GEND

Matin

| | | |
|---|---|---|
| 08:30 – 09:00 | *Café d'accueil* | |
| 09:00 – 09:15 | Introduction | *Massimo Lucrezia – IBM Genève* |
| 09:15 – 09:45 | Keynote speech: Bringing new data sources for your BI | *Henrique Säuberli – IBM Research THINKLab* |
| 09:45 – 10:30 | Big Data en Action: à travers des cas d'usage | *Vitor Lundberg, Solution Engineer EMEA, Hortonworks* |
| 10:30 – 11:00 | *Pause-Café* | |
| 11:00 – 11:45 | Hadoop and Spark: new framework for new insights | *Christophe Menichetti – IBM Montpellier* |
| 11:45 – 12:30 | Positioning Big Data Solution: NoSQL, Hadoop, In Memory puzzle | *Christophe Menichetti – IBM Montpellier* |
| 12:30 – 14:00 | *Repas* | |

GEND

brès-m

| | | |
|---|---|---|
| 14:00 – 14:45 | Big Data and Analytics: solutions integration and surroundings | *Oleg Ivanov & Jean-François Gonguet*<br>*IBM Genève* |
| 14:45 – 15:30 | Big Data on IBM Power Systems | *Vitor Lundberg & Christophe Menichetti* |
| 15:30 – 15 :45 | *Pause-Café* | |
| 15:45 – 16:15 | Big Data en entreprise: leçons apprises | *Alexandre Masselot - OCTO Technology* |
| 16:15 – 16:45 | Conception, demo, test avec IBM Client Center | *Alain Roy – IBM Montpellier* |

