Cognitive Computing and the Future of Science

DR. ALESSANDRO CURIONI
IBM FELLOW AND VICE PRESIDENT, EUROPE
IBM RESEARCH

@Ale_Curioni

IBM Research:

12 Labs on 6 Continents and 3,000 Scientists
IBM invests 6% of revenue in R&D annually


Six Nobel Laureates


Nine Medals of Technology


Five National Medals of Science


Three Kavli Prizes

2.5 quintillion bytes of data created every day.

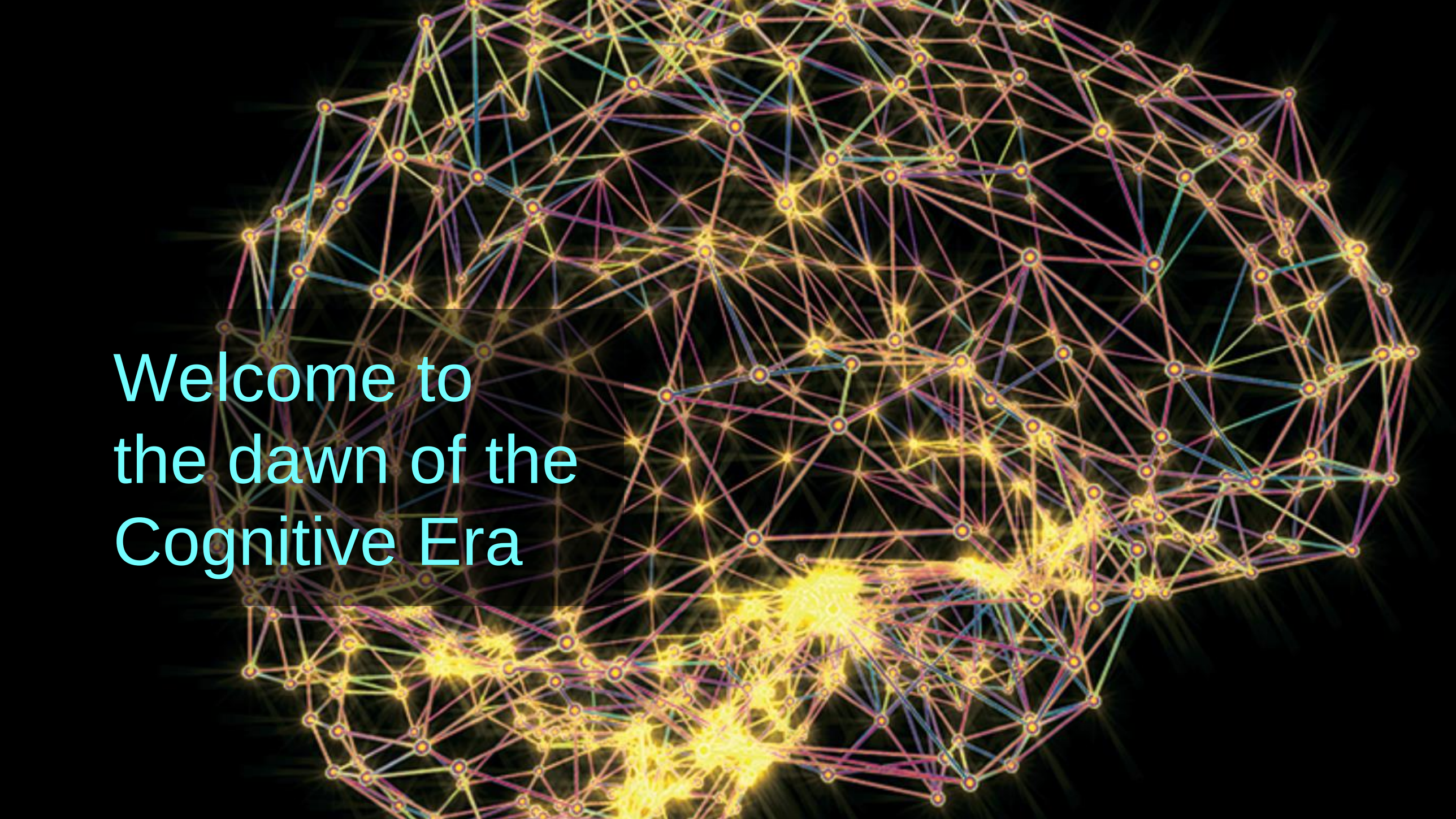
90% of the data in the world today has been created in the last **two years** alone.

