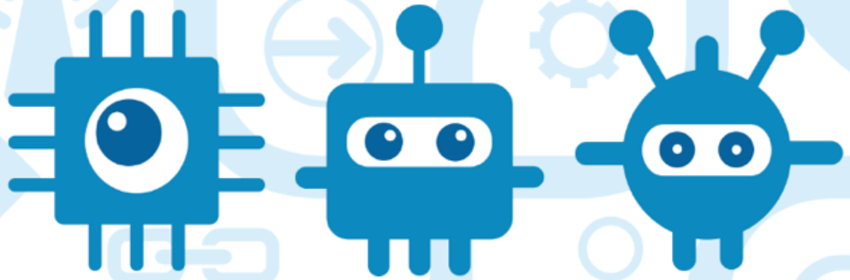
Big Data & Internet of Things

Innovate@SPEED



Big Data and the Internet of Things

Business opportunities, technologies and examples

Jukka Ruponen
Business Analytics Architect

+358-40-725-6086
jukka.ruponen@fi.ibm.com

Big data - Business point of view

Big Data

Big data is being generated by everything around us at all times. **Every digital process and social media exchange produces it. Systems, sensors and mobile devices transmit it.**

What differentiates **big data** from **traditional business data**?

- **Business data** is typically well structured, created by specifically designed business applications and processes (like CRM, ERP or HR systems), and managed in relational databases within the organization

- **Big data** is the data arriving from multiple external and/or internal sources at an alarming **Velocity**, **Volume** and **Variety**, typically not created by business applications and processes, not well organized or structured, and not stored in relational databases.

To **extract meaningful value from big data**, you need optimal **processing power**, **analytics capabilities** and **skills**.

What is changing in the realm of big data?

Big data is changing the way people within organizations work together. It is creating a new culture in which business and IT leaders must join forces to realize value from all data. Insights from big data can enable **all employees** to make **better decisions** — deepening **customer engagement**, **optimizing operations**, preventing **threats and fraud**, and capitalizing on **new sources of revenue**.

Escalating demand for insights from big data **requires a fundamentally new approaches to architecture, tools and practices.**

Competitive advantage
Data is emerging as the world's newest resource for competitive advantage.

Decision making
Decision making is moving from the elite few to the empowered many.

Value of data
As the value of data continues to grow, current systems won't keep pace.

Big Data & Analytics are making the world Smarter

Smarter Cities

Smarter Healthcare

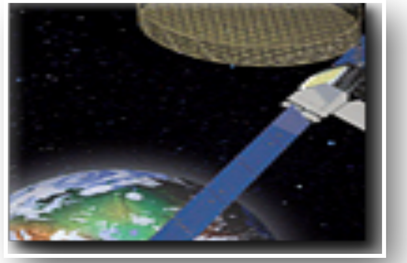
Smarter Grids

Smart Water

Smarter Traffic

Smarter Processes

Smarter Telecom



Smarter Public Safety

Smarter Finance

Smarter Retail

Smarter Living

Smarter Education

Real world big data use cases

BC Egg Marketing Board used IBM analytics capabilities to cut basic farm inspection workload by 66 percent and reduce the budget cycle from two months to less than two weeks.

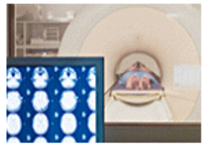
PDF (862.20KB)

Infinity Property & Casualty used predictive analytics to identify fraudulent claims, better identify subrogation potential, and improve right-tracking of claims to speed claim handling.

Video (00:03:17)

Daimler FleetBoard – Mercedes Benz uses an IBM analytics-based telematics solution to help optimize vehicle usage and routing, enable remote delivery of functionality to vehicles and lower customers' insurance premiums 10 percent.

PDF (80.85KB)

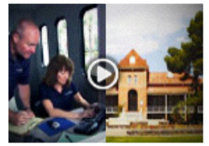


Discovery Health employed IBM predictive analytics to extract insights from clinical, demographic, billing and member data to find chronic health-risk patterns and help risk analysts improve predictive accuracy.

PDF (69.55KB)

Allianz Group implemented a centralized risk-management platform with IBM advanced analytics, gaining greater visibility into its equity basis and accelerating its economic portfolio reporting from six weeks to one day.

PDF (79.35KB)

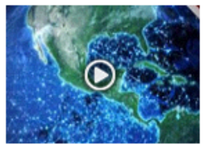


Security First Insurance employed the content analytics capabilities of IBM Watson Foundations on social media posts to boost productivity and customer satisfaction by bridging social media and the claims process.

PDF (88.18KB)

Video (00:04:22)

Visit the web page

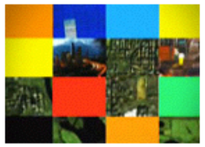


T-Mobile uses big data and analytics to gain insight into data flowing through its entire network in seconds, and enhance the veracity of that data to above 99.99 percent.

Video (00:03:10)

Sprint Velocity stores driver preferences in the cloud so drivers can use their smartphones in their vehicle compartments to create a more personalized experience.

Video (00:00:40)



Santam Insurance used IBM predictive analytics to improve claims categorization, reducing the time and cost of settling cases and processing 15% of claims within an hour —a 95% time savings.

Visit the web page

Westfield Insurance applied business intelligence and analytics to gain greater insights, make good risk selection, service customers, market and sell products, and differentiate the company.

PDF (932.04KB)

Video (00:04:01)

Insurance Bureau of Canada engaged IBM to perform a proof-of-concept project that successfully automated the detection of potential claim fraud and the identification of possible fraud rings.

PDF (695.68KB)

First Tennessee Bank increased marketing response rate by 3.1 percent by using predictive analytics to more accurately target offers to high-value customers.

PDF (357.75KB)

Video (00:04:12)

Pioneer West Virginia Credit Union reduced its loan delinquency ratio by 95% in one year by gaining the insights to understand operations in a near-real-time environment.

PDF (820.49KB)

Video (00:04:28)

Barclays used advanced analytics in a collaborative forum to reveal hot topics and sentiments in real time, helping ensure its global workforce aligned with the bank's customer-centric strategy.

PDF (932.83KB)

C Spire Wireless used IBM predictive analytics on customer-behavior and account-profile data to optimize its service mix and offers, increasing the efficacy of its customer retention campaigns by half.

PDF (77.68KB)

TEOCO enabled CSPs to access and analyze massive amounts of big data to uncover the source of cost and network issues, a move that is helping providers save millions of dollars.

PDF (746.6KB)

Real world big data use cases

Becker Underwood used IBM analytics to integrate and optimize its global agriculture chemicals supply chain, greatly improving its inventory turn and forecasting accuracy to sustain rapid business growth.

Video (00:03:45)

Recology used big data and analytics to improve operational efficiency, transporting waste to collection facilities more efficiently and reducing the amount of trash sent to the landfill by 50%.

Video (00:02:53)

Mueller, Inc. used IBM business intelligence analytics to measure and improve sales performance, create effective business plans, and gain deep insight into patterns and trends in its data.

PDF (628.67KB)

Jabil Circuit built a comprehensive suite of analytical tools that made its finance department more efficient and effective, reducing time to close and manual activity by 50 percent.

PDF (599.65KB)

Video (00:05:22)

City of Toulouse, France, uses IBM social media analytics to gain insights into citizens' needs from social media posts, helping the city identify and prioritize citizens' most pressing issues.

PDF (78.62KB)

Miami-Dade Police Department used advanced analytics to uncover insights in cold robbery cases, more quickly identifying suspects, making arrests and reducing repeat crimes.

PDF (2.19MB)

Video (00:03:32)

Memorial Healthcare System uses IBM content analytics to cull insights from vendor data, reducing invoice cycles by about 40 percent and revealing vendor relationship problems to mitigate fraud risks.

PDF (1.30MB)

Video (00:04:26)

Seton Healthcare Family minimized readmission of patients who had suffered congestive heart failure by uncovering insights trapped in unstructured data to identify at-risk patients.

Video (00:03:01)

Regina Police Service of Saskatchewan relies on sophisticated content analytics tools from IBM to identify and remove inappropriate content from social media sites, making them safe for the public.

PDF (750.75KB)

Durham, NC, Police Department reduced the amount of violent crime in a 2-square mile high-crime region of the city by over 50 percent from 2007 through 2011 with IBM analytics technology.

Video (00:03:20)

Ishinomaki City used analytics on fish species and location data to restore the local fishing industry, hurt by the 2011 earthquake and subsequent nuclear accident.

PDF (80.31KB)

Chickasaw Nation Division of Commerce used IBM predictive and patron analytics on its gaming operations, promotions and patron data to reduce its month-end close process by 50 percent and provide a better gaming experience.

Video (00:05:45)

Singapore Land Transit Authority used predictive tools and smart cards to provide citizens with a more convenient, smart public transportation system that helps reduce traffic and pollution.

Video (00:04:09)

State of North Carolina identified nearly USD200 million in suspicious Medicaid claims by using analytics to sort through and prioritize tens of thousands of providers and hundreds of millions of claims in minutes.

Video (00:02:13)

SHOP.CA, the largest e-commerce marketplace in Canada, uses IBM advanced analytics solutions to improve the search experience, target personalized emails and improve click-through rates 300 to 500 percent.

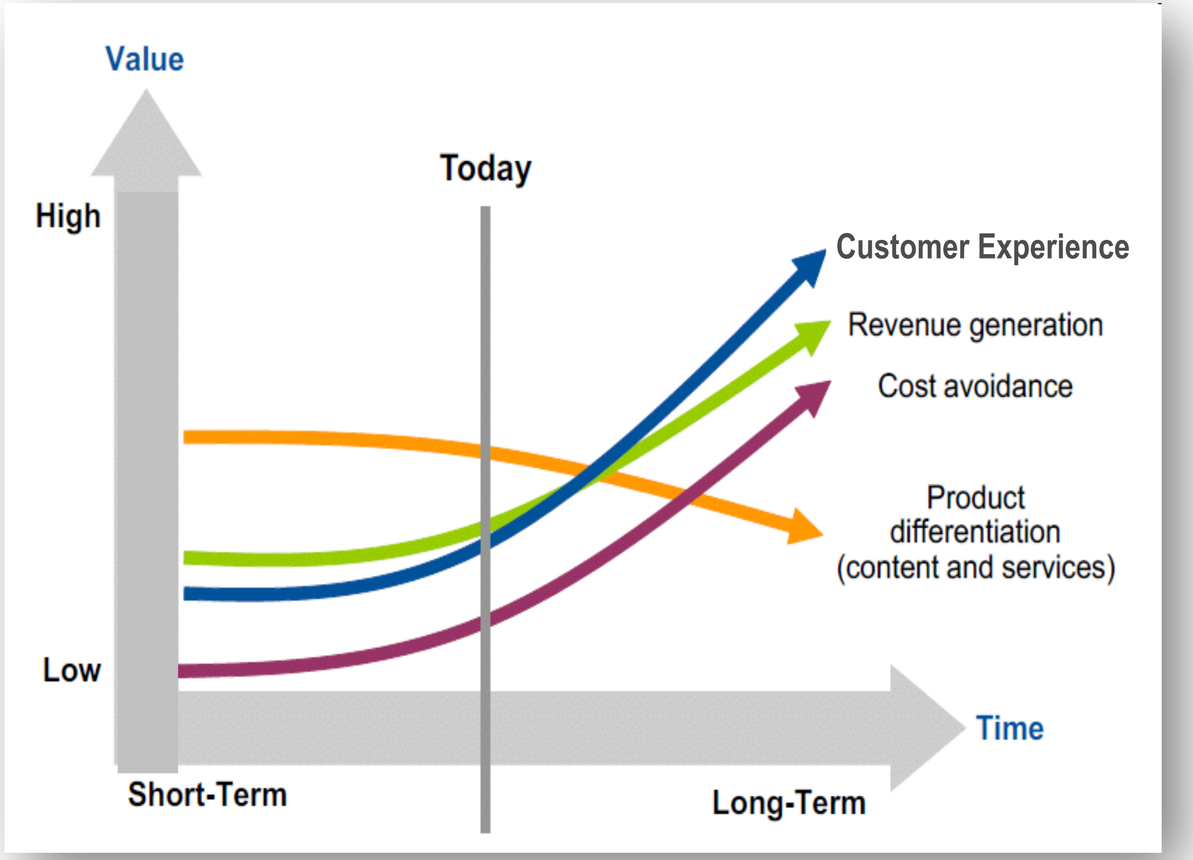
PDF (146.78KB)

Luxottica used IBM Big Data & Analytics capabilities and a solution from Aginity to identify their highest-value customers from nearly 100 million and create personalized marketing campaigns.

PDF (84.34KB)

The Real Value of Big Data

The benefits of connected people and devices will change over time, as connectivity becomes commodity and through higher experience & demand of the real value



Customer Experience

Improved, real-time customer interaction

New Revenue generation

Revenue from new applications and services

Cost avoidance / productivity

Reductions in warranty related costs (e.g. recalls) due to remote diagnostics

Product differentiation

Differentiation opportunities from content & service perspective (functionality)

How can businesses realize the value of big data?

New skills are needed to fully harness the power of big data. Though courses are being offered to prepare a new generation of big data experts, it will take some time to get them into the workforce. Meanwhile, leading organizations are developing new roles, focusing on key challenges and creating new business models to gain the most from big data.

Discovering the new role of data scientist

Gartner finds that by end 2015, the demand for data and analytics resources will reach 4.4 million jobs globally, but only one-third of those jobs will be filled. The emerging role of data scientist is meant to fill that skills gap.

Being proactive about privacy, security and governance

While big data can provide significant value, it also presents significant risk. Organizations must be proactive about privacy, security and governance to ensure all data and insights are protected and secure.

Create new business models with big data

From data-driven marketing and ad-targeting to the connected car and smart buildings, big data is fueling product innovation and new revenue opportunities for many organizations.

IBM

Questions to You

- What are the charasteristics of big data?
- What is the value of big data?
- What are the challenges with big data?

Big data - Information systems point of view

But first, some key technologies you'll need to understand…

Relational Database (“SQL data store”)

Relational databases stores data in **series of tables** (relations) with **pre-defined data structure** (attributes) and **well defined data types** (type of value). In other words, rows in tables define **records** and columns in tables define **fields** (or attributes) for that record. As result, each record in a table must share the same set of attributes. Another feature of relational database is “**relationships**” between the tables, meaning that records in different tables can be linked together by matching values in key columns (e.g primary keys).

Well-defined **Structured Query Language** (SQL) is used to **maintain and query** the database. Database **functions** and **procedures** are used to perform processes and calculations on data.

These well-defined charasteristics of relational databases makes them well suitable for storing and processing **numeric data** and allow them to be optimized in number of different ways to increase performance.