# MERCI!



IBM

common Romandie

# IBM Research - Zurich

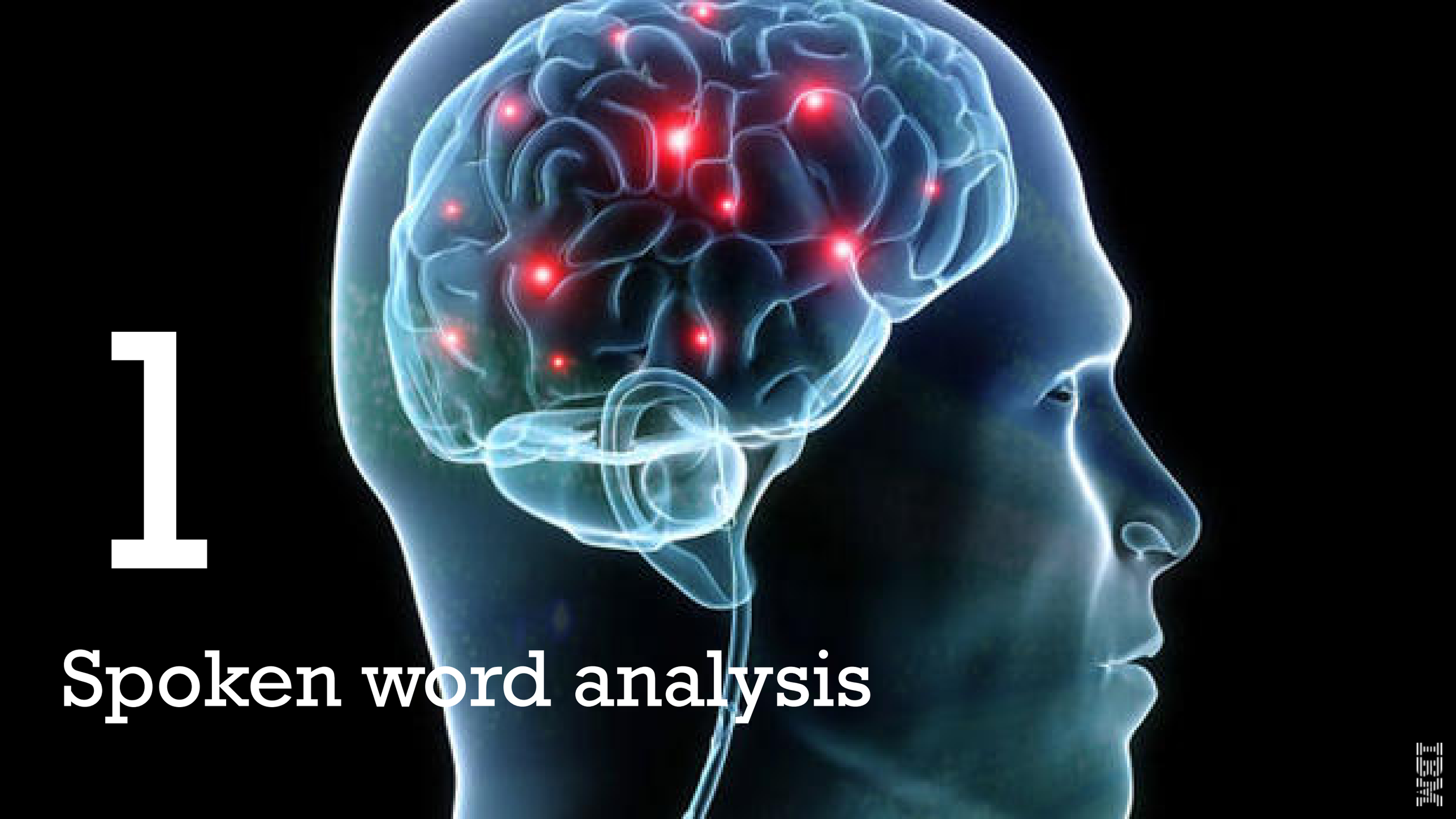
IBM Research - Zurich

From Atoms to Qubits to Big Data Analytics

Henrique Säuberli
Executive Consultant

last updated 08 May 2017


The world
is our laboratory
IBM



# 1

# Spoken word analysis

# 2

# Superhero sight

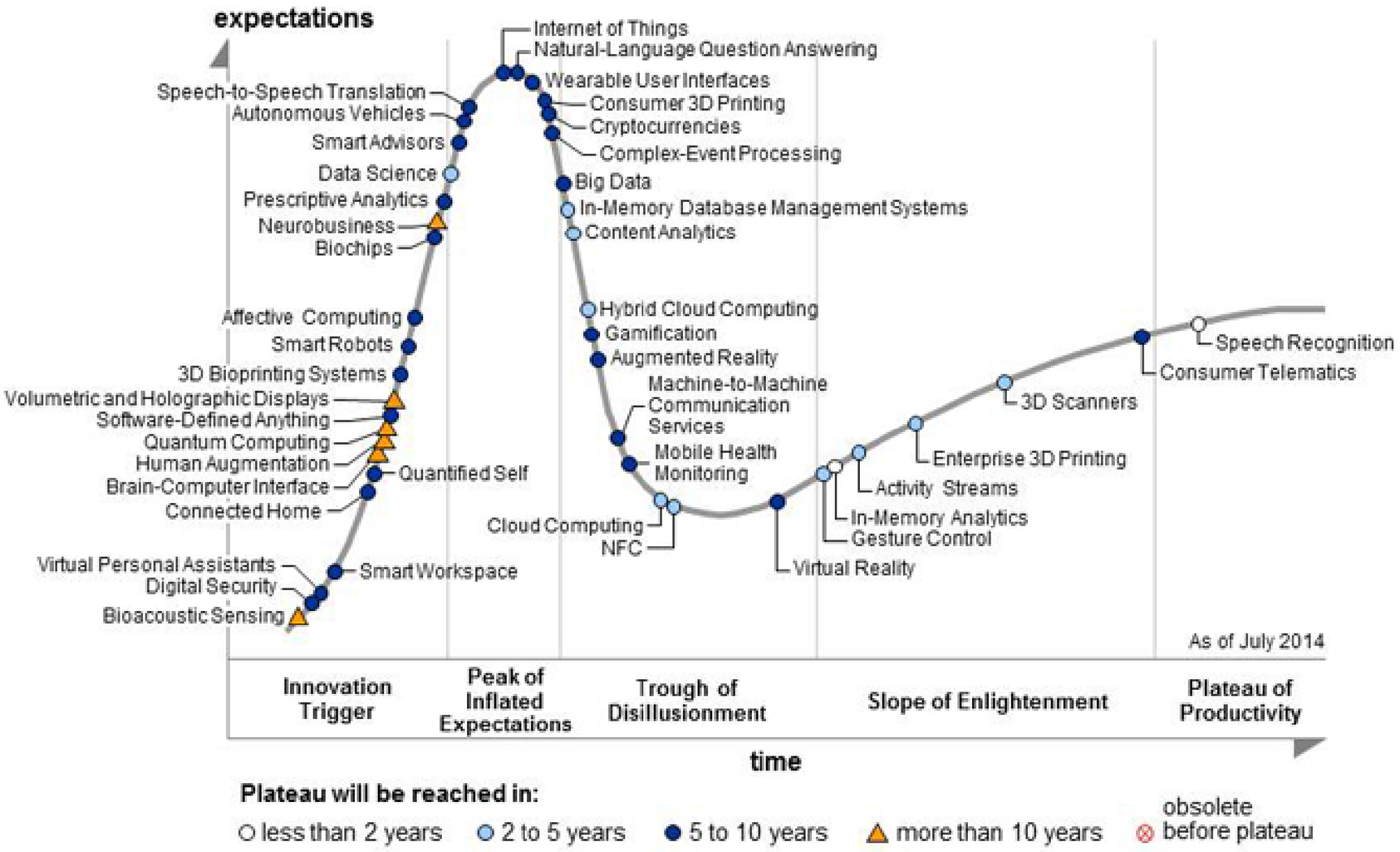

3
Macroscopic

# 4

# Mini medical laboratory

# 5 Enviromental sensors

**"Big data has quickly moved over the Peak of Inflated Expectations and has become prevalent in our lives across many hype cycles."**



# The FOUR V's of Big Data

From traffic patterns and music downloads to w
history and medical records, data is record
stored, and analyzed to enable the technolo
and services that the world relies on every d
But what exactly is big data, and how can the
massive amounts of data be used?

As a leader in the sector, IBM data scientis
break big data into four dimensions: **Volu
Velocity, Variety and Veracity**

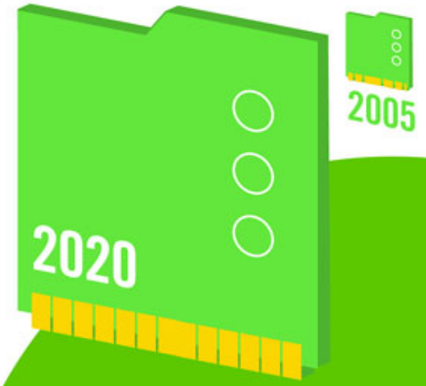
## Volume

### SCALE OF DATA

**40 ZETTABYTES**

[ 43 TRILLION GIGABYTES ]

of data will be created by 2020, an increase of 300 times from 2005




It's estimated that

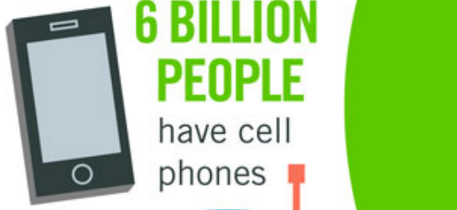
**2.5 QUINTILLION BYTES**

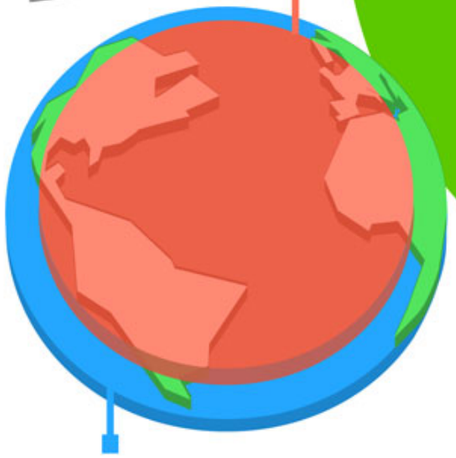
[ 2.3 TRILLION GIGABYTES ]

