语种自然语言处理
- 多语种自动翻译

基础设施：
- 全球各大洲高可靠云服务能力

解决方案：
- 全球舆情洞察
- 社交网络传播主体分析
- 媒体用户画像

全球多渠道数据采集能力：
- 15+主要国家和地区
- 15+主要语种
- Facebook、Twitter等主流社交媒体

微软AI的应用

TiD2017

数字化转型+ 人工智能

Conversational agents
Customized experiences
Customer analytics

Engage your customers

Transform your products

Enable your employees

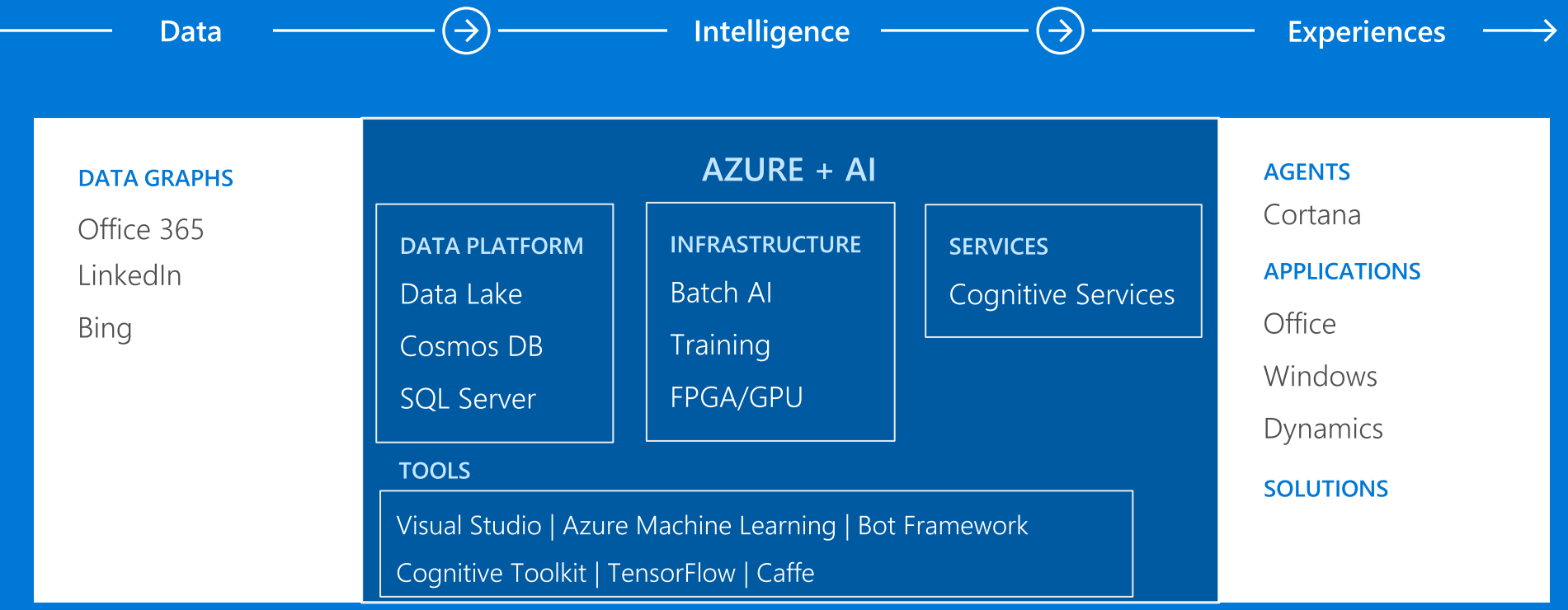
Optimize your operations

Product innovation
Differentiated experiences
New scenarios

Employee productivity
Business data differentiation
Organizational knowledge

Intelligent predictions
Operational efficiency
Deep insights

微软人工智能

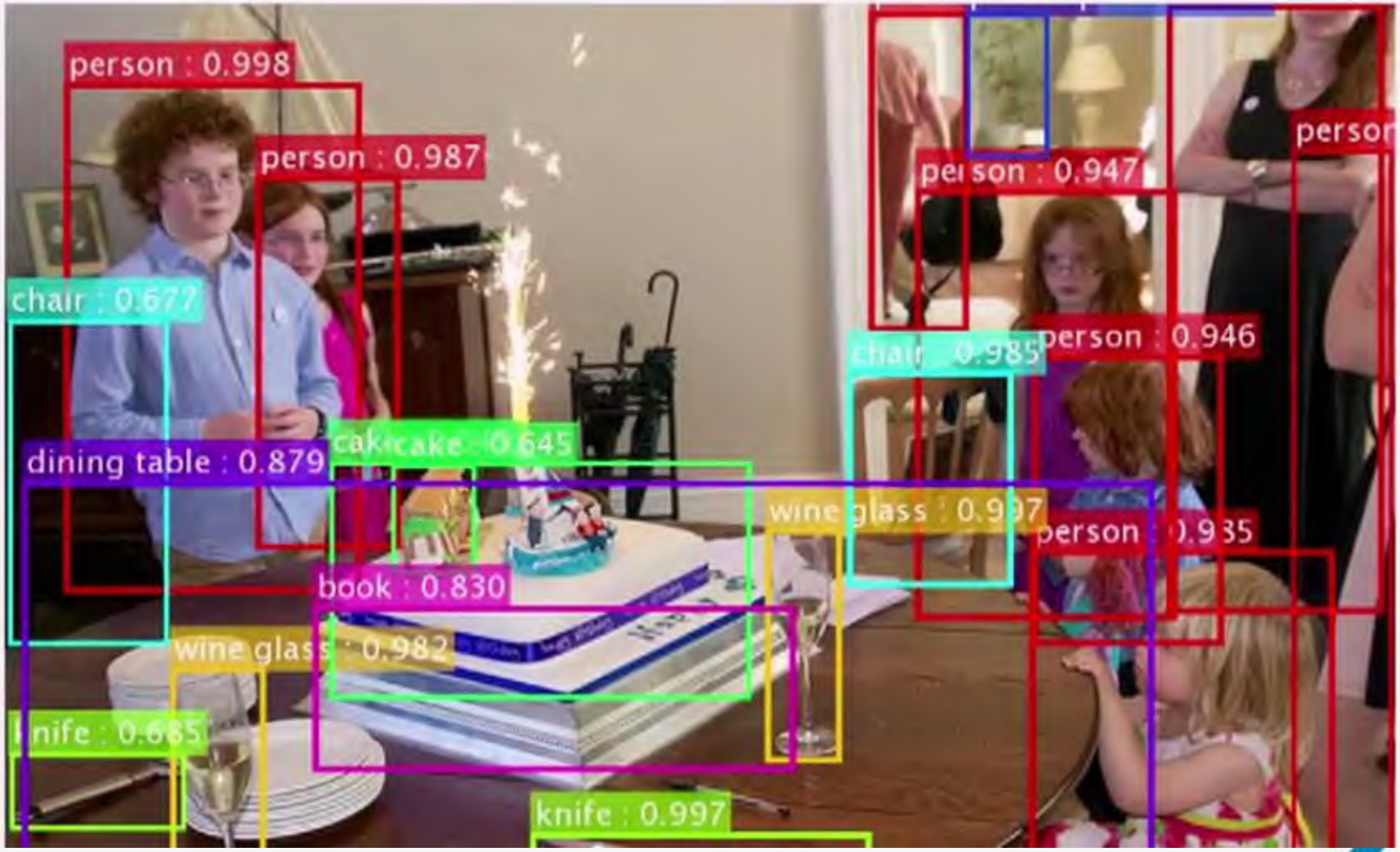

Data
Intelligence
Experiences
DATA GRAPHS
Office 365
LinkedIn
Bing
AZURE + AI
DATA PLATFORM
Data Lake
Cosmos DB
SQL Server
INFRASTRUCTURE
Batch AI
Training
FPGA/GPU
SERVICES
Cognitive Services
TOOLS
Visual Studio | Azure Machine Learning | Bot Framework
Cognitive Toolkit | TensorFlow | Caffe
AGENTS
Cortana
APPLICATIONS
Office
Windows
Dynamics
SOLUTIONS

微软认知服务

Vision	Language	Speech	Search	Knowledge
Computer vision	Text analytics	Speaker recognition	Web search	Academic knowledge
Face	Spell check	Speech	Image search	Entity linking service
Emotion	Web language model		Video search	Knowledge exploration
Content Moderator	Linguistic analysis		News search	Recommendations
Video	Translator		Autosuggest	QnA maker
NEW Video Indexer				
NEW Cognitive Services Labs				
NEW Custom Vision Service	Custom Language Understanding	Custom Speech Service	NEW Custom Search	NEW Custom Decision Service