Relational databases are **not** well suitable, however, to store unstructured information (such as text, audio, video), large amounts of data (petabytes) or data that requires flexibility in it’s structure (like generic sensor data).

Examples: IBM DB2, Oracle, Microsoft SQL, mySQL

Hypothetical Relational Database Model

PubID	Publisher	PubAddress
03-4472822	Random House	123 4th Street, New York
04-7733903	Wiley and Sons	45 Lincoln Blvd, Chicago
03-4859223	O'Reilly Press	77 Boston Ave, Cambridge
03-3920886	City Lights Books	99 Market, San Francisco

Primary Key

Primary Key

AuthorID	AuthorName	AuthorBDay
345-28-2938	Haile Selassie	14-Aug-92
392-48-9965	Joe Blow	14-Mar-15
454-22-4012	Sally Hemmings	12-Sept-70
663-59-1254	Hannah Arendt	12-Mar-06

ISBN	AuthorID	PubID	Date	Title
1-34532-482-1	345-28-2938	03-4472822	1990	Cold Fusion for Dummies
1-38482-995-1	392-48-9965	04-7733903	1985	Macrame and Straw Tying
2-35921-499-4	454-22-4012	03-4859223	1952	Fluid Dynamics of Aquaducts
1-38278-293-4	663-59-1254	03-3920886	1967	Beads, Baskets & Revolution

Primary Key Foreign Key Foreign Key

SELECT B.Title, A.AuthorName from BOOKS B, AUTHORS A where PUBLISHERS.PubID in ’03-4859223’

Title	AuthorName
Fluid Dynamics of Aquaducts	Sally Hemmings

Key-Value Database (“noSQL” data store)

A key-value store, or key-value database, is a data storage paradigm **designed for storing, retrieving, and managing associative arrays**, a data structure more commonly known today as a **dictionary** or **hash**. Dictionaries contain a **collection of objects**, or records, which in turn **have many different fields** within them, each containing data. These records are stored and retrieved using a **key that uniquely identifies the record**, and is used to quickly find the data within the database.

Key-value systems treat the data as a single opaque collection which **may have different fields for every record**. This offers considerable **flexibility** and more closely follows modern concepts like **object-oriented programming**. Because optional values are not represented by “placeholders” in the structure, as in most RDBs, key-value stores often use far **less memory to store the same database**, which can lead to large performance gains in certain workloads. Data is maintained and queried by applications using the database programming interfaces, e.g REST API (http).

Key-value stores are **not** well suitable for storing data that requires high performance in numeric calculations and, due to lack of pre-defined structure, they tend to duplicate data.

Examples: Cloudant, MongoDb

e.g JSON format:

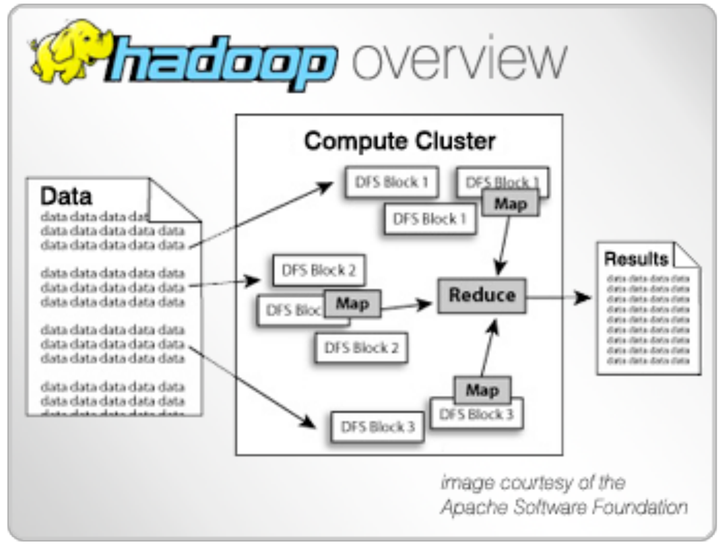
```
{ “ISBN” : ”2-35921-499-4”,
  “Title” : ”Fluid Dynamics of Aquaducts”,
  “Date” : 1952,
  “Publisher” :  [
   { “PubID” : ”03-4859223”,
     “PublisherName” : ”O’Reilly Press”,
     “PubAddress” : ”77 Boston Ave, Cambridge”}],
   “Author” : [
    { “AuthorID” : ”454-22-4012”,
      “AuthorName” : ”Sally Hemmings”,
      “AuthorBDay” : ”12-Sept-70”}]
}
{ “ISBN” : ”1-34532-482-1”,
  “Title” : ”Cold Fusion for Dummies”,
  “Revision” : 2,
  “Date” : 1990,
  “Publisher” :  [
   { “PubID” : ”03-4472822”,
     “PublisherName” : ”Random House”,
     “PubAddress” : ”123 4th Street, New York”}],
   “Author” : [
    { “AuthorID” : ”345-28-2938”,
      “AuthorName” : ”Haile Selassie”,
      “AuthorBDay” : ”14-Aug-92”,
      “AuthorTitle” : “Mr”}]
}
```

Hadoop and MapReduce

Hadoop is an open source Apache project and framework technology for **distributed computing that addresses data intensive analytic needs in big data domain.**

Allows storing of **vast amounts of data** in **any format** on a **distributed Hadoop File System** (HDFS), to be processed with **massively parallel MapReduce programming model** for enhanced analytic review.

Can be **integrated with other technologies** which enhance gathering, analysis, text analytics and visualization of massive data sets



Examples:
Apache Hadoop,
IBM BigInsights,
Hortonworks,
Cloudera, MapR

Traditional serial processing (single file):

Count the number of words in the following paragraph.
Make sure you do not count "Stop Words" (the, a, is)

file: *The quick brown fox jumps over the lazy dog*
A dog is a man's best friend

quick => 1
brown => 1
fox => 1
jumps => 1
over => 1
lazy => 1
dog => 2
mans => 1
best => 1
friend => 1

MapReduce parallel processing (multiply sequences of the file):

seq 1: *The quick brown fox jumps over the lazy dog*

seq 2: *A dog is a man's best friend*

Map 1

quick => 1
brown => 1
fox => 1
jumps => 1
over => 1
lazy => 1
dog => 1

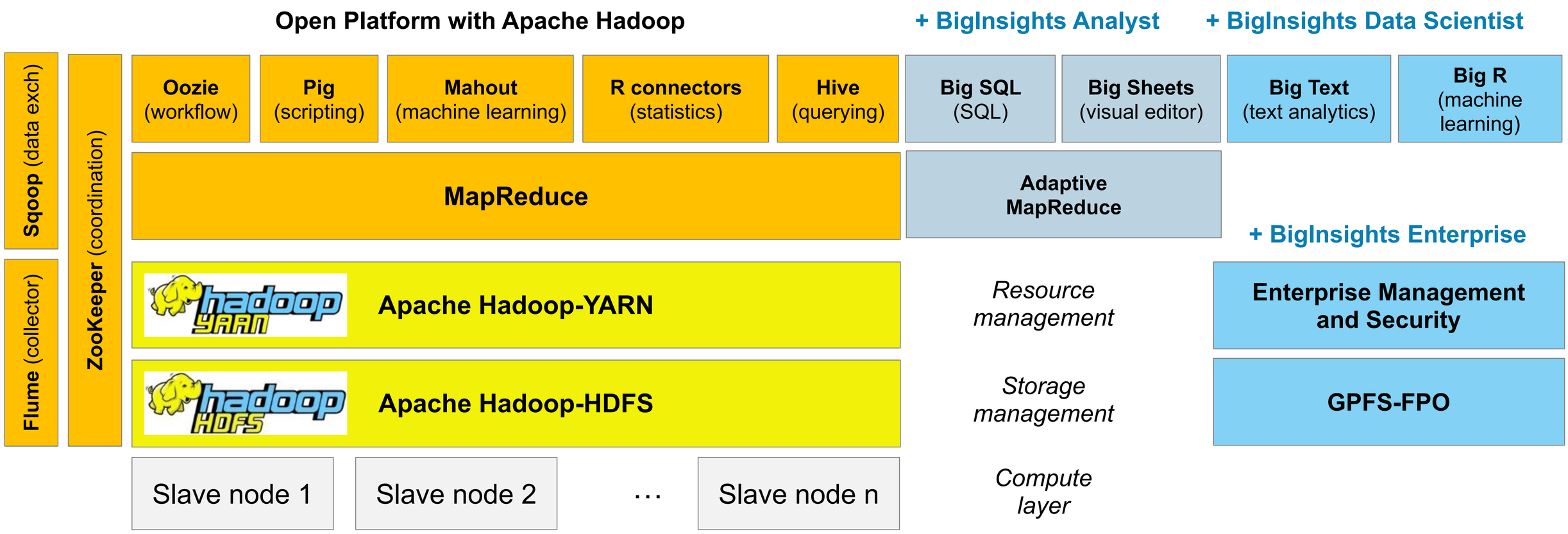
Map 2

dog => 1
man's => 1
best => 1
friend => 1

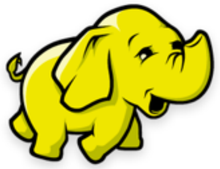
quick => 1
brown => 1
fox => 1
jumps => 1
over => 1
lazy => 1
dog => 2
mans => 1
best => 1
friend => 1

Sources:
http://www-01.ibm.com/software/ebusiness/jstart/hadoop/
http://www.techspot.co.in/2011/07/mapreduce-for-dummies.html

MapReduce on top of Hadoop



Hadoop Advantages



Unlimited Scale

- Multiple data sources
- Multiple applications
- Multiple users

- Reliability
- Resiliency
- Security

Enterprise Platform

Wide Range of Data Formats

- Files
- Semi-structured
- Databases

MapReduce Challenges

- Need deep Java skills
- Few abstractions available for analysts

Ease of Development

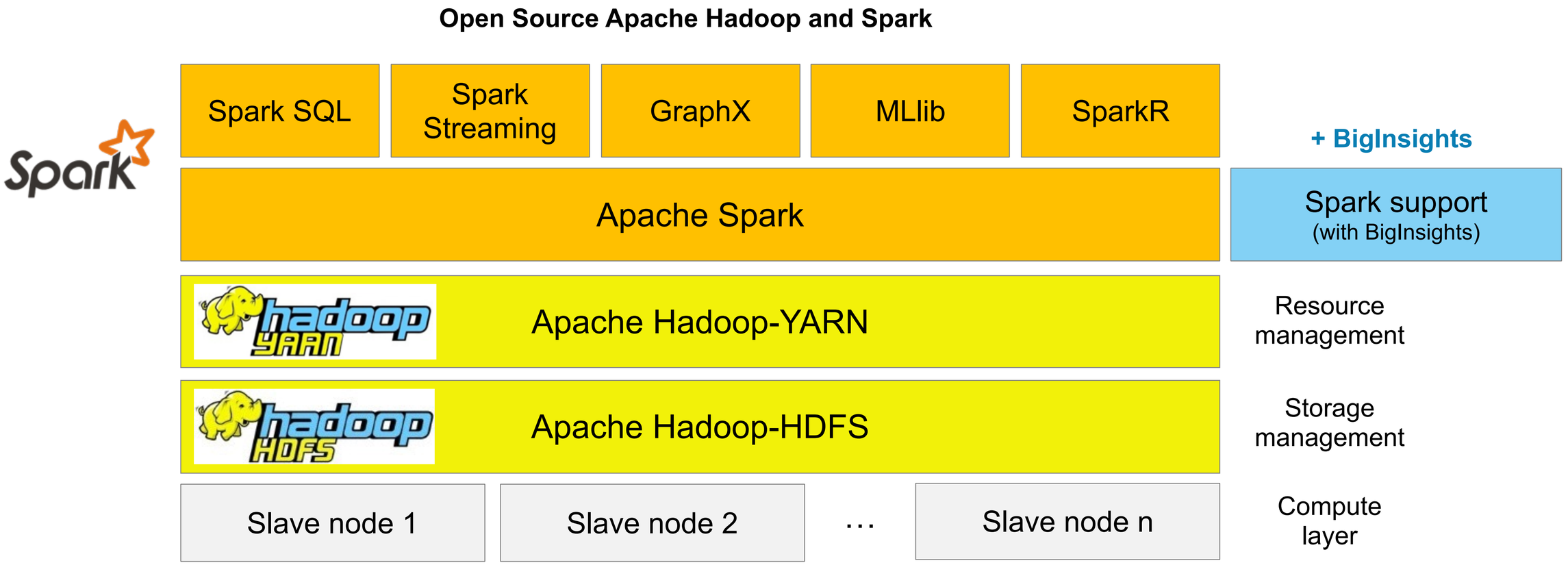
Disk-based Performance

- No in-memory framework
- Application tasks write to disk with each cycle

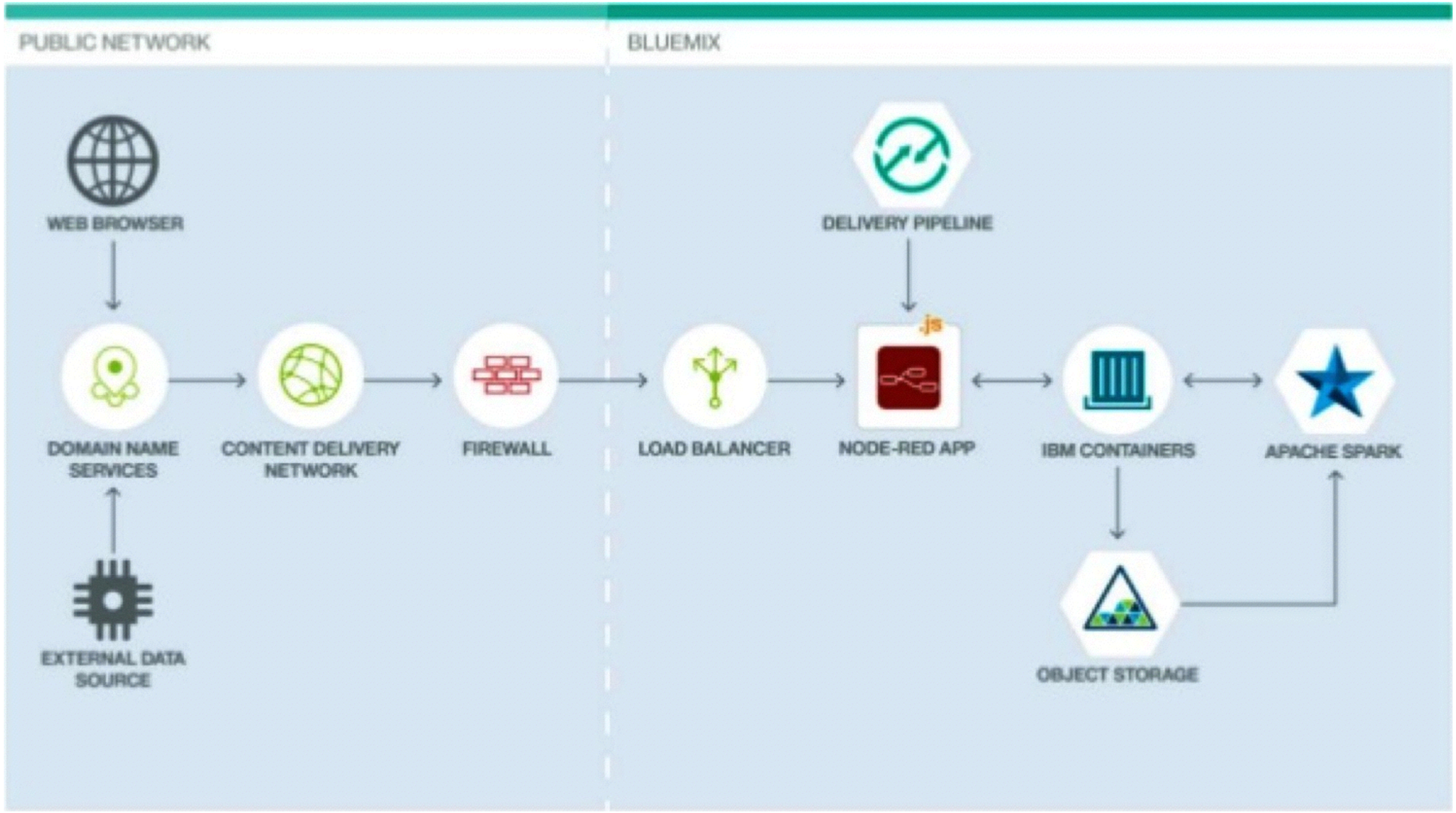
- Only suitable for batch workloads
- Rigid processing model