of data are created each day

**6 BILLION PEOPLE**

have cell phones





WORLD POPULATION: 7 BILLION

Most companies in the U.S. have at least

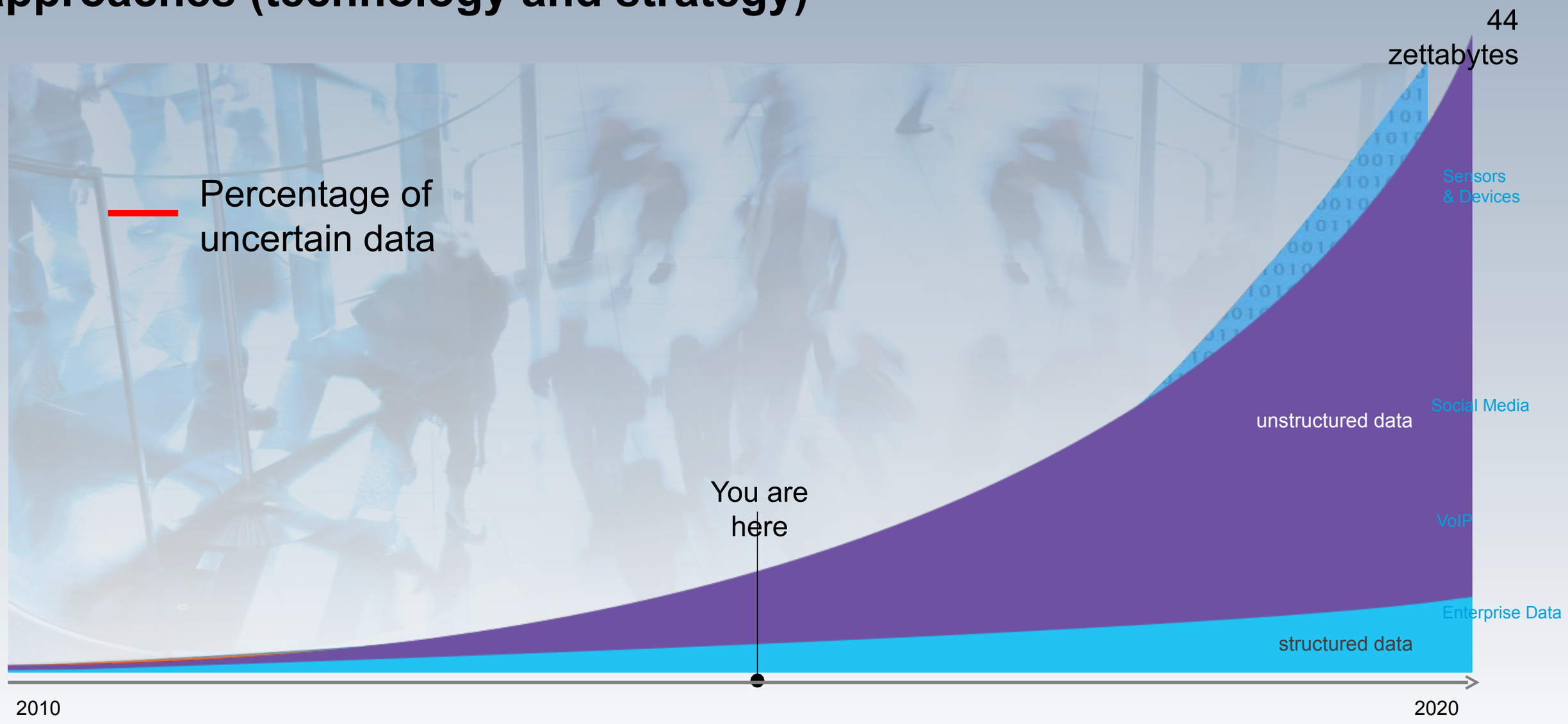
**100 TERABYTES**

[ 100,000 GIGABYTES ]

of data stored

Modern cars have close to

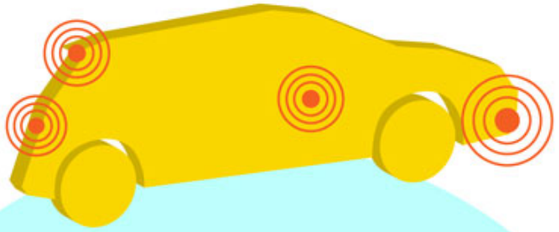
# Data is growing exponentially and demands new approaches (technology and strategy)



The New York Stock Exchange captures

**1 TB OF TRADE INFORMATION**

during each trading session



Modern cars have close to

**100 SENSORS**

that monitor items such as fuel level and tire pressure

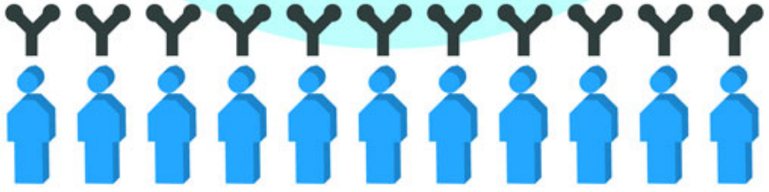
# Velocity

## ANALYSIS OF STREAMING DATA

By 2016, it is projected there will be

**18.9 BILLION NETWORK CONNECTIONS**

– almost 2.5 connections per person on earth



**Velocity, Variety and Veracity**

Depending on the industry and organization, b
data encompasses information from multip
internal and external sources such as transactio
social media, enterprise content, sensors a
mobile devices. Companies can leverage data
adapt their products and services to better me
customer needs, optimize operations a
infrastructure, and find new sources of revenu