Every minute, **1.7 megabytes** of data is created for **every person** on the planet. **All 7.3 billion of us.**

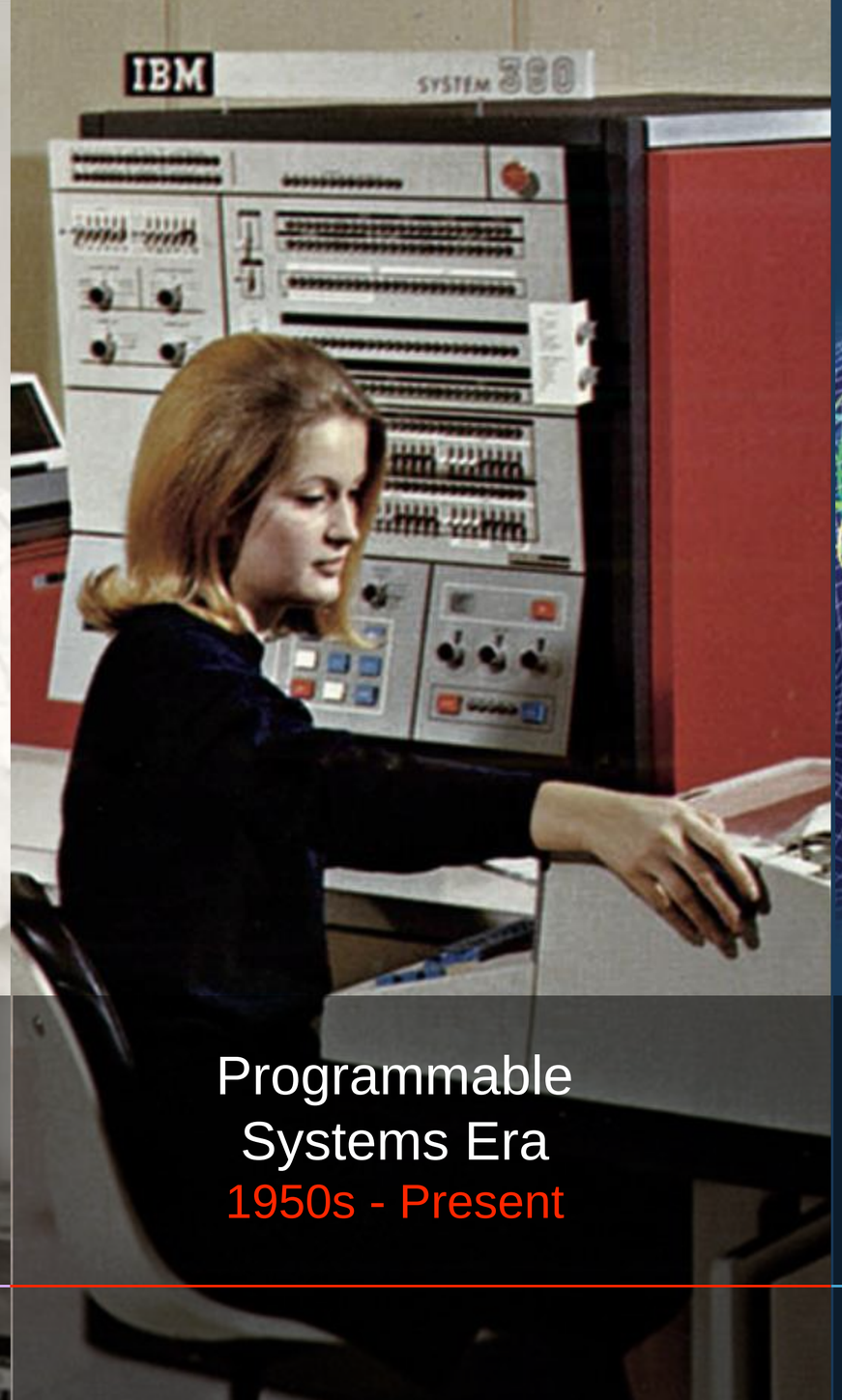
Unstructured data — “dark data” — accounts for 80% of all data generated today.