TiD2017

物体识别与检测



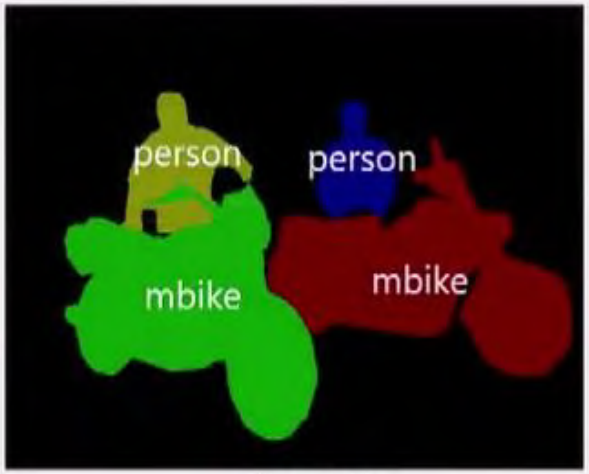
TiD2017

实例分割

- 比物体检测和语义分割更为困难

物体检测


语义分割



实例分割

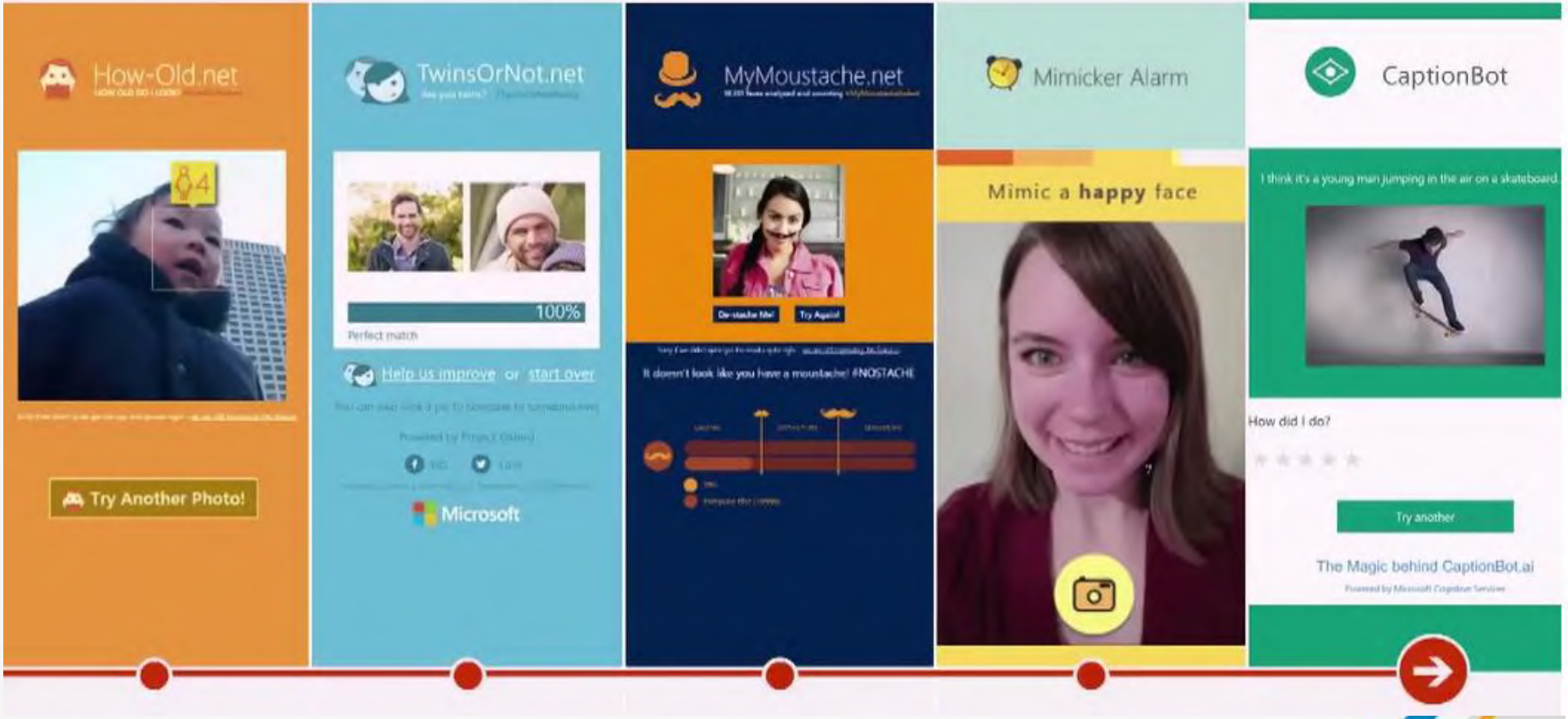
语义分割

Recognizing pixels

TiD2017

人工智能应用

各种有趣的App



TiD2017

标题Bot

www.captionbot.ai

CaptionBot

I think it's a young man jumping in the air on a skateboard.