By 2015

**4.4 MILLION IT JOBS**

will be created globally to support big data, with 1.9 million in the United States

**Sources:** McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS

# Data at the edge is changing how we look at data

**90%**

of data created over the last 10 years was never captured or analyzed

**2X**

Rate of data creation compared to the expansion of bandwidth over the past decade

**60%**

of valuable sensory data loses value in milliseconds

**by 2017**

The collective computing and storage capacity of smartphones will surpass all worldwide servers

# The FOUR V's of Big Data

m traffic patterns and music downloads to web
tory and medical records, data is recorded,
red, and analyzed to enable the technology
d services that the world relies on every day.
t what exactly is big data, and how can these
ssive amounts of data be used?

a leader in the sector, IBM data scientists
eak big data into four dimensions: **Volume,**
**ocity, Variety and Veracity**

As of 2011, the global size of data in healthcare was estimated to be

**150 EXABYTES**

[ 161 BILLION GIGABYTES ]

## Variety

**DIFFERENT FORMS OF DATA**

**30 BILLION PIECES OF CONTENT**

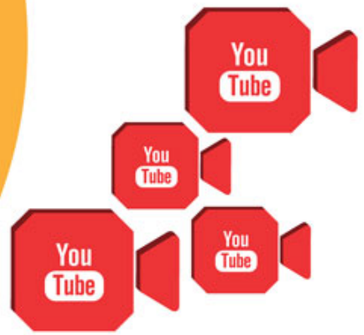
are shared on Facebook every month

By 2014, it's anticipated there will be

**420 MILLION WEARABLE, WIRELESS HEALTH MONITORS**

**4 BILLION+ HOURS OF VIDEO**

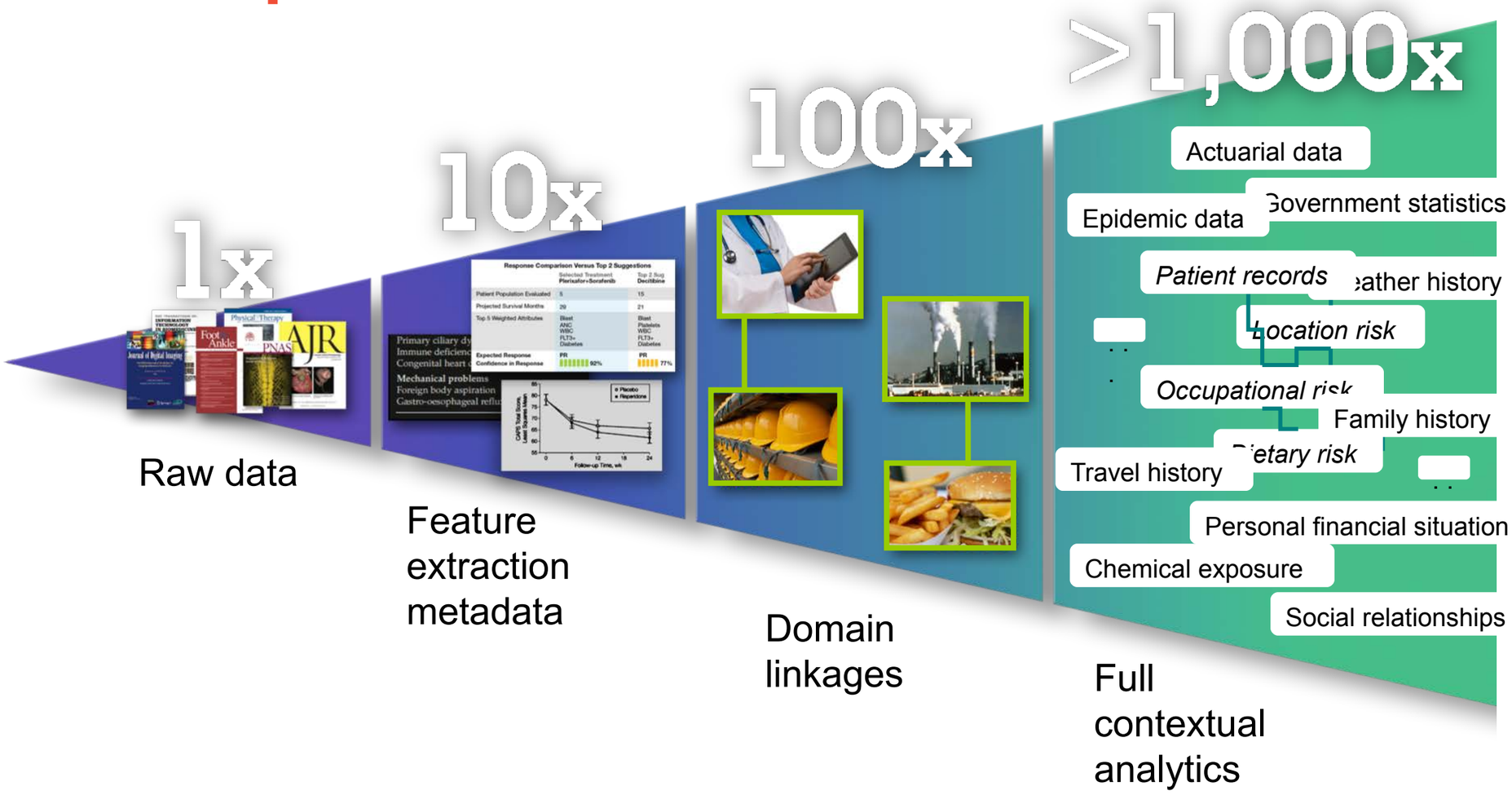
are watched on YouTube each month




**400 MILLION TWEETS**

are sent per day by about 200 million monthly active users

# Context Multiplier Effect



pending on the industry and organization, big
ta encompasses information from multiple
ernal and external sources such as transactions,
cial media, enterprise content, sensors and
bile devices. Companies can leverage data to
apt their products and services to better meet
stomer needs, optimize operations and
rastructure, and find new sources of revenue.