Combine Workflows

Spark on Hadoop



Apache Spark in IBM Bluemix cloud



Apache Spark Advantages

Ease of Development
- Easier APIs
- Python, Scala, Java

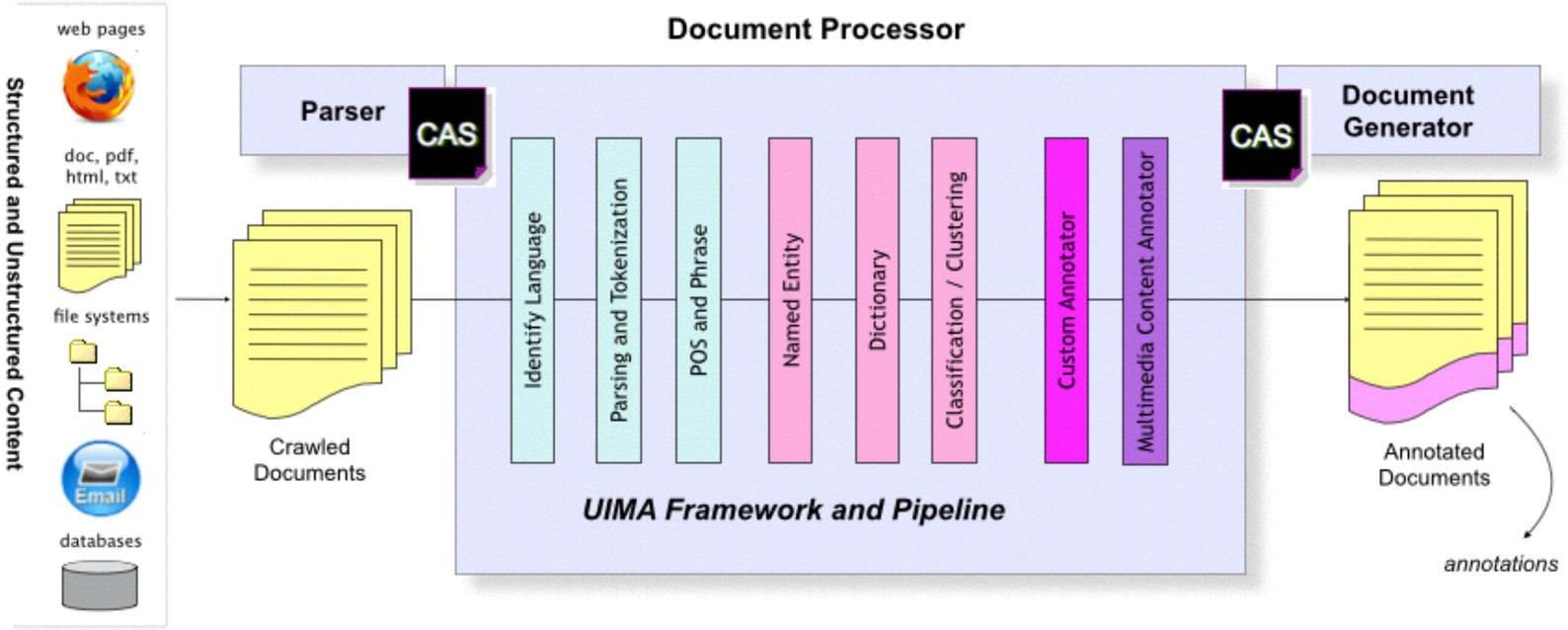
In-Memory Performance
- Resilient Distributed Datasets
- Unify processing
- 10 to 100 times faster than MapReduce

Combine Workflows
- Batch
- Interactive
- Iterative algorithms
- Micro-batch

Unstructured Information Management Architecture (UIMA)

- UIMA is an **open standard (OASIS) to manage unstructured data**.
- Typical use case for UIMA is **text analytics**, i.e to create a structure upon an unstructured information.
- **Annotators** are developed to process the information.
- Information sources (and targets) can be relational databases, noSQL datastores, HDFS etc.

- Examples:
  Apache UIMA
  IBM Watson Content Analytics



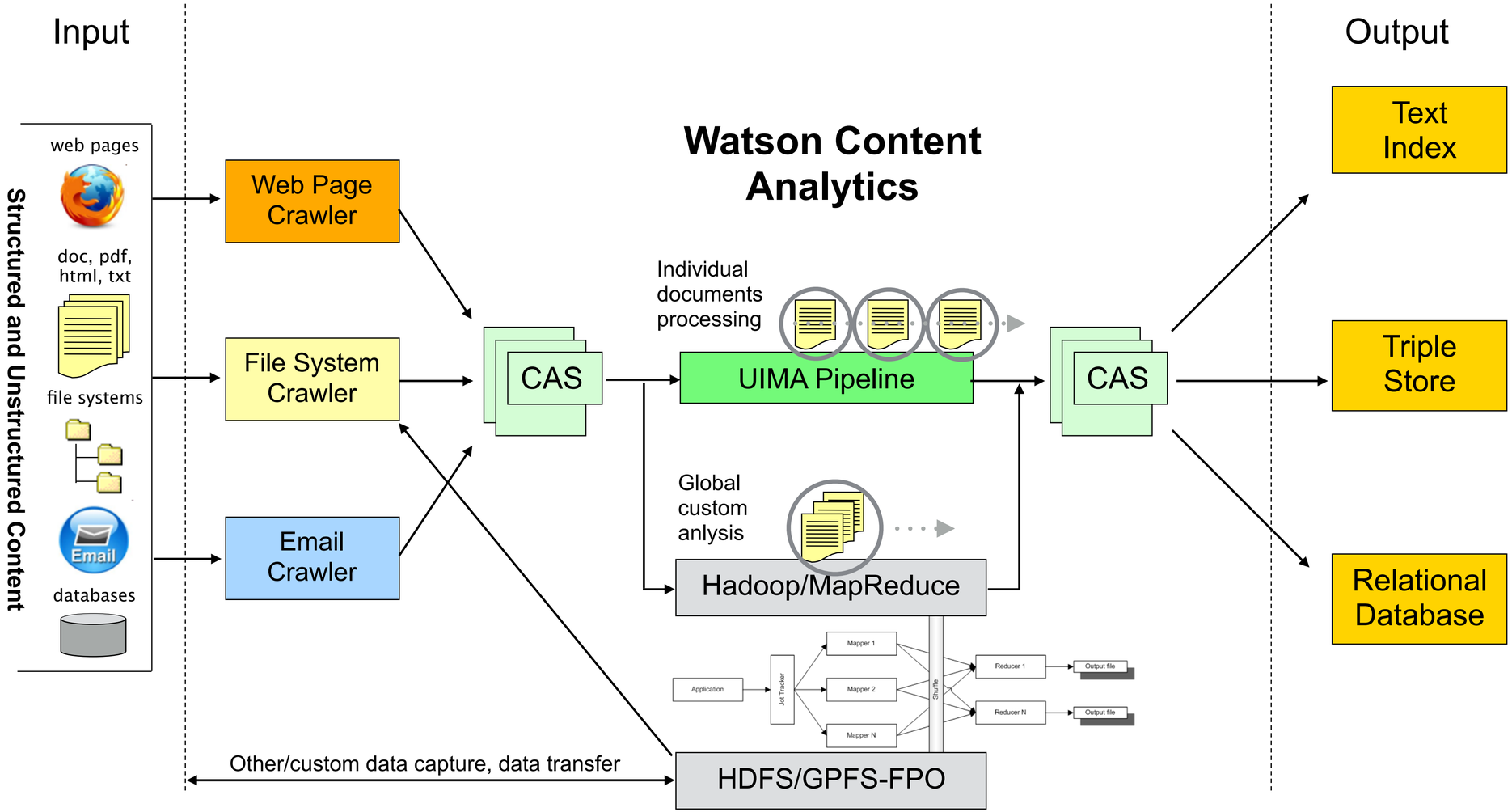
Inbound Unstructured Data

PC 143 (Hunter)
15 June 2006 23:47
Suspect identified himself as John Setsuko. Matched description given by night club doorman (IC1, Male, Ag 22-24 yrs, blue Everton shirt). Stopped whilst driving White Ford Mondeo, W563 WDL. Address given as 22 East Dene Ridge, Copdock, Ipswich. Searched at scene and found in possession of 1oz Cannabis Resin and lockable pocket knife.

Extracted / Derived Information

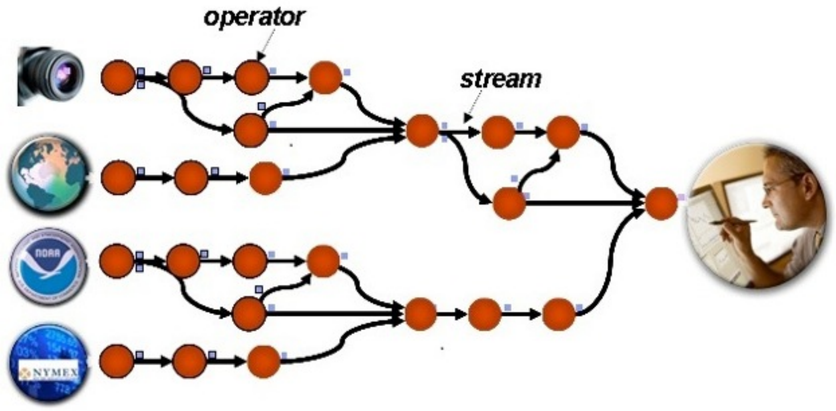
Arresting_Officer	PC 143
Arrest_Date_Time	15/06/2006 : 23:47
Suspect_Forename	John
Suspect_Surname	Setsuko
Suspect_VRN	W563WDL
Suspect_Vehicle_Colour	White
Suspect_Vehicle_Make	Ford Mondeo
Suspect_Addr_Street	22 East Dene Ridge
Suspect_Addr_Town	Ipswich
Evidence_1_Description	1 oz Cannabis Resin
Evidence_2_Description	Lockable pocket knife

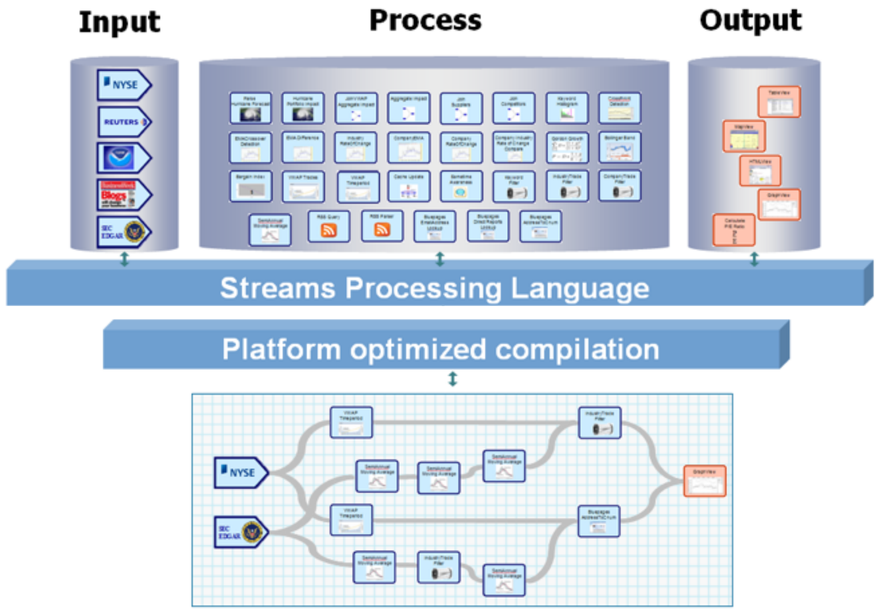
UIMA with Hadoop



Stream Computing

- Stream Computing (or Realtime Analytic Processing) is a technology to **continuously analyze massive volumes of streaming data in near real-time**
- In addition to capture streams of data from network, data sources can also include persistent data stores, such as relational databases, noSQL databases, Hadoop etc.
- A stream computing system is typically capable of receiving **hundreds of thousands transactions per second**, and can leverage **sub-millisecond latencies** in analytics to immediately react upon events and trends
- **Can perform complex but fast analytics of heterogeneous data types** including text, images, audio, voice, VoIP, video, police scanners, web traffic, email, GPS data, financial transaction data, satellite data, sensors, and any other type of digital information
- **Can adapt to rapidly changing data forms and types**.
- Examples: IBM Streams





Predictive Analytics and Stream Computing in Healthcare

Capabilities:

- IBM Stream Computing
- IBM SPSS Predictive Analytics

Real-time analytics and correlations on physiological data streams

Heart Rate deviation, Blood pressure, Temperature, EKG, Blood oxygen saturation etc.,

Improvements:

Early detection of the onset of potentially life-threatening conditions, up to 24 hours earlier than current medical practices

Early intervention leads to lower patient morbidity and better long term outcomes

Technology also enables physicians to verify new clinical hypotheses

Asian telco reduces billing costs and improves customer satisfaction

Capabilities:

IBM Stream Computing
Analytic Accelerators

Real-time mediation and analysis of
6B CDRs per day

Improvements:

Data processing time reduced from
12 hrs to 1 sec

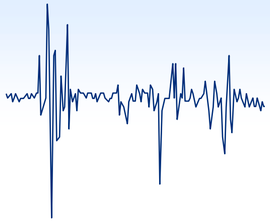
Hardware cost reduced to 1/8th

Proactively address issues (e.g. dropped calls) impacting customer satisfaction.