TiD2017

语言翻译

Microsoft Translator - 成熟和实用的人工智能

大规模使用神经网络运算模型

翻译质量比传统的统计学算法有显著提高，并会不断提升
大规模部署微软AI超级计算机和认知工具包 CNTK
同时发布10种语言DNN 翻译，快速规模化新技术的应用

面向中国市场

新增粤语的翻译支持
即将部署到位于中国的数据中心
不断增长的中国的消费者和企业用户

elex 360 www.360.com 網易 NETEASE www·163·com YOUZU.COM 游族网络

TiD2017

图片识别案例 – 宝贝回家

图片识别案例 – 宝贝回家

助 宝贝回家

微软认知服务 - 人脸识别技术

1-2秒匹配照片库中几万张照片

27 个面部重要特征点识别

云端部署，跨平台、跨地域访问

微软Bot 框架

Accelerate development

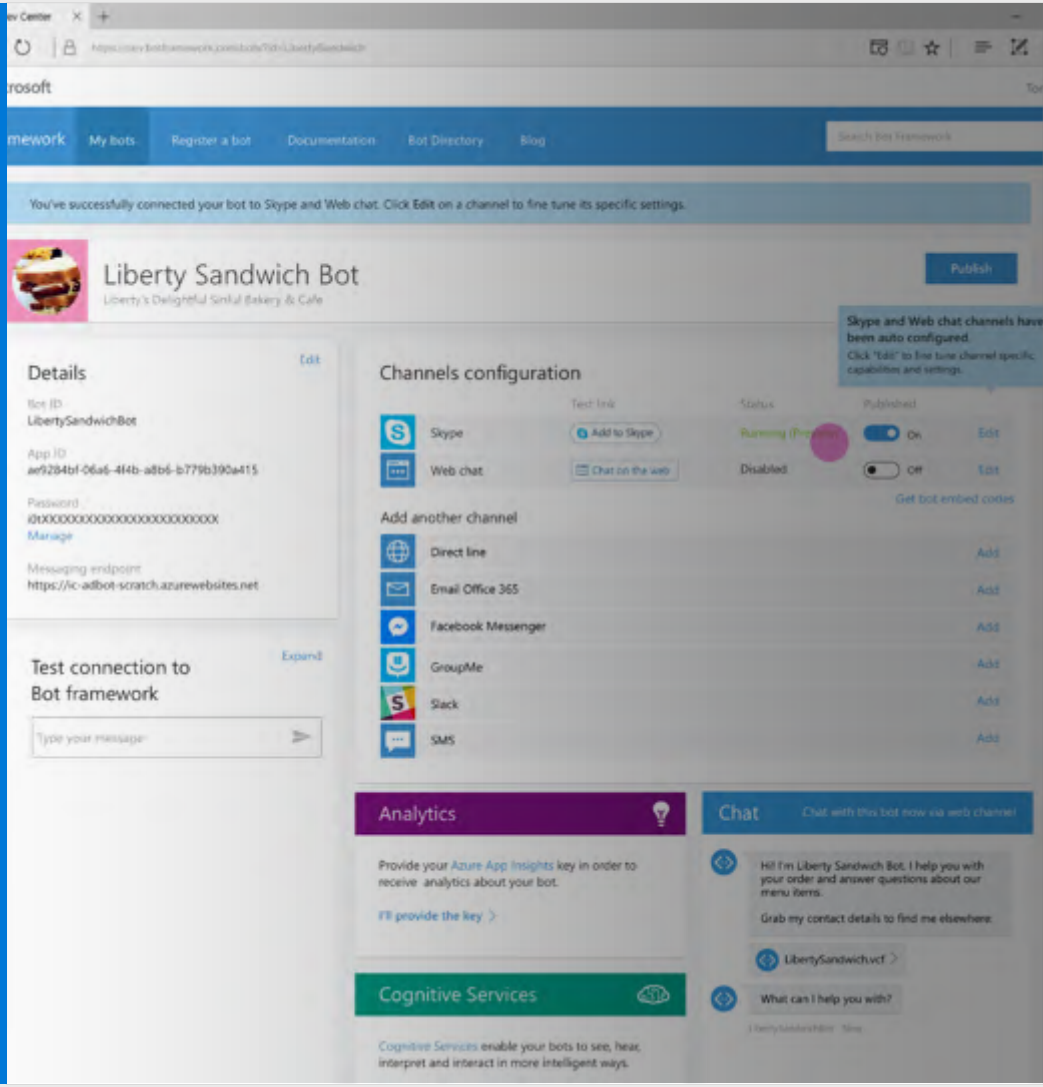
Framework, tools and services to build great bots that converse naturally