2015

**4 MILLION IT JOBS**

l be created globally to support big data,
h 1.9 million in the United States

**1 IN 3 BUSINESS LEADERS**

don't trust the information they use to make decisions

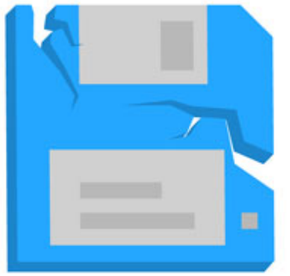

in one survey were unsure of how much of their data was inaccurate

# Veracity

## UNCERTAINTY OF DATA

Poor data quality costs the US economy around

**$3.1 TRILLION A YEAR**



# A vast amount of untapped data could have a great impact on our health - yet it exists outside medical systems.

**60%**
Exogenous Factors

**1100 Terabytes**
Generated per lifetime

Volume, Variety, Velocity, Veracity

**30%**
Genomics Factors

**6 Terabytes**
Per lifetime

**10%**
Clinical Factors

**0.4 Terabytes**
Per lifetime

IBM **Watson Health** // SOURCE: ©2015 J.M. McGinnis et al., "The Case for More Active Policy Attention to Health Promotion," Health Affairs 21, no. 2 (2002):78–93

Data is transforming industries and professions.

How, and why now?

## Oil & Gas

Modern facilities have more than **80,000 sensors** in place, and a single reservoir will produce more than **15 petabytes** of data in its lifetime.


**Reta**
**il**

Consumers post **500 million** tweets and **55 million** Facebook updates each day.

#CognitiveEra

**Public Safety**

New York City surveillance cameras and sensors generate **520 TB** of data per day, largely unstructured.

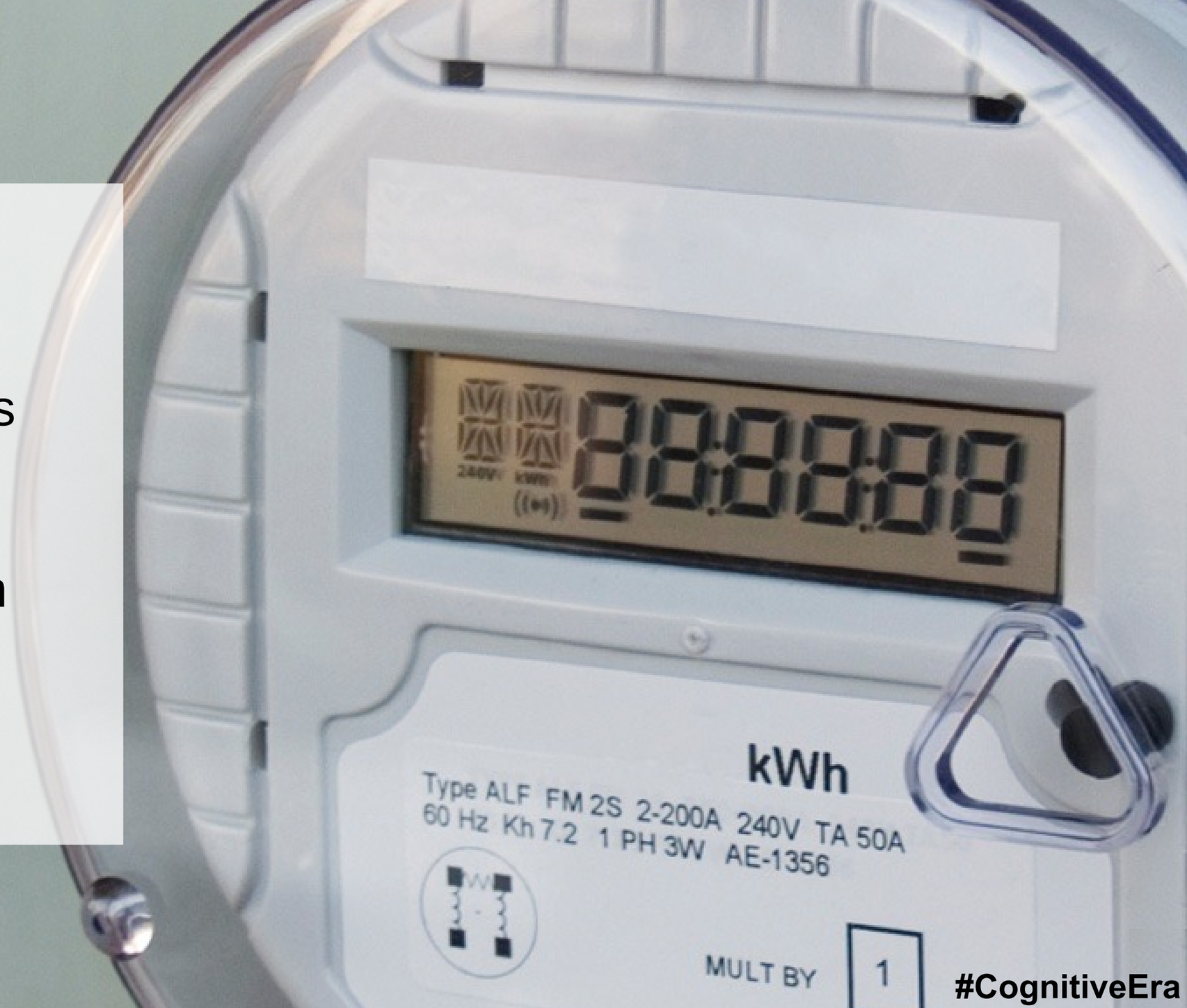
**Securit**
**y**

In 2014, more than **1 billion** personal data records were compromised by cyberattacks.

#CognitiveEra

**Energy and Utilities**

More than **680 million** smart meters will be installed globally by 2017 — producing more than **280 PB** of new data to be analysed and acted upon.




#CognitiveEra

**Healthcare**

Healthcare data recently reached **150 exabytes**, if this continues at projected rates it will grow to fill a stack of DVDs that would stretch from **Earth to Mars**.