This is expect to grow to 93% by 2020.

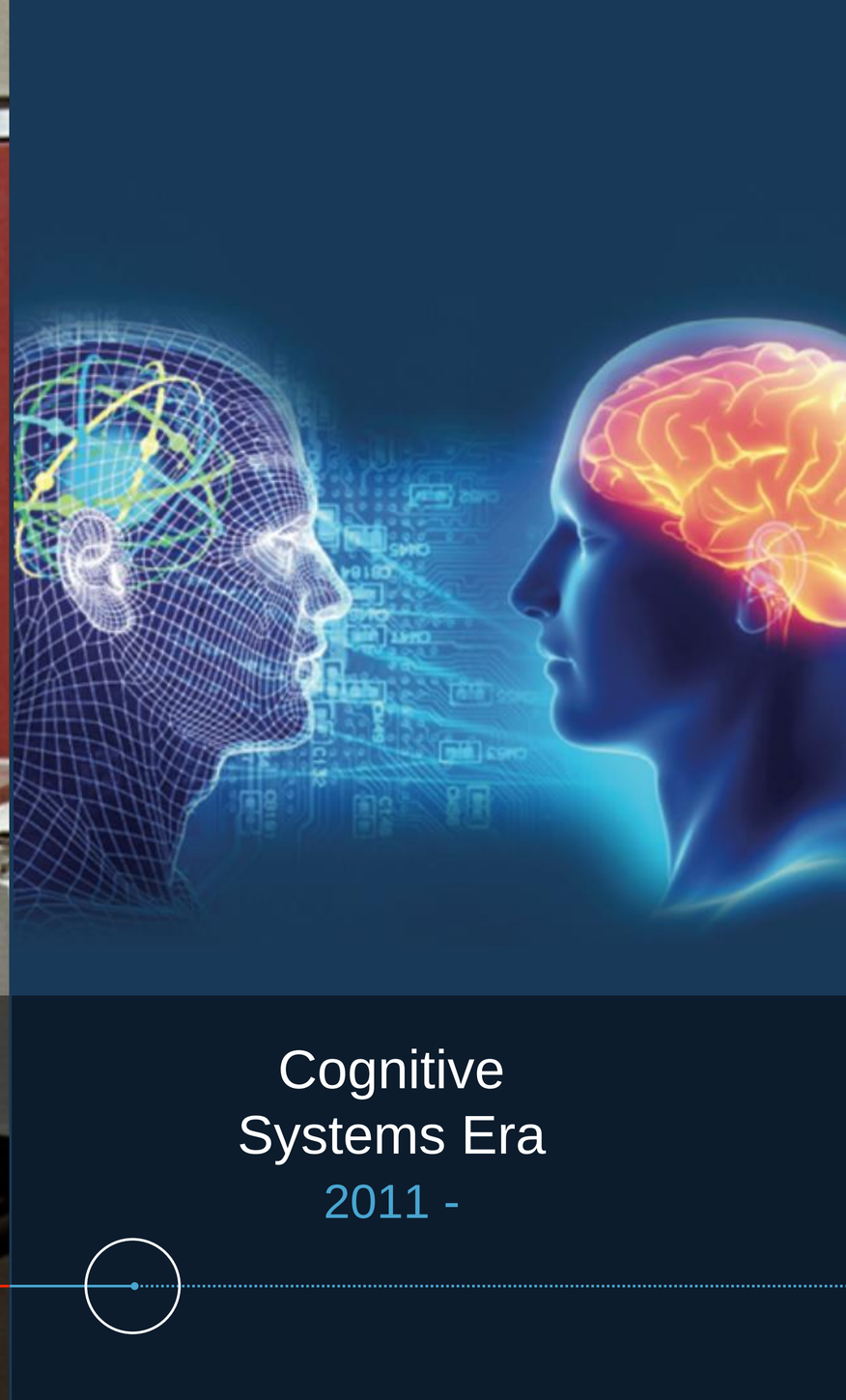

The price of not knowing.