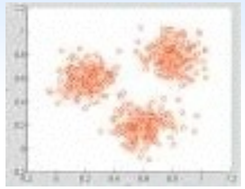
Streams Analytic Accelerators

Text
(listen, verb),
(radio, noun)

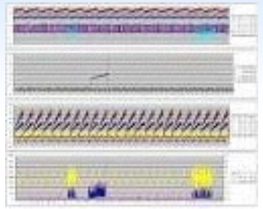
Simple & Advanced Text Analytics



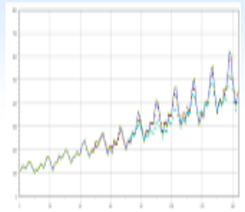
Acoustic



Mining in Microseconds



Advanced Mathematical Models



Predictive

$$\sum_{population} R(s_t, a_t)$$

Statistics

Geospatial

Image & Video

Information management architectures for big data …

Systems of Insight create New Value

Any Data

New/Enhanced Applications

Systems of Record

Managing enterprise transactions and data

Systems of Insight

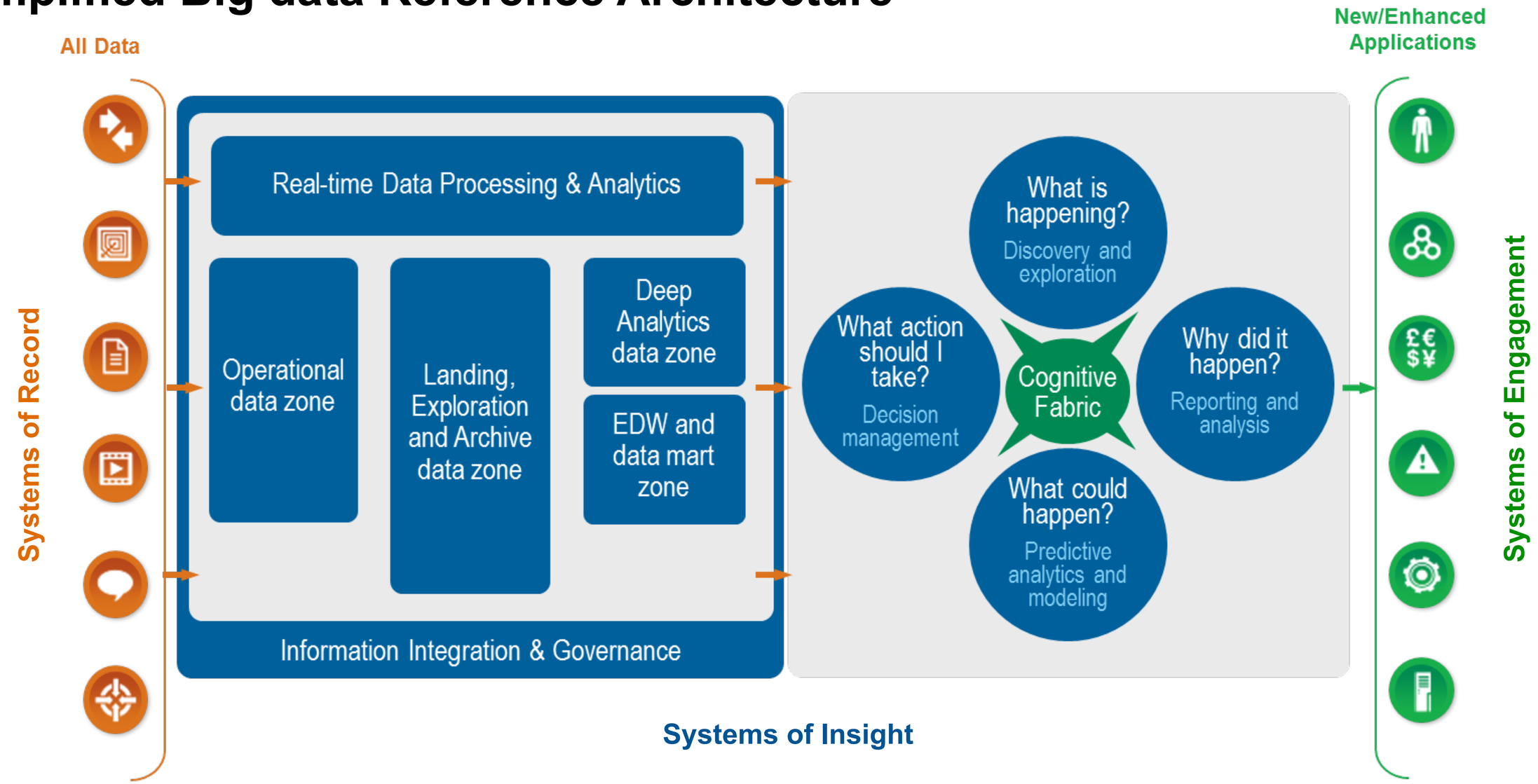
- Drive every decision
- Fuel every interaction
- Power every process

Using analytics to create new insights

Systems of Engagement

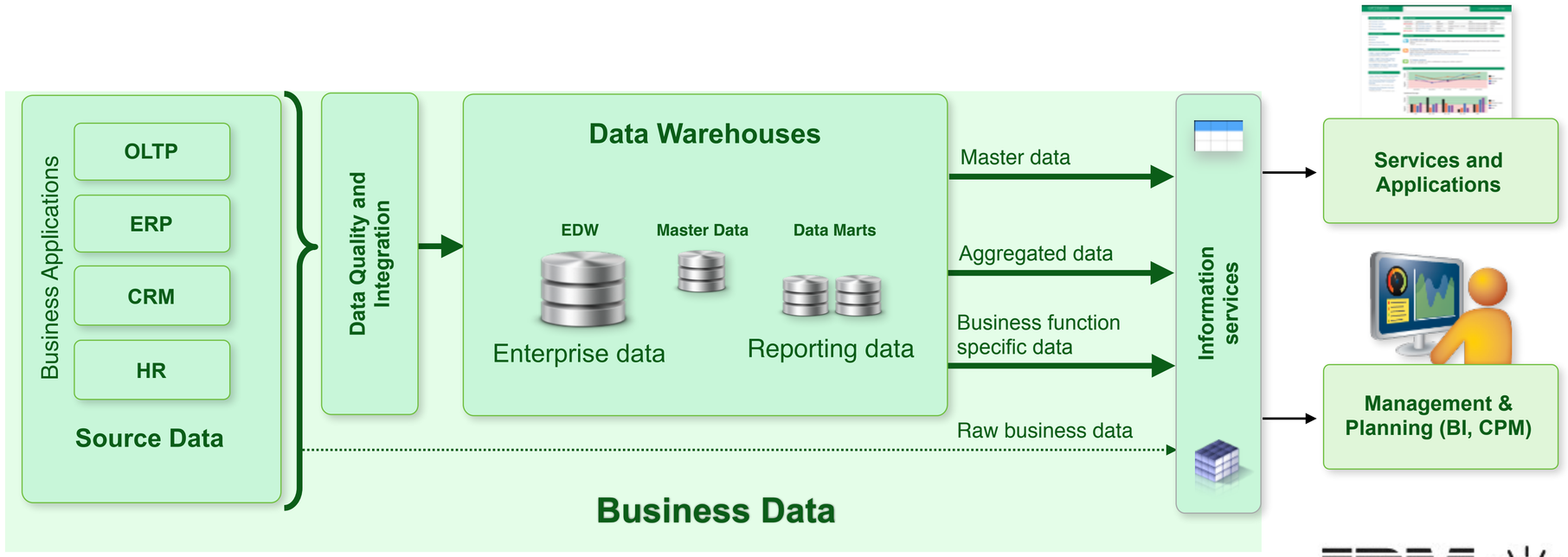
Providing user experience, engaging customers/employees

Simplified Big data Reference Architecture



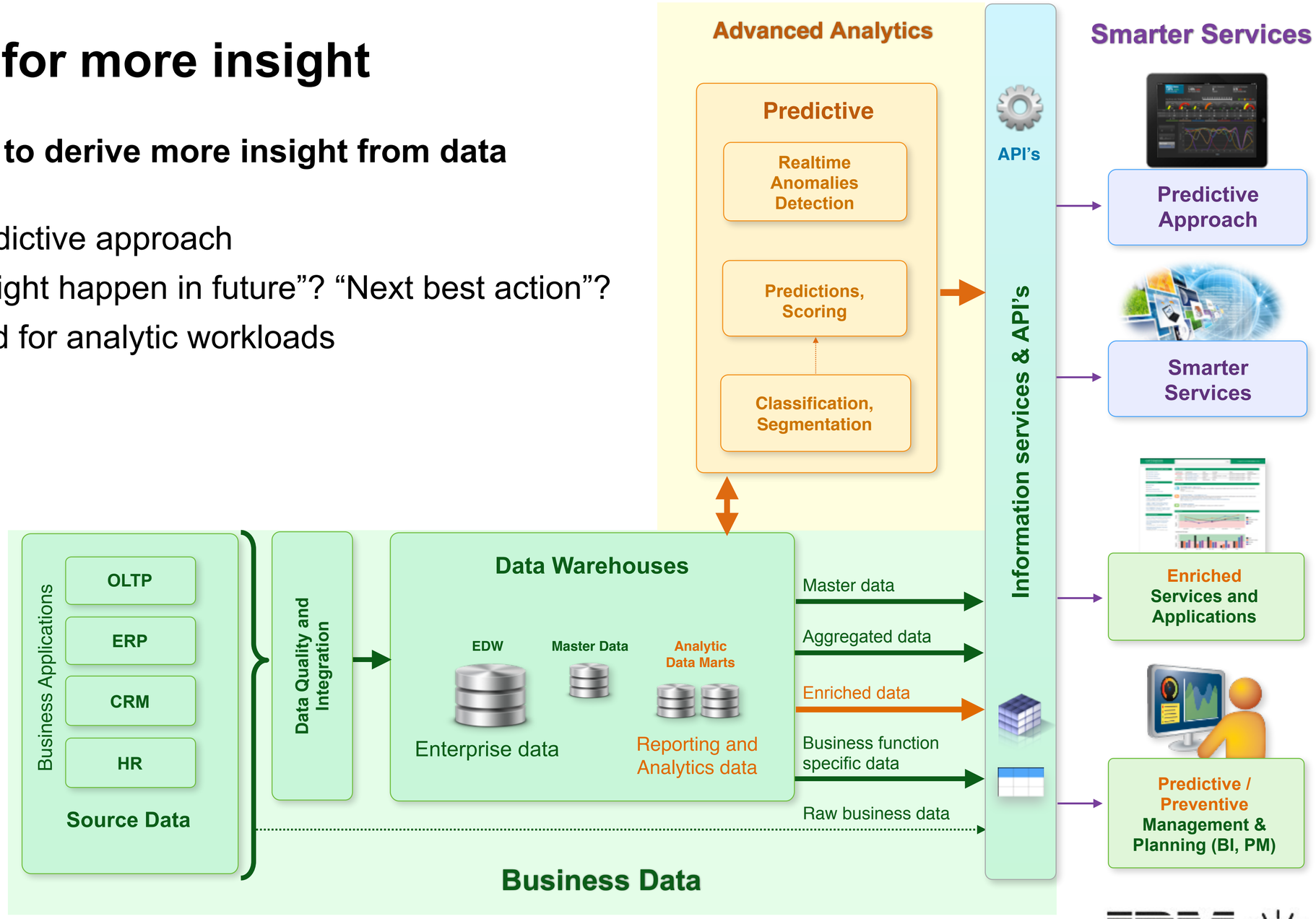
1. Traditional business information systems

- **The focus is to deliver application independent, trusted information to the business** (decision making process)
- **DW**: Based on well structured data, organized in Relational Databases (SQL-based)
- **BI**: “How are we doing”? “Why is it so”? “What should we do about it”?



2. Adding analytics for more insight

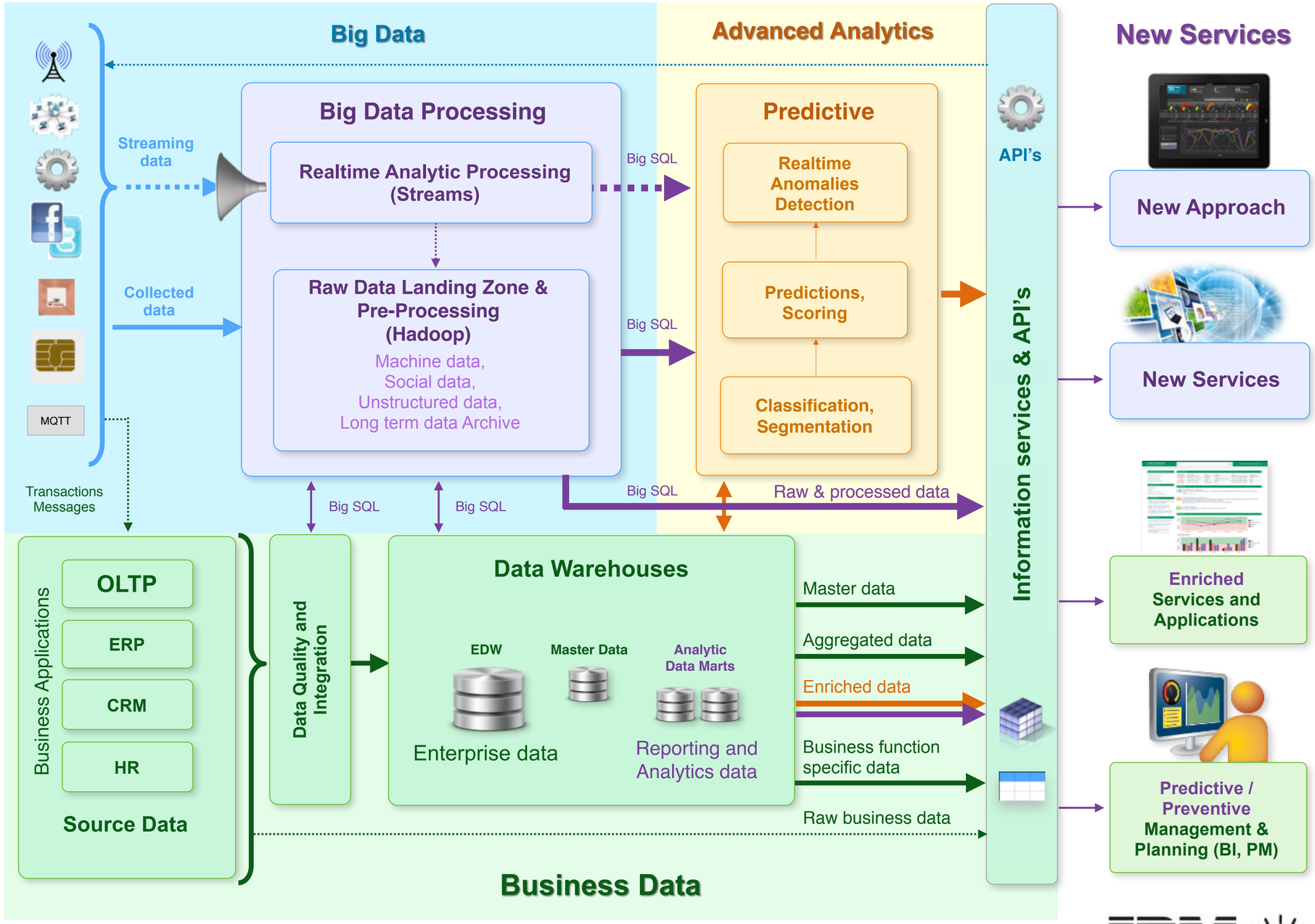
- **The focus is on being able to derive more insight from data** (business optimization)
- Enrich existing data with predictive approach
- **PA**: “More insight”? “What might happen in future”? “Next best action”?
- **DW**: Increase power & speed for analytic workloads