Reach any user

Develop once and expose through any channel including Skype, Facebook, web

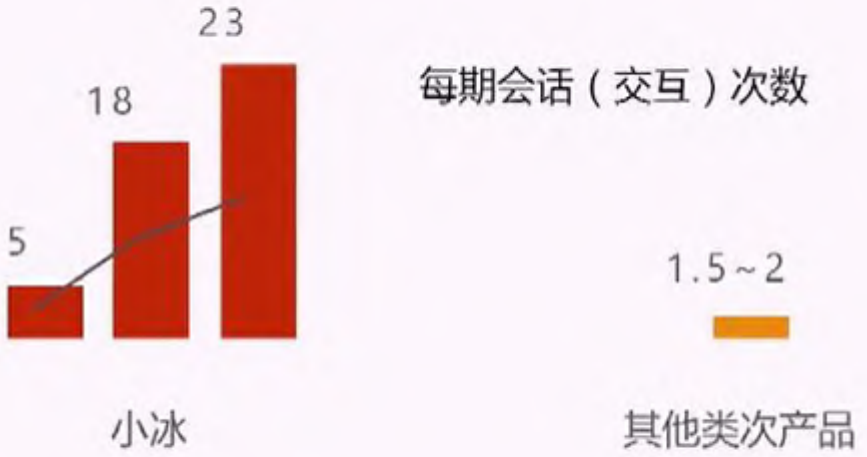
Enrich the experience

Add advanced capabilities through close integration with Cognitive Services



小冰 & Rinna

- 用户
  - 超过4千万注册（中国 + 日本）
  - 超过2万个公众号 (中国)
- 扩展功能
  - 图片和录像评论
  - 情绪分析
  - 京东，中国东方航空，Lawson， 用户互动
  - 东方卫视，天气预报
- 扩展地区
  - 中国
  - 日本
  - 美国

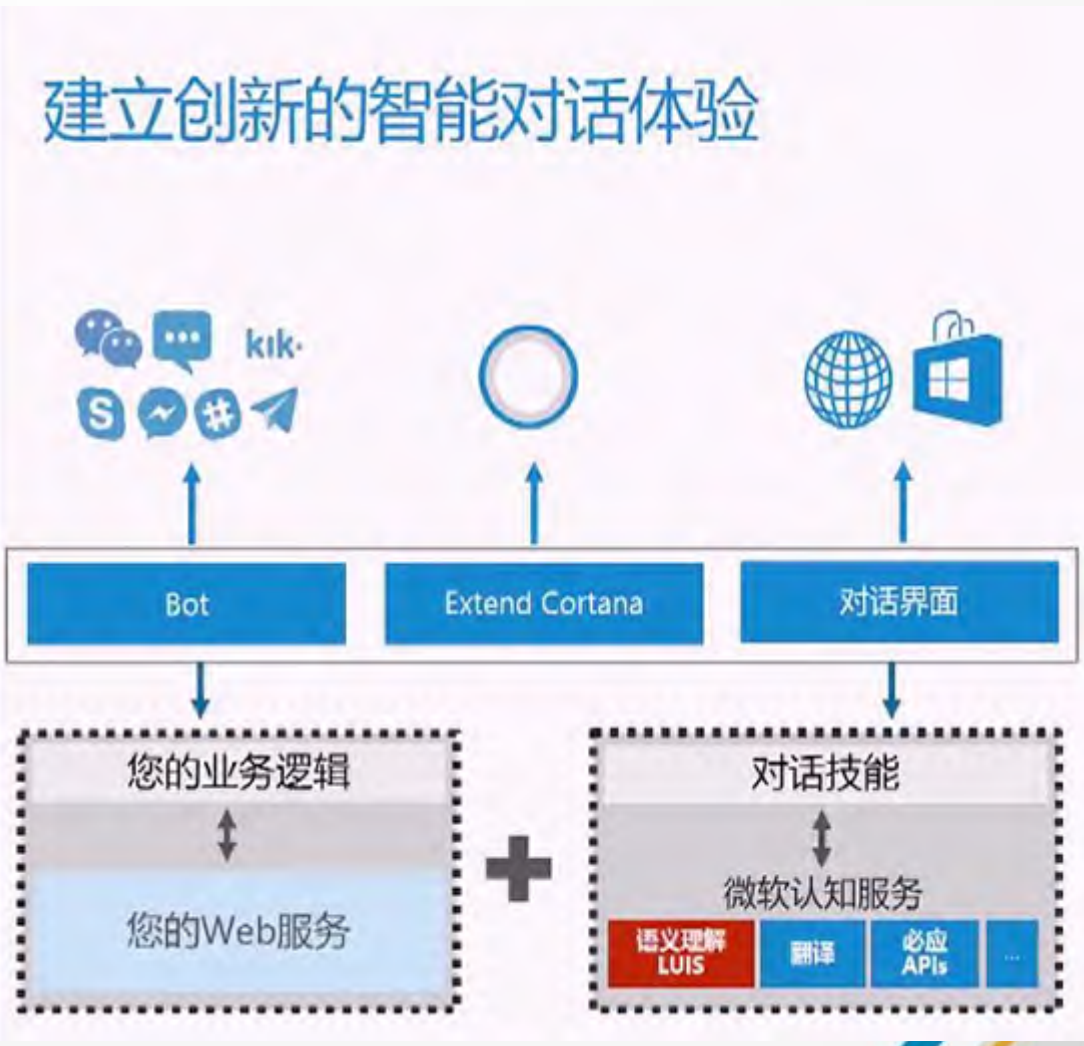


智能对话

Bots 给业务和用户粘度提供了新的机会和渠道:

- 随时随地、不受平台与设备限制地为客户提供服务
- 以一种更自然的方式与用户进行交互：在上下文语境中与用户会话
- 在用户量巨大的各种消息平台上开拓您的应用与服务
- 提升用户体验，减少人工帮助的请求

建立创新的智能对话体验



微软AI工具

Azure Machine Learning

Integrated environment

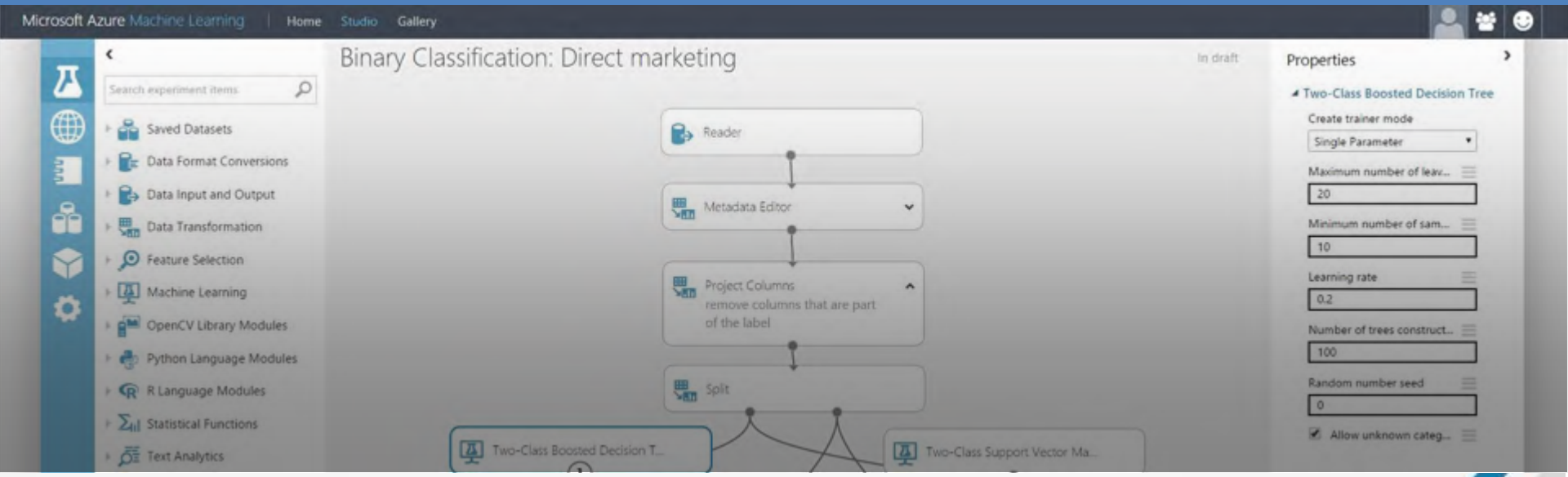
Simple drag, drop, and connect interface for Data Science

Best-in-class algorithms

Built-in collection of best-of-breed algorithms. R, CRAN packages, and Python

Deploy in minutes

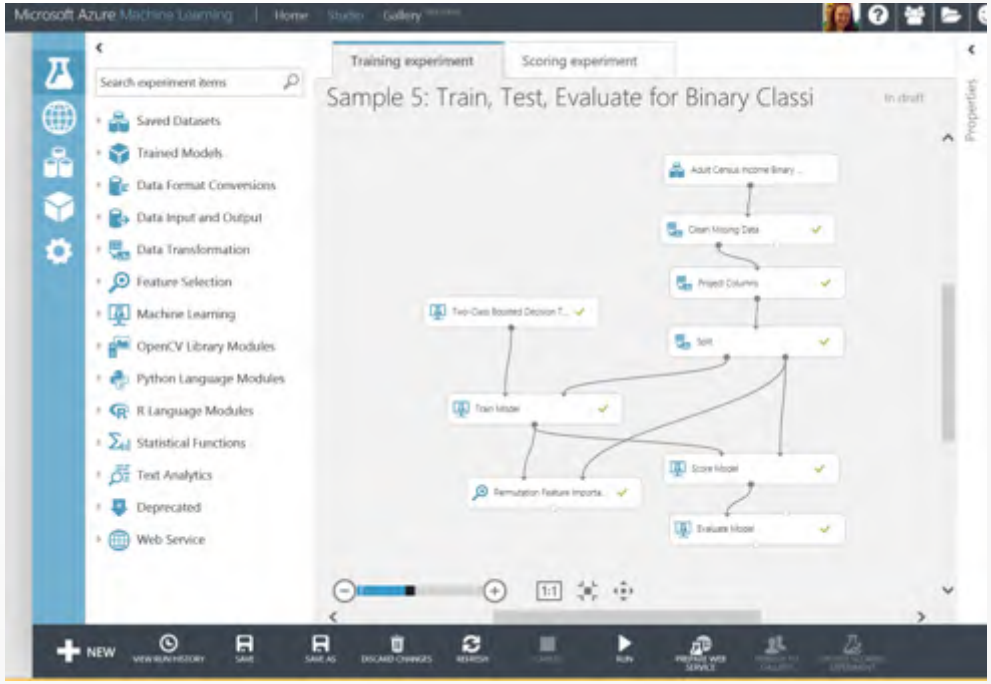
Train and operationalize models with a single click