Welcome to
the dawn of the
Cognitive Era

Tabulating
Systems Era
1900 - 1940s



Programmable
Systems Era
1950s - Present



Cognitive
Systems Era
2011 -


PIENSE
THINK
DENKE
PENSER
$18,200
$21,040
WATSON
$5,600
BRAD

Watson in 2011

System Specifications

2880 Processing Cores

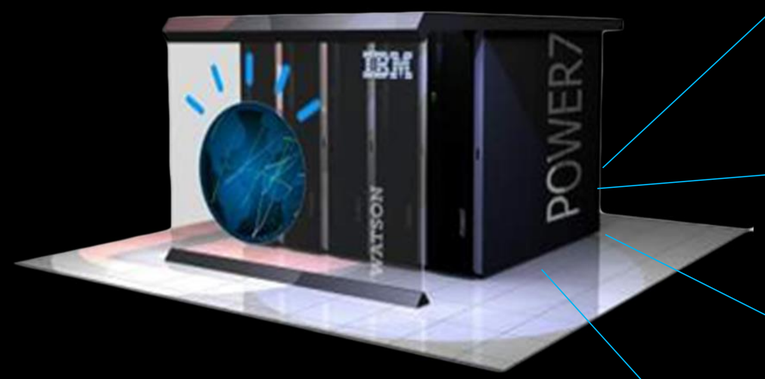
90 IBM P750 Servers

POWER7+™ BUILT ON Power™

16 Terabytes Memory (RAM) – 20TB Disk

80 Teraflops (80 trillion operations per second)

Workload Optimized Systems



IBM Technology Depth

Content Analytics

Business Analytics