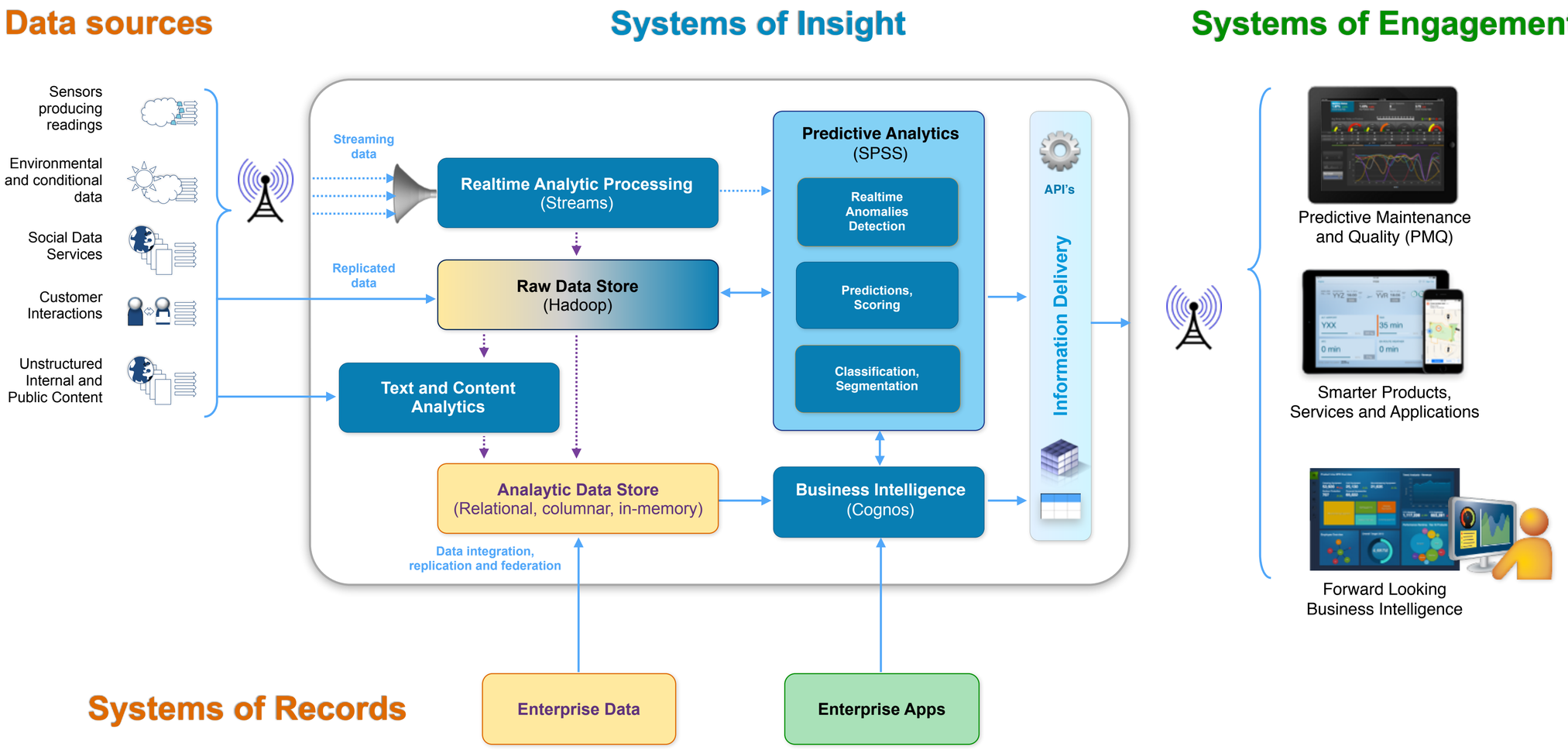
3. Adding new data sources

- The focus is on being able to **analyze Any/All data** and **take big data approach** in order to differentiate, stay ahead and create new business models (business innovation)

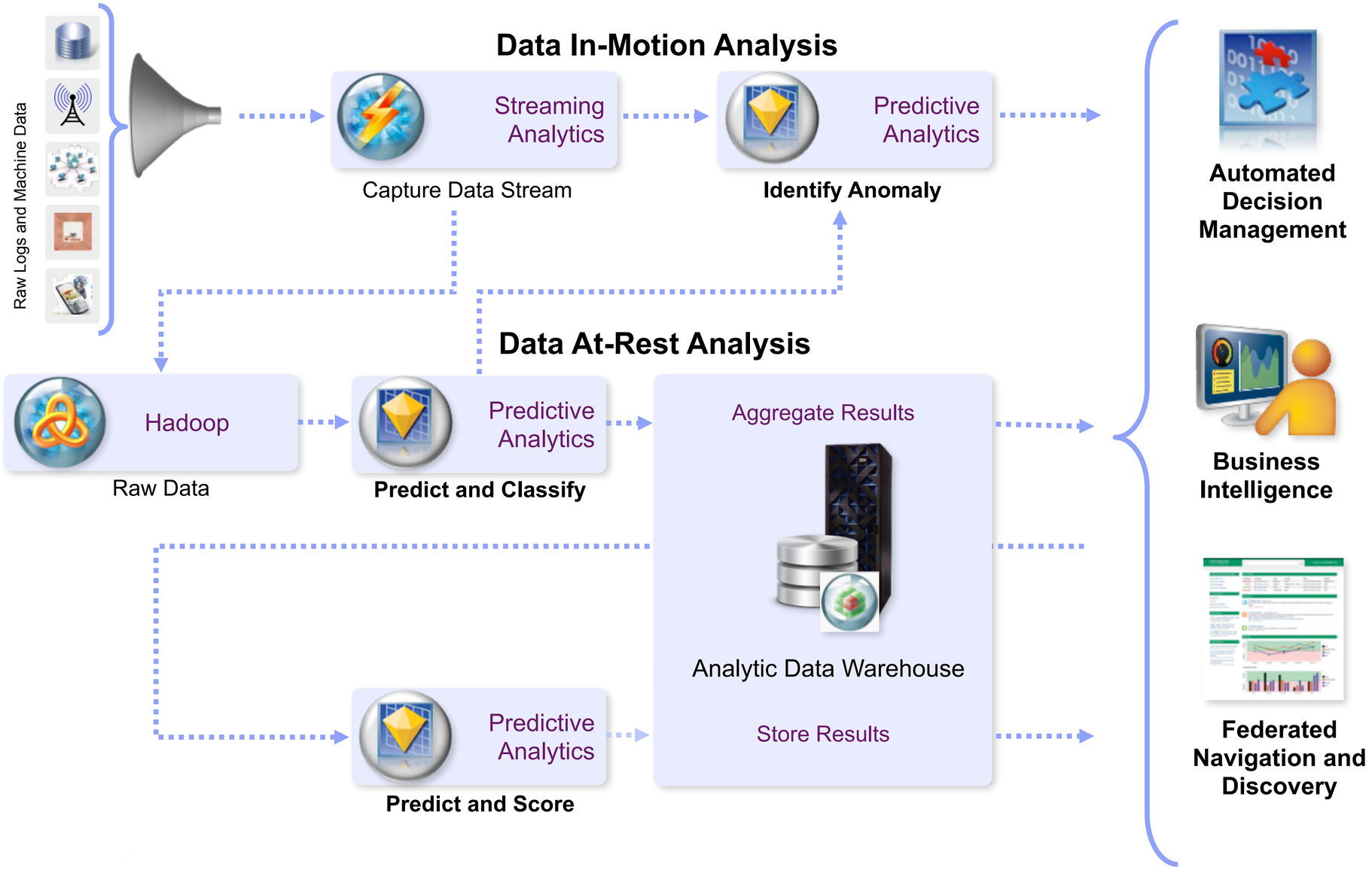
- **DW & Streams**: Dynamic data warehousing with **added noSQL and Streaming data processing capabilities** to handle **Volume**, **Variety** and **Velocity** of data



Example of operational analytics, supporting big data sources



Example data flow



Questions to You

- How can you analyze big data? What kinds of data?
- Why can’t traditional business information systems handle big data very well?
- What kind of architectural changes (or technologies) are needed in business information systems in order to start leveraging big data?

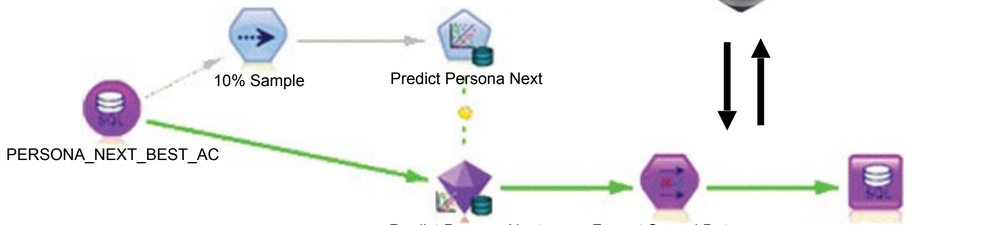
Two examples of business solutions, leveraging big data

Predictive Customer Intelligence (PCI)

- Analyze customer footprints – purchase history, in-store browsing, clickstream data, social media posts and other digital exhaust – to create a three dimensional customer view.

"THINGS"

Realtime location, condition or event

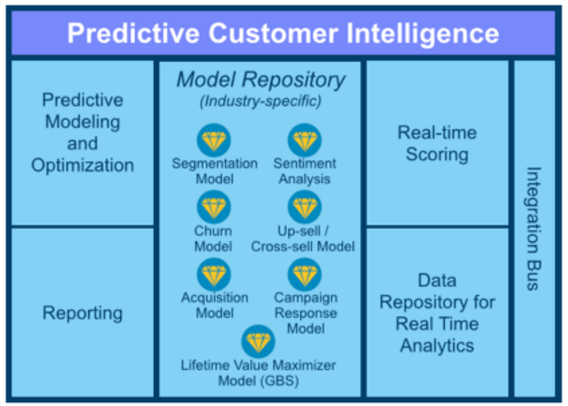


Targeted campaigns and actions

Shopping & online browsing history

- Previous product searches
- React upon offers
- Activity
- Personal preferences and habits

"System of Insight"



Product availability, current campaigns etc

Predictive Customer Intelligence (PCI)

Big Data Sources

WHAT?
Behavioral data
- Orders
- Transactions
- Payment history
- Usage history

WHO?
Descriptive data
- Attributes
- Characteristics
- Self-declared information
- Geographic demographics

HOW?
Interaction data
- Email & chat transcriptions
- Call center notes
- Web clickstreams
- In-person dialogues

WHY?
Attitudinal data
- Opinions
- Preferences
- Needs and desires

Predictive Customer Intelligence

- Acquisition models
- Campaign response models
- Churn models
- Customer lifetime value
- Market basket analysis
- Price sensitivity
- Product affinity models
- Segmentation models
- Sentiment models
- Up-sell / Cross-sell models
- Predictive Customer Intelligence available both inbound (real-time) and outbound (batch)