Azure Machine Learning Studio

Azure上基于GUI，用于创建和操作机器学习工作流的集成开发环境。

- 通过浏览器访问
- 通过可视化的模块E2E支持数据流
- 提供多种ML算法和模型库
- 支持R和Python扩展;
- 支持Web API快速部署和伸缩



Deep Learning

Open environment

Cognitive Toolkit: performance, scale

Full support for TensorFlow, Caffee

Productive tools

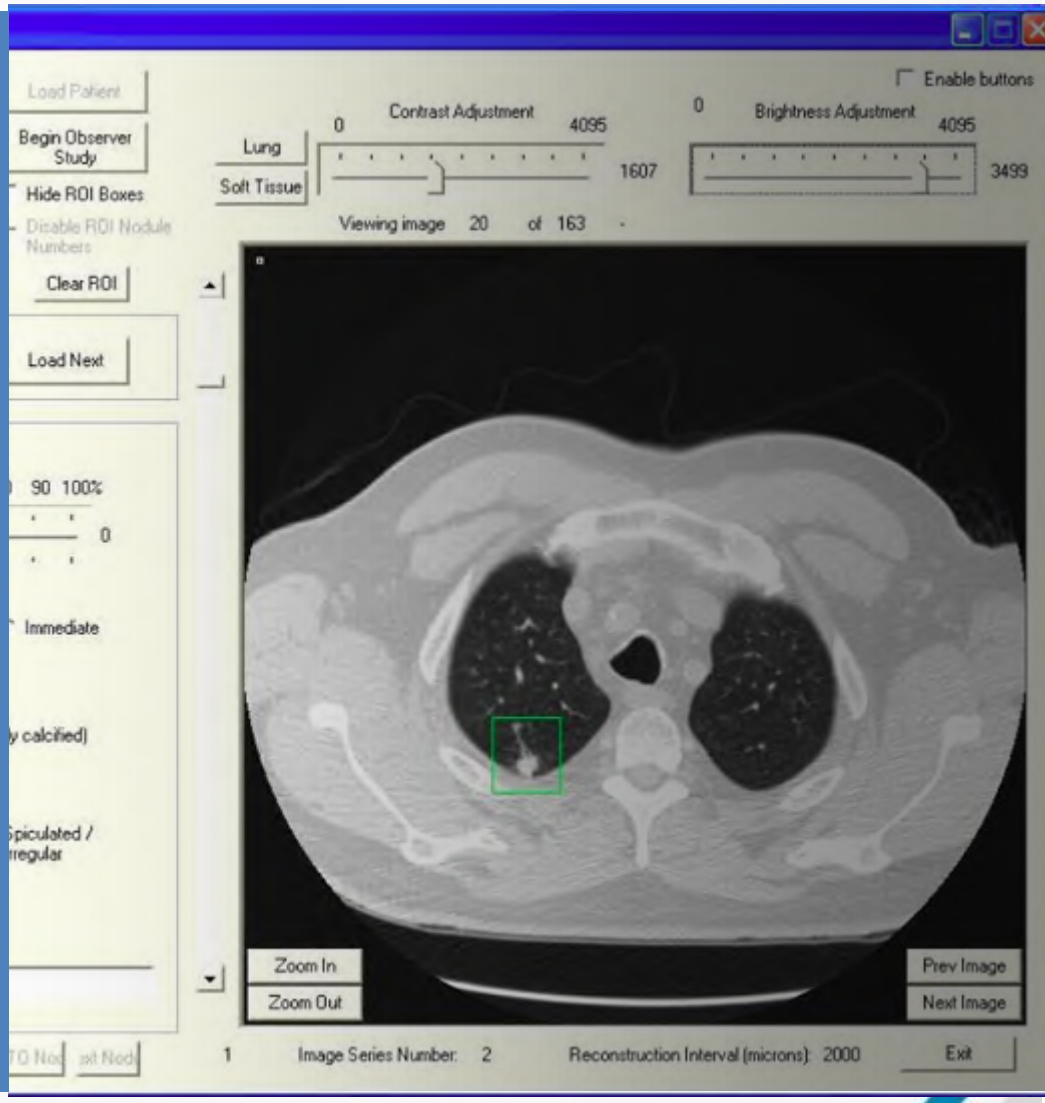
Visual Studio

Jupyter Notebook

Azure AI infrastructure

The AI supercomputer: CPU, GPU, and (coming next) FPGA

Managed service with Batch AI Training



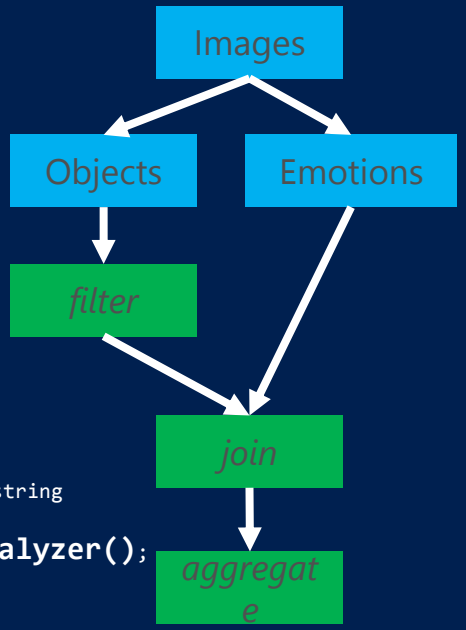
谢谢！

Jenny Jiang jejiang@Microsoft.com
Frank Chen frank.chen@microsoft.com

Appendix

U-SQL/Cognitive Example

- Identify objects in images (tags)
- Identify faces and emotions and images
- Join datasets – find out which tags are associated with happiness




```
REFERENCE ASSEMBLY ImageCommon;
REFERENCE ASSEMBLY FaceSdk;
REFERENCE ASSEMBLY ImageEmotion;
REFERENCE ASSEMBLY ImageTagging;

@objects =
    PROCESS MegaFaceView
    PRODUCE FileName, NumObjects int, Tags string
    READONLY FileName
    USING new Cognition.Vision.ImageTagger();

@tags =
    SELECT FileName, T.Tag
    FROM @objects
         CROSS APPLY
             EXPLODE(SqlArray.Create(Tags.Split(';')))
                AS T(Tag)
    WHERE T.Tag.ToString().Contains("dog") OR
          T.Tag.ToString().Contains("cat");
```

```
@emotion_raw =
    PROCESS MegaFaceView
    PRODUCE FileName string, NumFaces int, Emotion string
    READONLY FileName
    USING new Cognition.Vision.EmotionAnalyzer();

@emotion =
    SELECT FileName, T.Emotion