**Transportation**

Gartner predicts there will be **250 million** connected vehicles on the road by **2020.**

# Disruption is upon us.

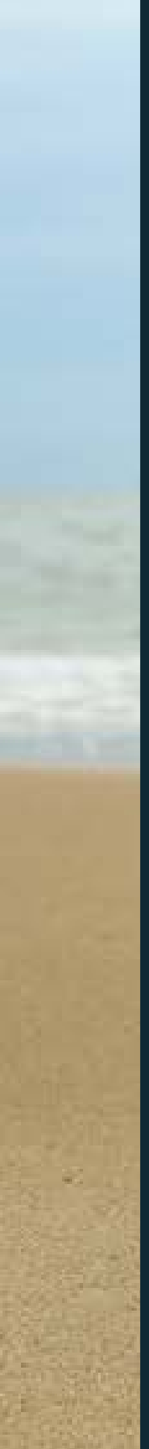
The biggest taxi company
owns no cars.

The largest accommodation company
owns no real estate.


The largest retailer
carries no inventory.


The biggest media company
owns no content.

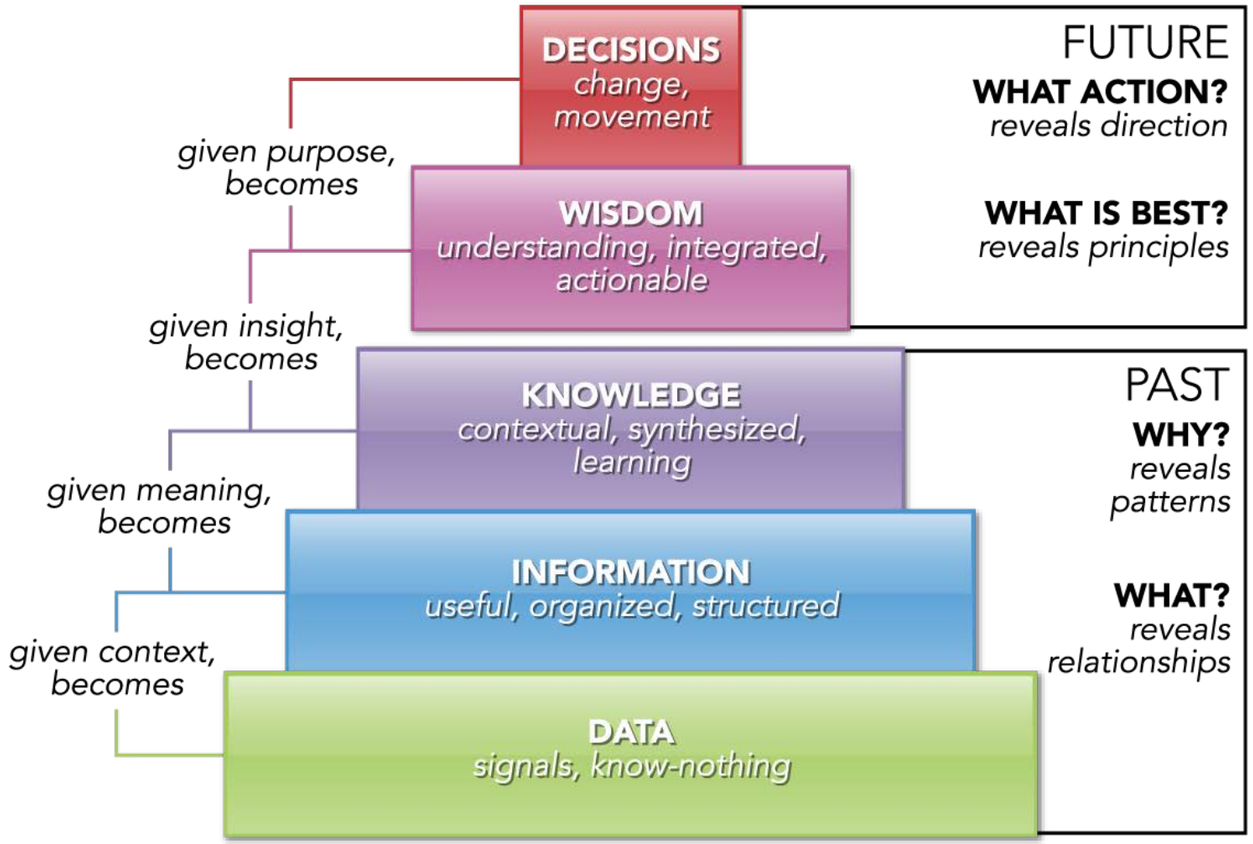
*Unstructured data — "dark data" — accounts for 80% of all data generated today.*