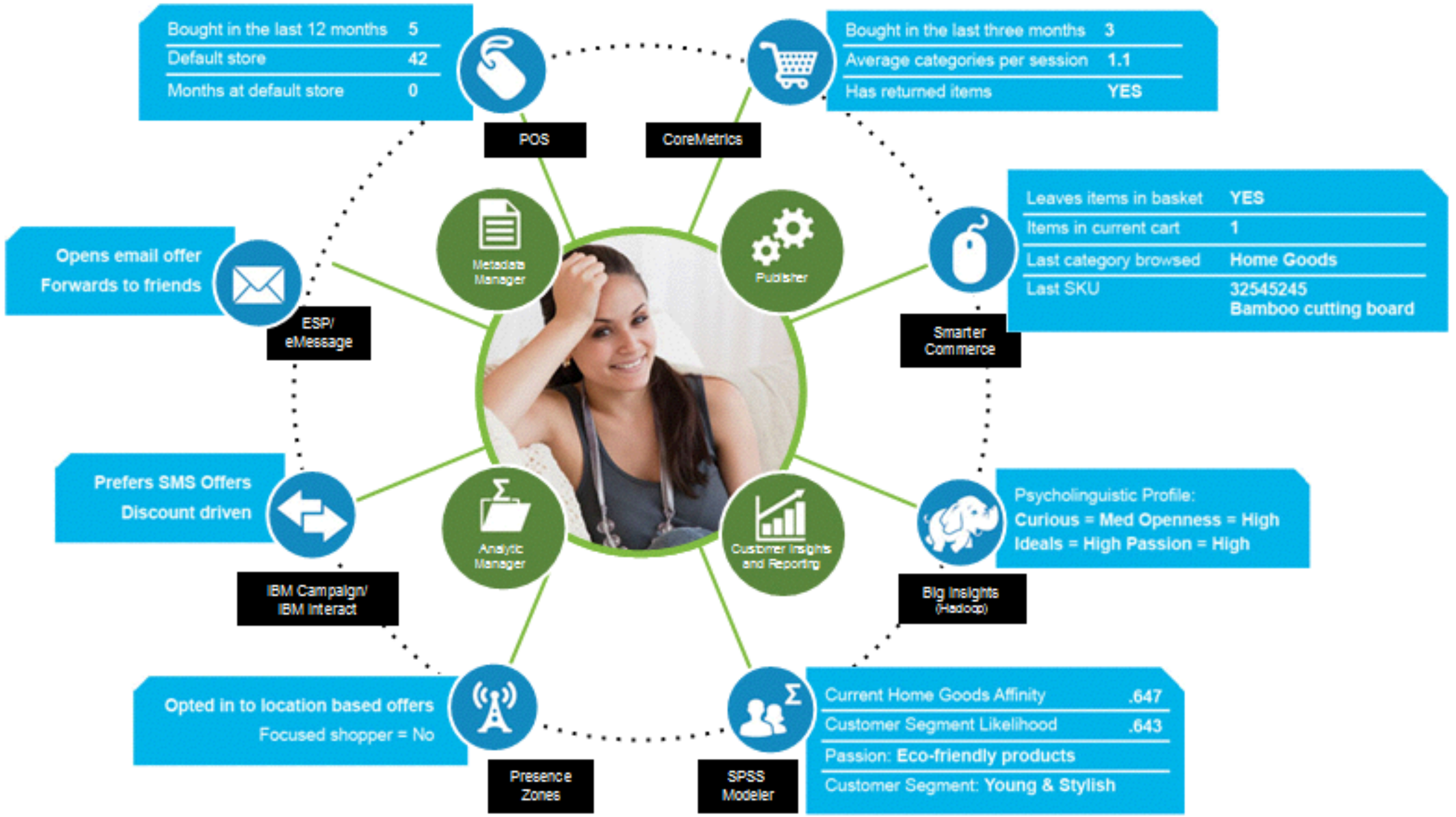
Enterprise Marketing

- Campaigns
- Offers
- Messaging
- Lead Management
- Cross-channel Campaign Mgmt
- Real-time Marketing
- Marketing Event Detection
- Digital Marketing

Multi-channel Customer Interactions

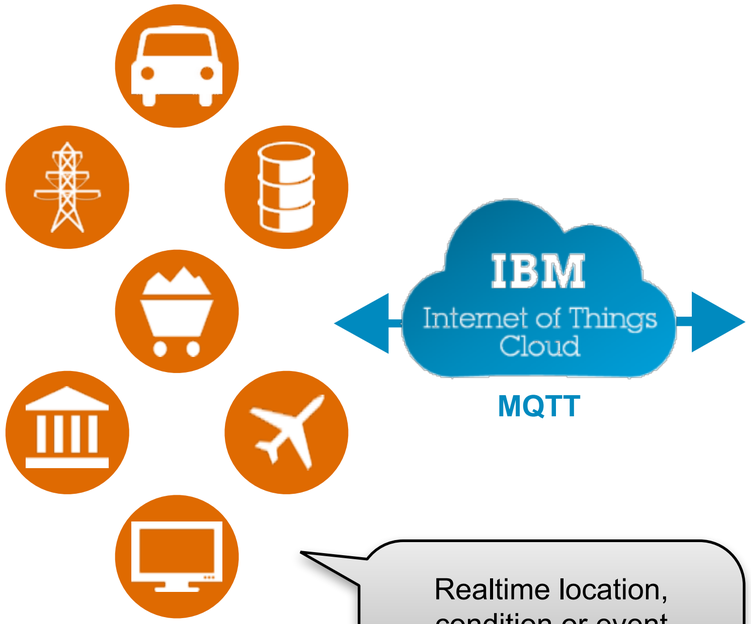
- Web
- Interactive Voice Response
- Mobile apps
- Short Message Service
- Social media
- Chat
- Voice
- Email

The Value of Predictive Customer Insight (PCI)



Predictive Maintenance and Quality (PMQ)

- Analyzes asset & service data from multiple sources, predicts failures and events, and provides recommended (preventive) actions



Monitor, maintain, and optimize assets

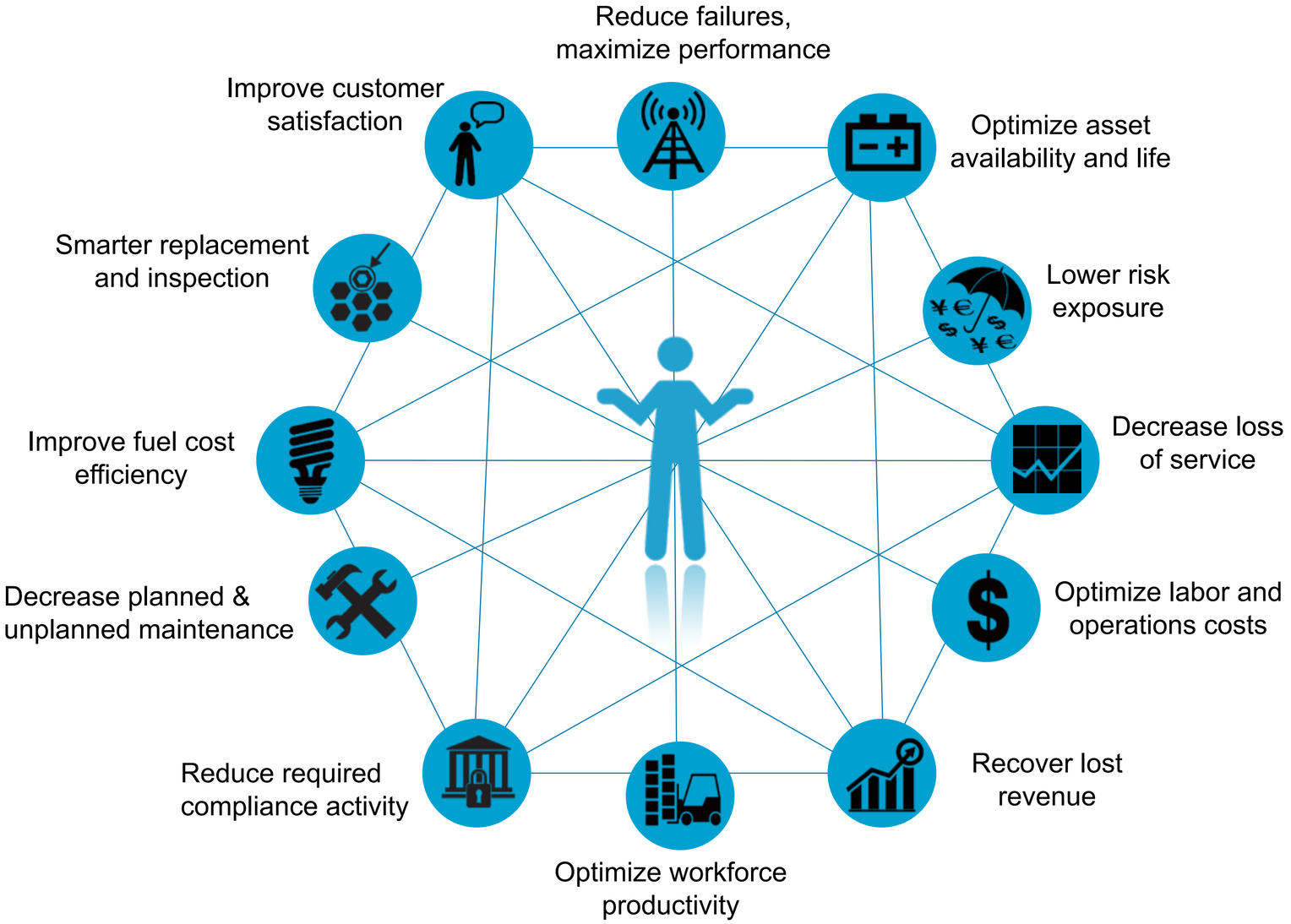
Predict asset failure and identify poor quality parts

Reduce guesswork

"System of Insight"

Data related to asset sales, components, service history and operations

The Value of Predictive Maintenance & Quality



Unplanned downtime

An aircraft engine manufacturer uses predictive analytics to prevent costly aircraft-on-ground engine events.

Challenge:

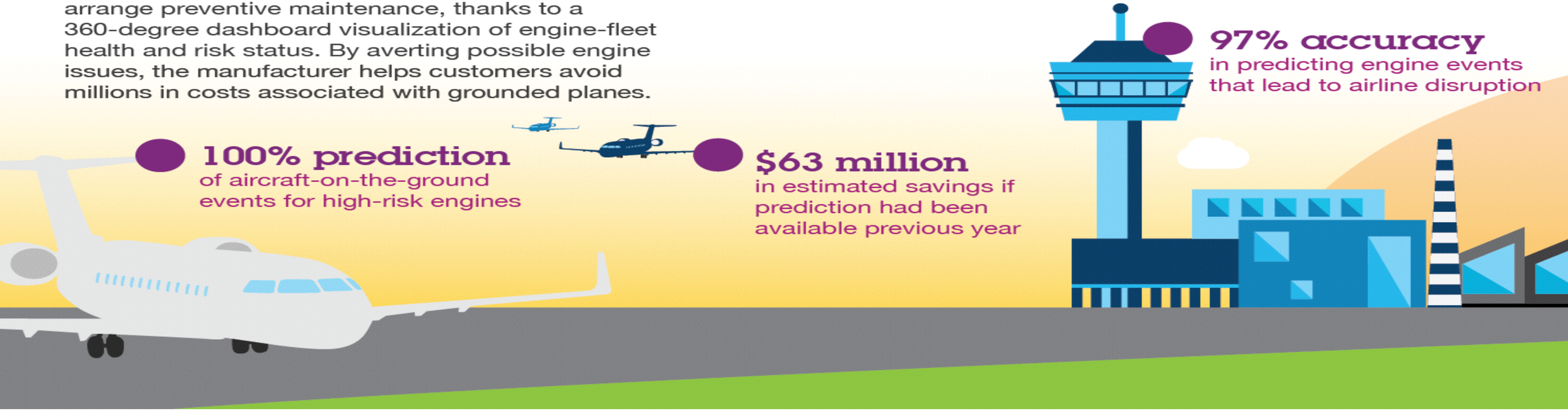
Collect large amounts of data about company's engines through various methods, while fully analyzing all available information to proactively address issues.

Lesson:

Analytics solutions provide a quick answer to an important question: How is our business doing today?

Solution:

An analytics platform automatically alerts the manufacturer to different types of impending engine events. The company then works proactively to arrange preventive maintenance, thanks to a 360-degree dashboard visualization of engine-fleet health and risk status. By averting possible engine issues, the manufacturer helps customers avoid millions in costs associated with grounded planes.



100% prediction
of aircraft-on-the-ground events for high-risk engines

$63 million
in estimated savings if prediction had been available previous year

97% accuracy
in predicting engine events that lead to airline disruption

Scrap reduction

Global automotive manufacturer uses advanced analytics to reduce scrap rates in manufacturing process.

Challenge:

An auto manufacturer noticed reject and scrap rates for engine blocks were much higher than expected. Minute cracks were identified, but officials didn't know how they were caused.

Solution:

The company used an advanced analytics platform to analyze various types of existing information. This analysis showed that high scrap and reject rates happened when liquid metal was poured into the mold during afternoon hours, because the temperature around the manufacturing line spiked several degrees during that time. The simple solution was to install blinds on a window, saving several million dollars for the organization.

Lesson:

Big problems can sometimes have simple solutions. Root-cause analysis can accurately identify process anomalies.

IBM

Yield on production

A Middle East electricity provider increases grid reliability.

Challenge:

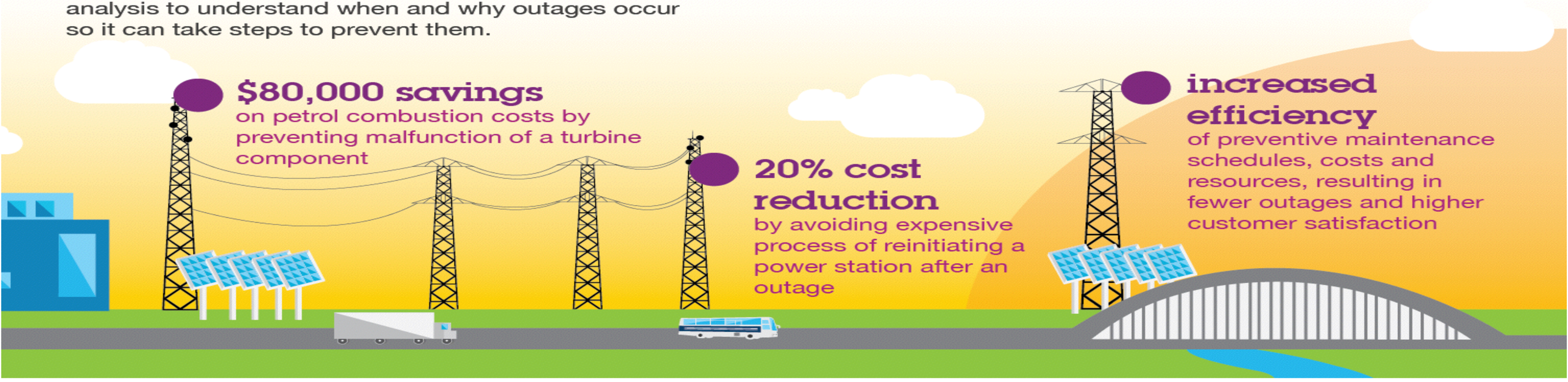
The power generator's research institute must improve safety and reliability of power generation and transmission while remaining innovative. Besides planning for disruptive events such as solar storms, the company has a duty to improve efficiency, incorporate new sources of renewable energy, and analyze volumes of data from an increasingly smart grid.

Solution:

The energy provider uses powerful predictive analysis to understand when and why outages occur so it can take steps to prevent them.

Lesson:

Predictive analytics evaluates possible problems to be resolved before they occur.



Questions to You

- What types of applications or opportunities you'd see with big data? What would be required to make it happen?