    FROM @emotion_raw
         CROSS APPLY
             EXPLODE(SqlArray.Create(Emotion.Split(';')))
                AS T(Emotion);

@correlation =
    SELECT T.FileName, Emotion, Tag
    FROM @emotion AS E
         INNER JOIN
             @tags AS T
         ON E.FileName == T.FileName;
```

Text Analysis

```
REFERENCE ASSEMBLY [TextCommon];
REFERENCE ASSEMBLY [TextSentiment];
REFERENCE ASSEMBLY [TextKeyPhrase];

@WarAndPeace =
    EXTRACT No int,
            Year string,
            Book string, Chapter string,
            Text string
    FROM @"/usqlext/samples/cognition/war_and_peace.csv"
    USING Extractors.Csv();

@sentiment =
    PROCESS @WarAndPeace
    PRODUCE No,
            Year,
            Book, Chapter,
            Text,
            Sentiment string,
            Conf double
    USING new Cognition.Text.SentimentAnalyzer(true);

OUTPUT @sentinment
    TO "/sentiment.tsv"
    USING Outputters.Tsv();
```

Imaging

```
REFERENCE ASSEMBLY ImageCommon;
REFERENCE ASSEMBLY FaceSdk;
REFERENCE ASSEMBLY ImageEmotion;
REFERENCE ASSEMBLY ImageTagging;
REFERENCE ASSEMBLY ImageOcr;

@imgs =
    EXTRACT FileName string, ImgData byte[]
    FROM @"/images/{FileName}.jpg"
    USING new Cognition.Vision.ImageExtractor();

// Extract the number of objects on each image and tag them
@objects =
    PROCESS @imgs
    PRODUCE FileName,
            NumObjects int,
            Tags string
    READONLY FileName
    USING new Cognition.Vision.ImageTagger();

OUTPUT @objects
    TO "/objects.tsv"
    USING Outputters.Tsv();
```