*This is expect to grow to 93% by 2020.*

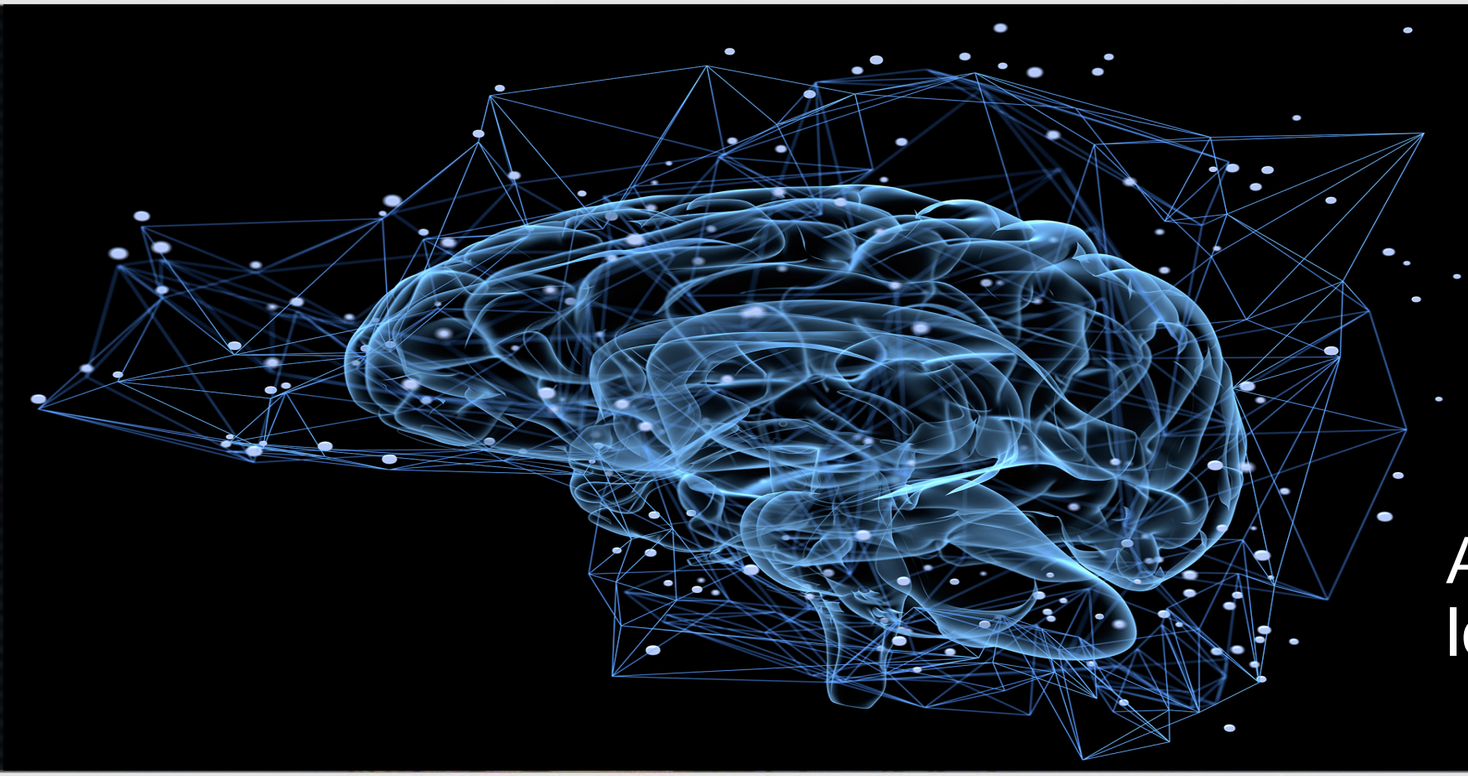
#CognitiveEra

# From data to decision




source: http://www.pursuant.com/blog/redefining-fundraising-data/

Cognitive is **driving** new research …



**Cognitive Computing**

A partnership between people and learning, reasoning machines

What **search** is to **information retrieval**,
**Cognitive computing** is to **decision making**