- What is the real value of big data? (for e.g people, society, business)

- What obligations or responsibilities do you see around big data? (concerning e.g individuals, organizations, business, legal, technology…)

Internet of Things

Internet of Things (IoT) connects Physical and Digital worlds

Smart Scales
Track health in outpatients, remote measurements

Connected Car
Track location and status

Shipment Tracking
End to end container tracking, prevent tampering

Health Care
Monitor patients at home

Vending Machines
Conditions, shelf life, stock reporting

Smart Deliveries
Track parcel, monitor and open door remotely on arrival

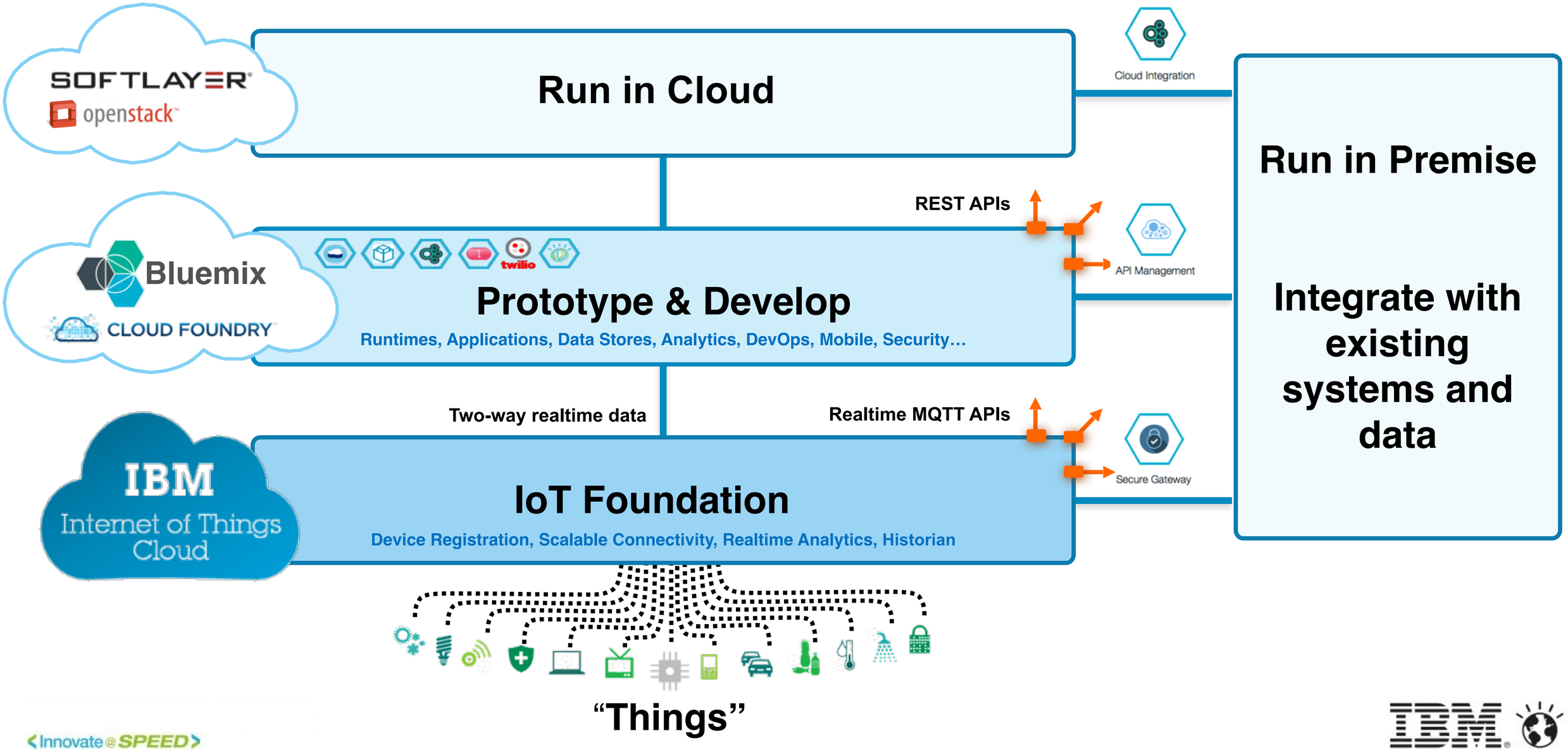
Improved Security
Facial recognition, remote notification

Smart Metering
Track and control usage

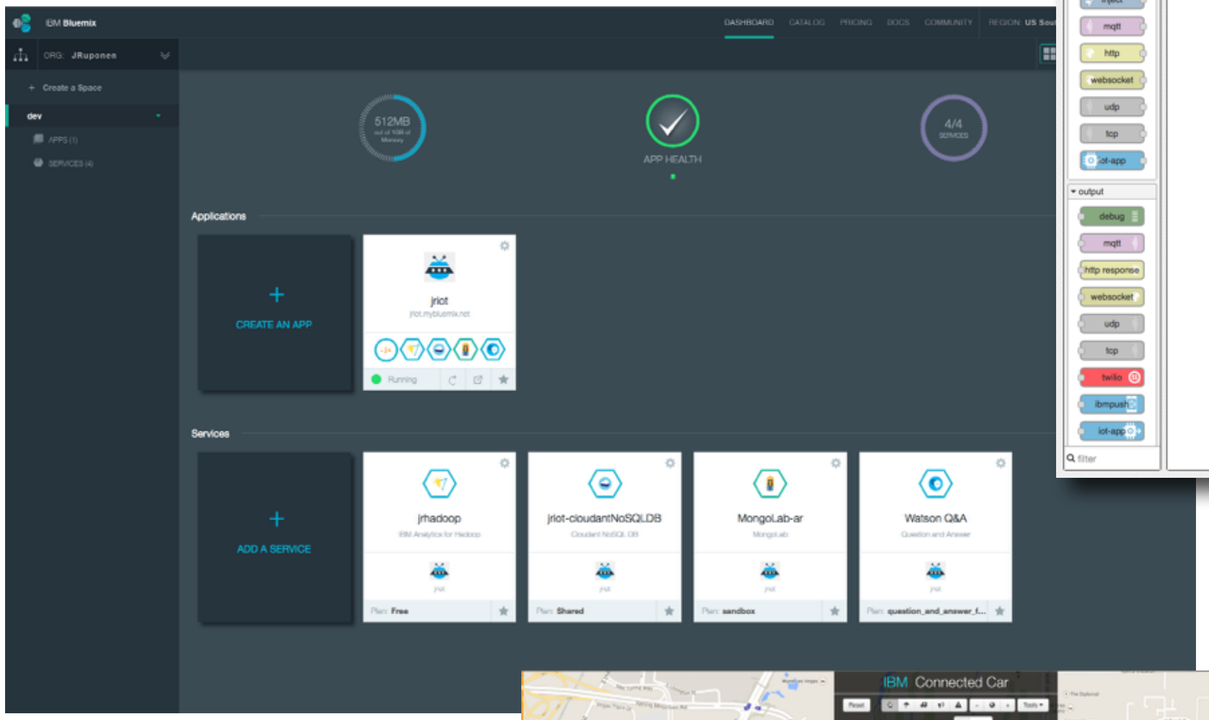
Mobile Transactions
Mobile payments, signaling & controlling

Smart Buildings
Maximum efficiency using presence detection, weather predictions and remote control

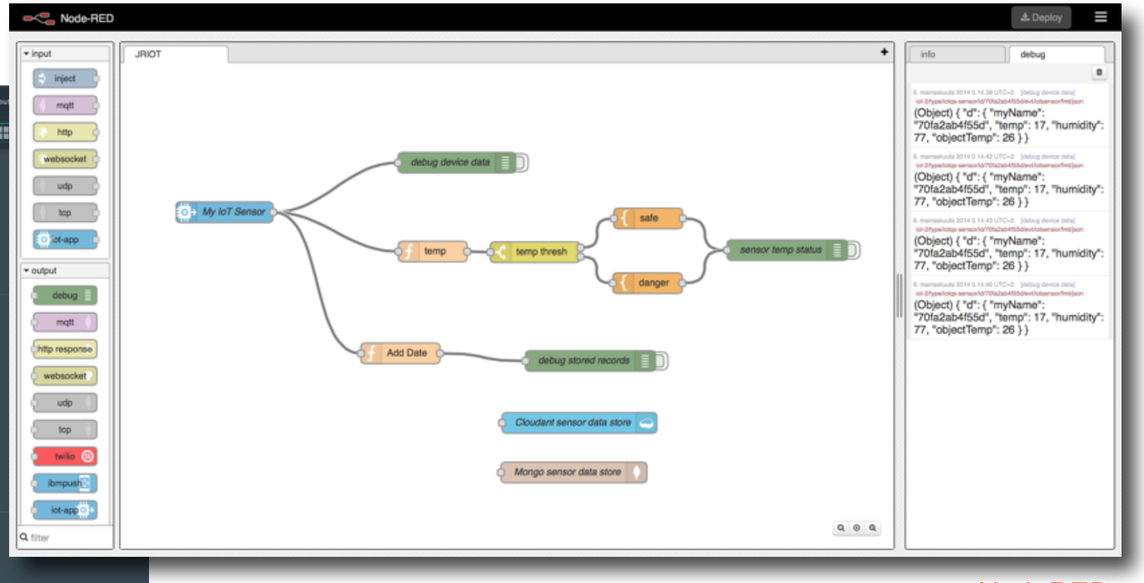
IBM IoT Foundation, Bluemix and Softlayer



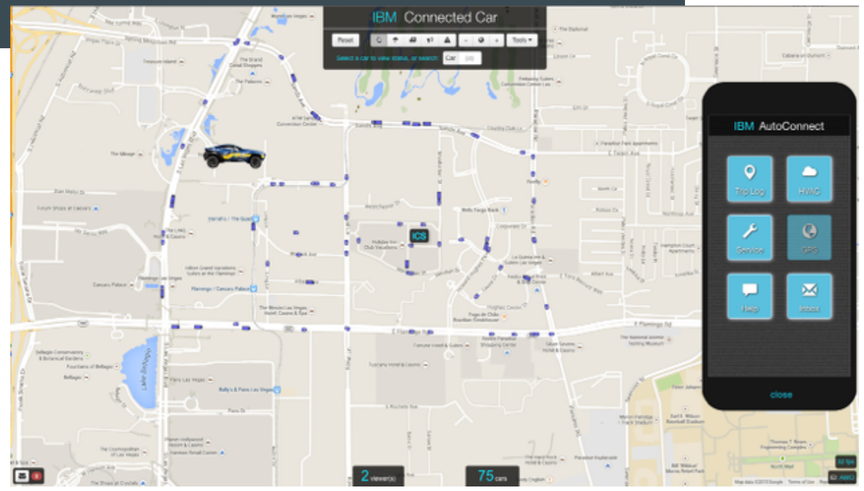
Some demos



bluemix.net



NodeRED



Recipes

Use case example: **Context-Aware Systems of Insight** (prototypes)

Returns context aware information, e.g. alerts and notifications

Context Aware APIs

Intelligent **Mobile First** applications

IoT Foundation
5-10 sec motion samples with MQTT messaging

Predictive Analytics
Association Rules, Pattern matching

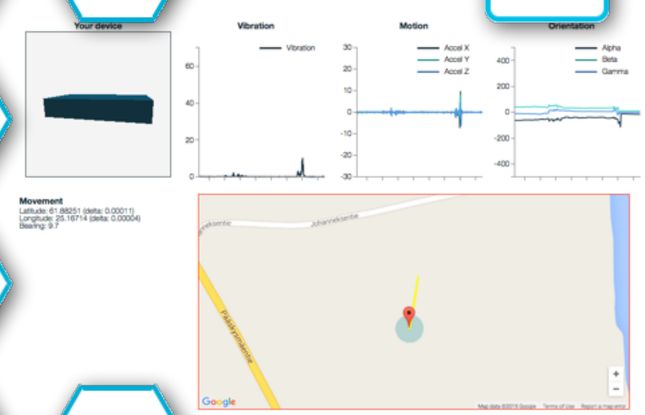
API Management
New API's for intelligent alerting and notifications

Geospatial Analytics
Real-time geofencing, based on location, type-of-motion and prevailing conditions

Insights from Social Data
Social sentiment analysis

BigInsights Hadoop for raw & history data. Data mining & pre-processing with text analytics. User preferences & profiling. Data federation and enrichment of structured data.

Insights for Weather
Weather data analytics



Road conditions

Train tracking

Digitraffic & Digiroad
REST API or WebSocket

Liikennevirasto

Digitraffic & Digiroad. Real time traffic information for road and rail traffic, e.g travel time system, automatic measuring devices (LAM), road weather stations and road weather & surface cameras.

Route planning

Bus tracking

Digitransit
REST API or WebSocket

Digitransit. Time table & travel guide for public transportation. Combines data from different providers (e.g HSL).

Locations, events

Linked Events
REST API

HELSINKI REGION INFOSHARE

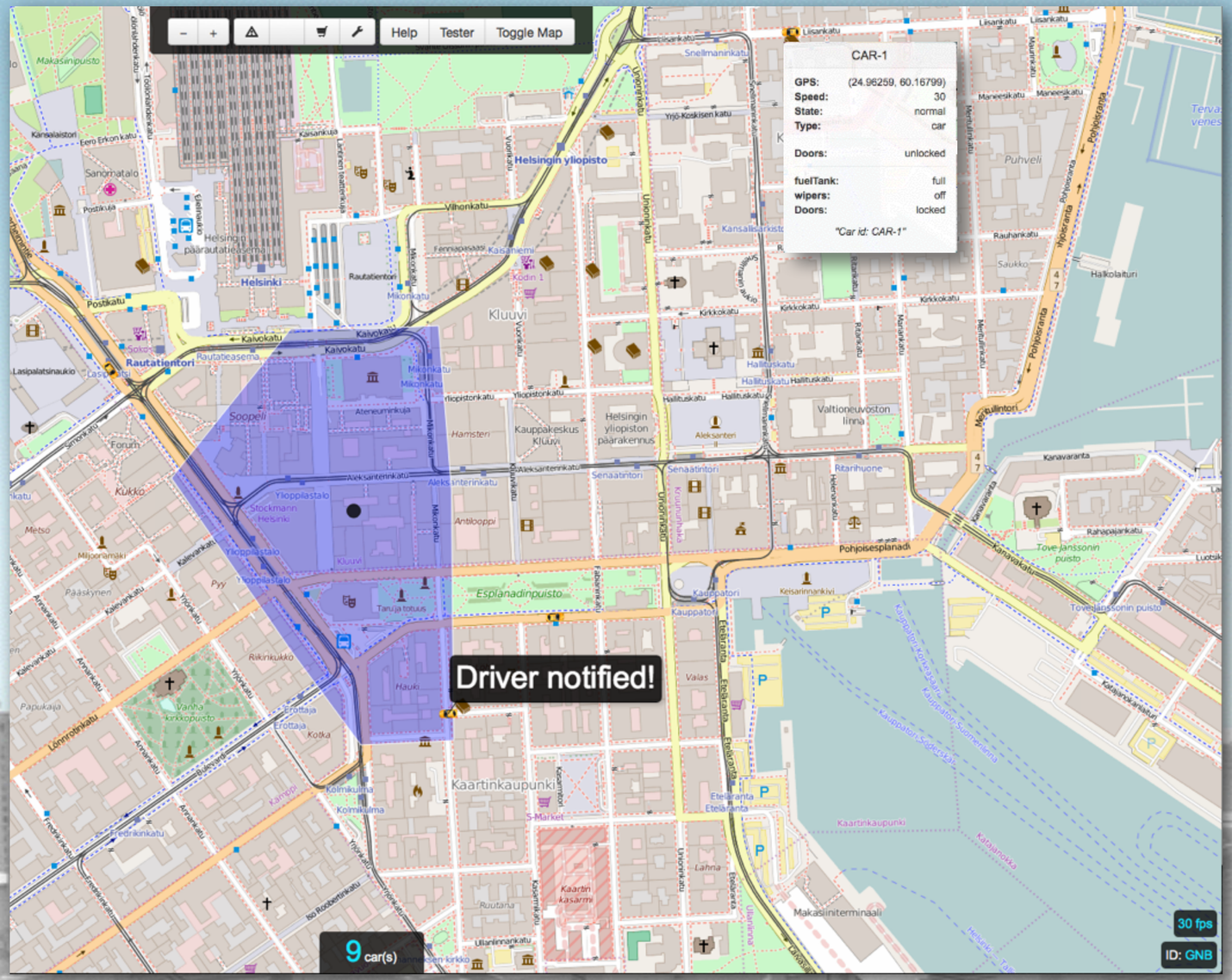
Linked Events providing categorized data on events and places for the Helsinki capital region.

Reittiopas
REST API or WebSocket

HSL HRT

Provides HSL real-time public traffic information, exceptions, realtime vehicle information and routing statistics.

Use case example: **Location & Condition aware Alerting & Notification**



Explanation:
Stream computing is used here with Geospatial Analysis to notify or alert people (or vehicles) as they enter (or leave) a region of interest under certain conditions. Location can be detected from GPS on a mobile device or in vehicle. Geospatial Regions can be created programmatically, based on events and conditions acquired from open API's, user preferences and/or user input.

Example types of Geospatial Regions:
Road conditions, traffic conditions, service announcements, accidents, user alerts, help requests, business opportunities…

Examples of Analytic services:

Streaming Analytics. Ingest, analyze, monitor, and correlate data streams as they arrive from real-time data sources. Analytic models for e.g. statistics, predictive, text, acoustic/sound and video.

Geospatial Analytics. Connect to data sources that support the MQTT protocol and monitor devices as they move into geographic regions of interest. Define geographic regions and control monitoring of regions using the geospatial application programming interface.

Predictive Analytics. SPSS analytics platform that developers and data scientists will use integrate predictive capabilities with applications (segmentation, scoring, classification, pattern detection, association rules, anomalies detection etc).

API Management. Using existing API's and creating new API's for intelligent alerting and notifications

Connected Car platform

with MQTT messaging, Stream computing (Geospatial & Geofencing) and Mobile Apps for Predictive Maintenance etc.

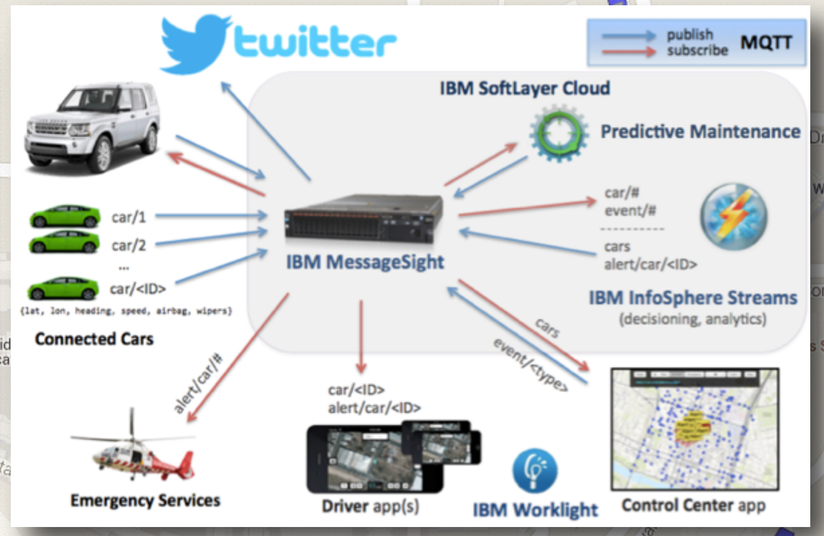
Reset | Black Ice | Accident | Emergency | Roadwork | − | + | Tools

Select a car to view status, or search: Car (id)

Car 66

GPS:	(-115.16249, 36.11857)
Heading:	271.5°
Speed:	11 mph
Airbags:	Off
Wipers:	Off
Emergency Vehicle:	On / OFF

Send Message
Show Messages
Drive
AutoConnect



< IBM AutoConnect

SERVICE

8/9/2013

10000 miles ✓
Routine maintenance
1/2/2014

15000 miles
Detailed maintenance and inspection

20000 miles
Routine maintenance

25000 miles
Routine maintenance

30000 miles
Detailed maintenance, inspection, and filter replacement

close

1 viewer(s)

75 cars

23 fps

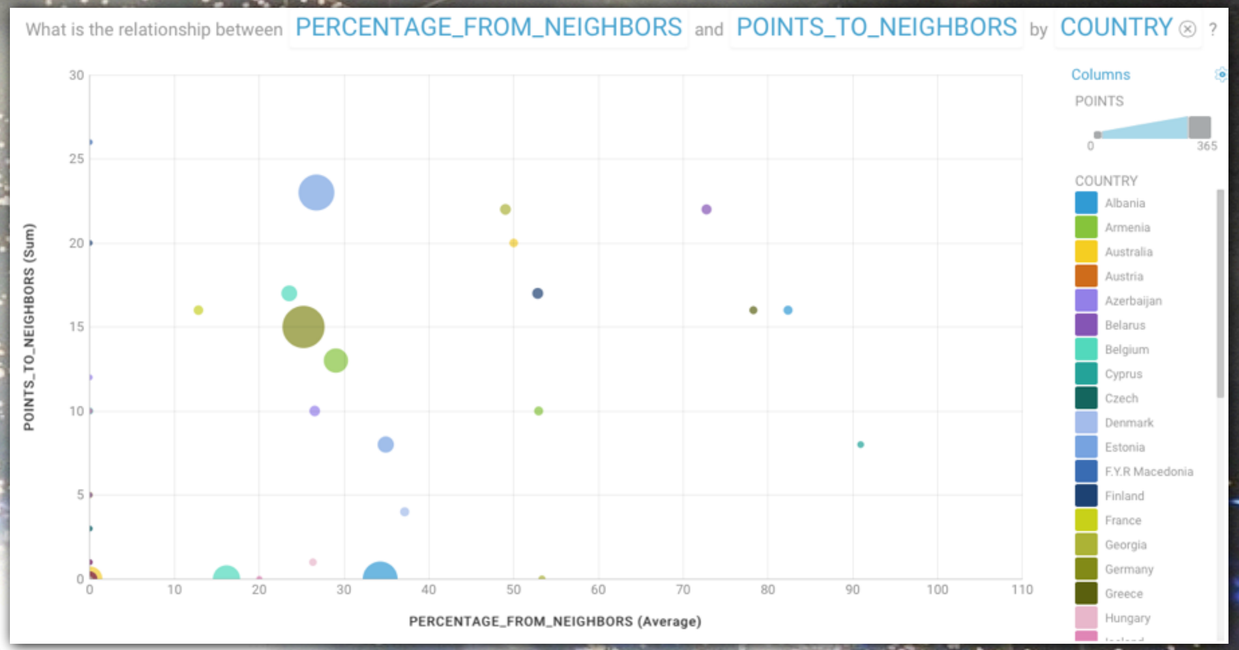
ID: EZV

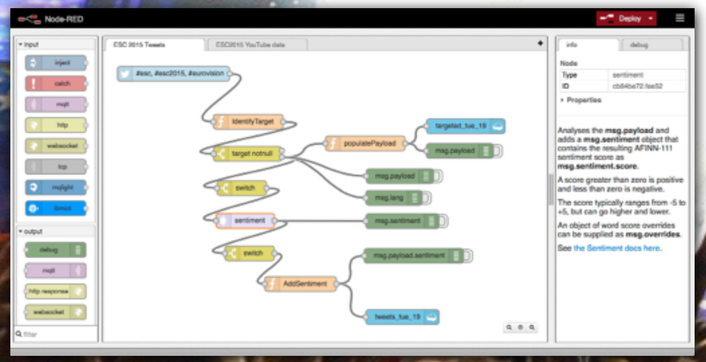
Use case example: **Predicting contest results based on Social Sentiment**

IBM Watson Analytics

EUROVISION SONG CONTEST

YouTube



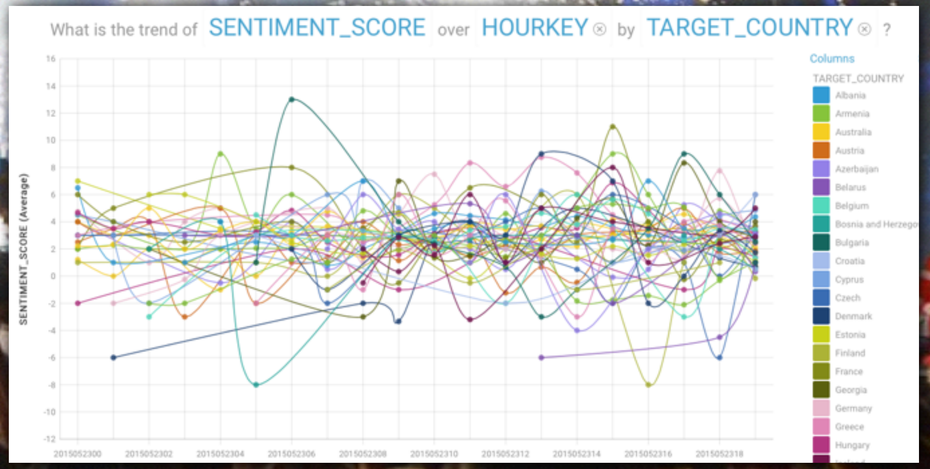


IBM Bluemix & Node-RED

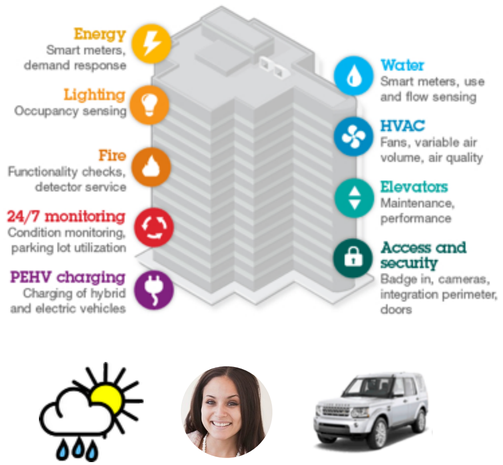
Cloudant NoSQL DB

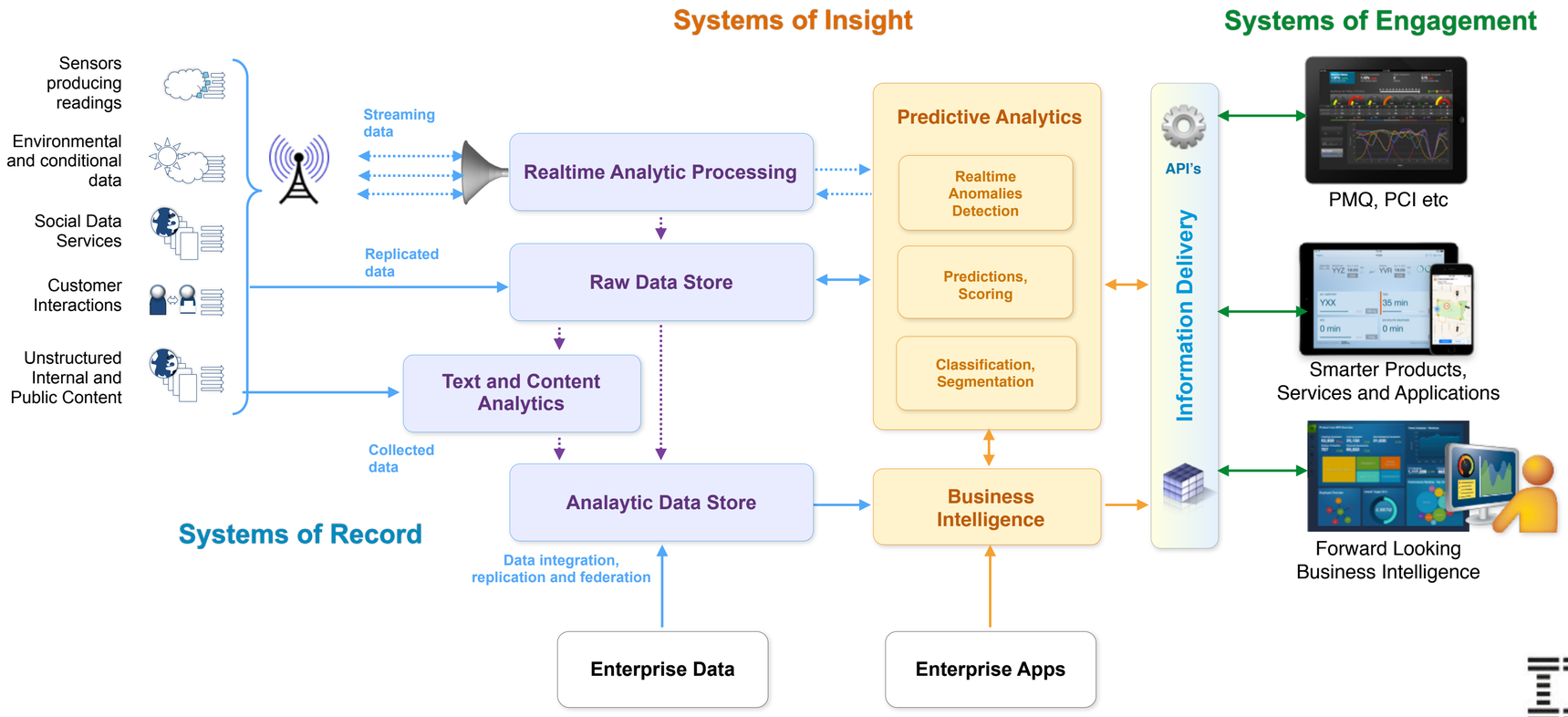
dashDB

IBM Bluemix Cloudant + dashDB



Using IBM cloud platforms from Fast Prototyping to Production





Innovate@SPEED

Thank You

Jukka Ruponen
Business Analytics Architect

jukka.ruponen@fi.ibm.com
+358-40-725-6086