# Human + Machine

Self-directed goals

Common sense

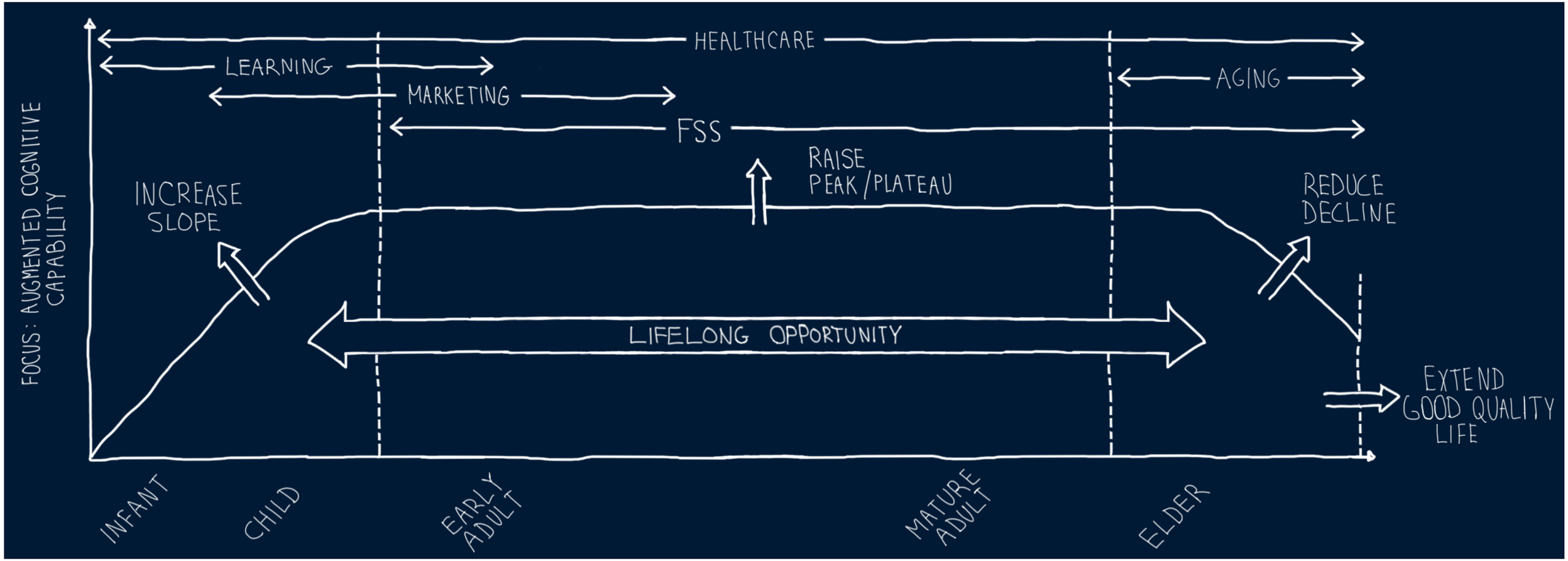
Value judgement

Large-scale math

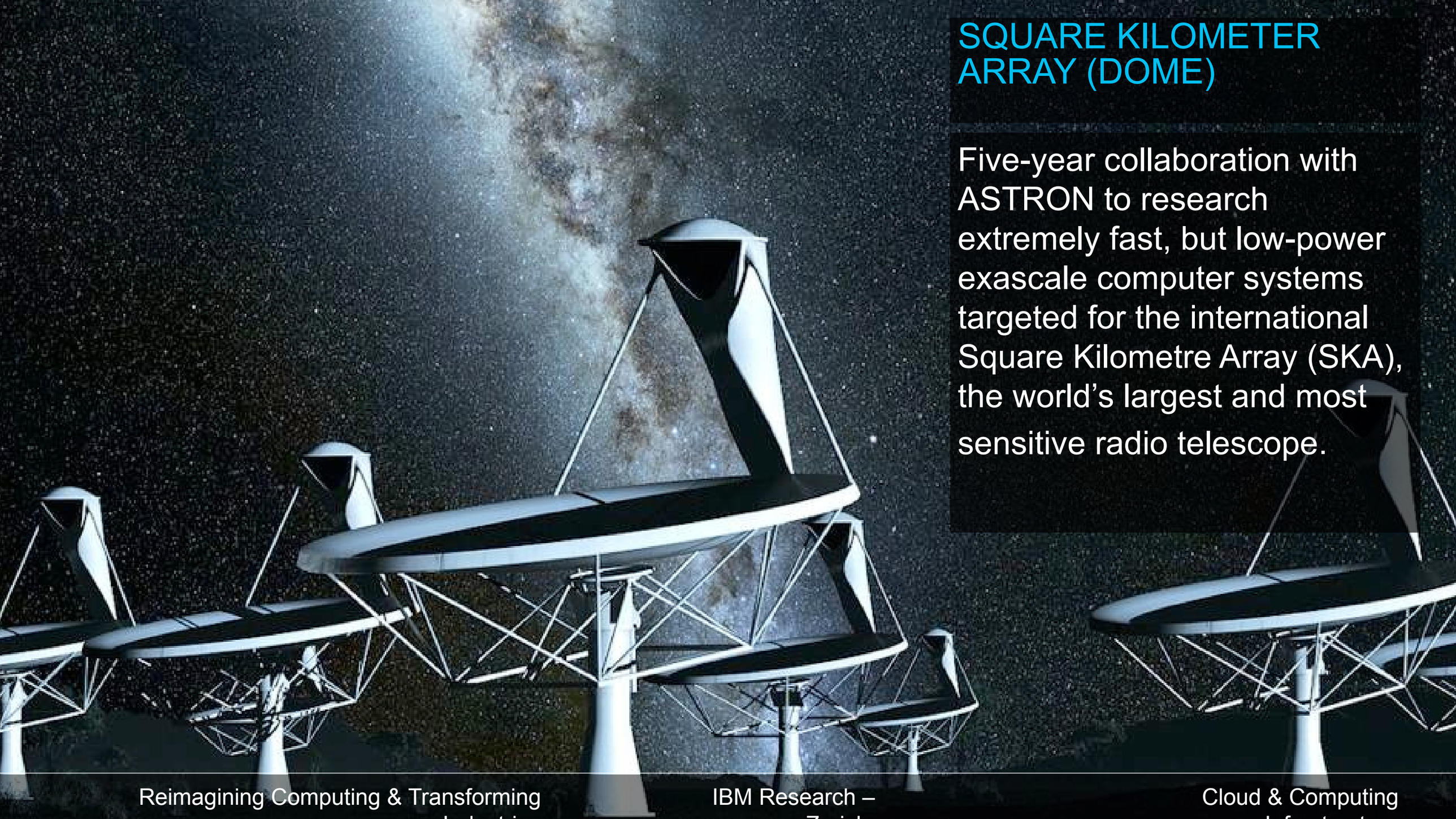
Pattern discovery

Statistical reasoning

# Cognitive Business is the Basis of Lifelong Engagement



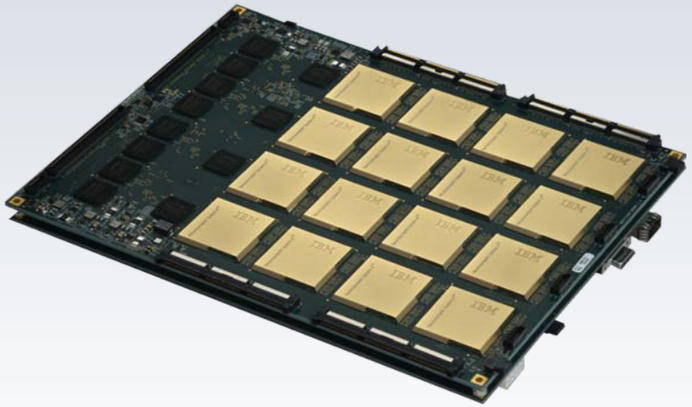


# SQUARE KILOMETER ARRAY (DOME)

Five-year collaboration with ASTRON to research extremely fast, but low-power exascale computer systems targeted for the international Square Kilometre Array (SKA), the world's largest and most sensitive radio telescope.

# IBM SyNAPSE chip


| | 2011 | 2014 |
|---|---|---|
| Programmable neurons | 256 | 1 million |
| Programmable synapses | 262,144 | 256 million |
| Neurosynaptic cores | 1 | 4,096 |



Board:

16 million neurons

4 billion synapese
(deutsch 4 Milliarden)

100 billion neurons
(deutsch 100 Milliarden)

100 trillion synapses
(deutsch 100 Billionen)

# Questions?


IBM

*Executive Briefing*
*Consultant*
*IBM Research THINKLab*

*IBM Research GmbH*
*Säumberstrasse 4*
*8803 Rüschlikon,*
*Switzerland*

*Phone* *+41 44 724 87 27*
*Mobile* *+41 78 767 94 96*
*sae@zurich.ibm.com*


Stay up-to-date about IBM Research:

www.ibm.com/blogs/research/

www.ibm.com/research