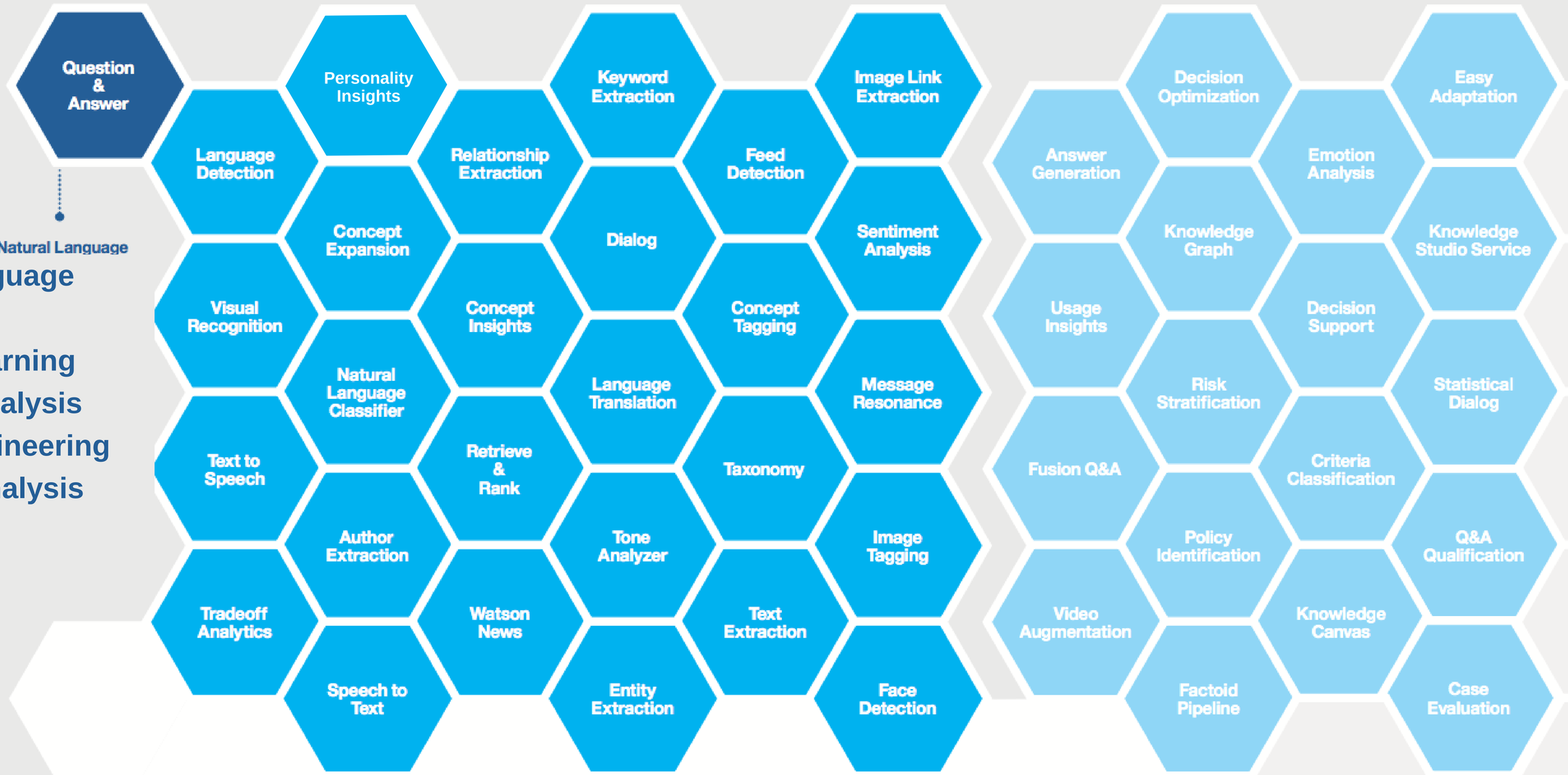
Big Data

Databases / Data Warehouses

Watson in 2016

Since then, Watson has grown to a family of **28 APIs**.

By the end of 2016, there will be nearly **50 Watson APIs**— with more added every year.

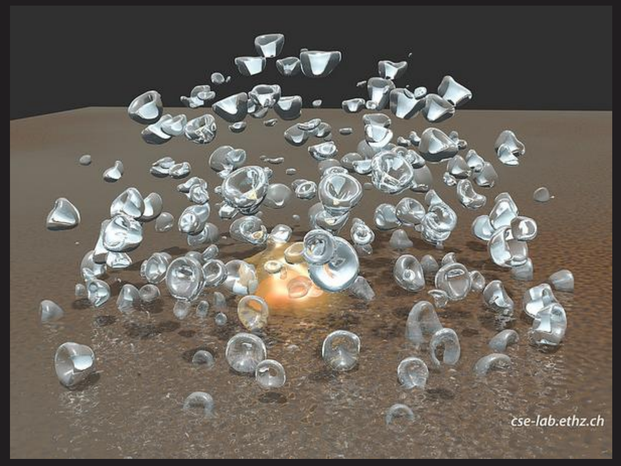


Natural Language Processing
Machine Learning
Question Analysis
Feature Engineering
Ontology Analysis

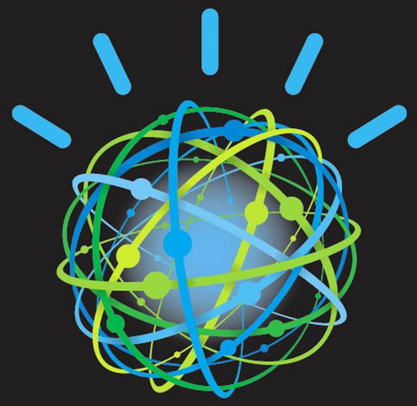
More Data + New Technologies = Scientific Discovery

1000s Years Ago
Theory

Last 100s of Years
Experimentation



Recent Decades
Computer Simulation

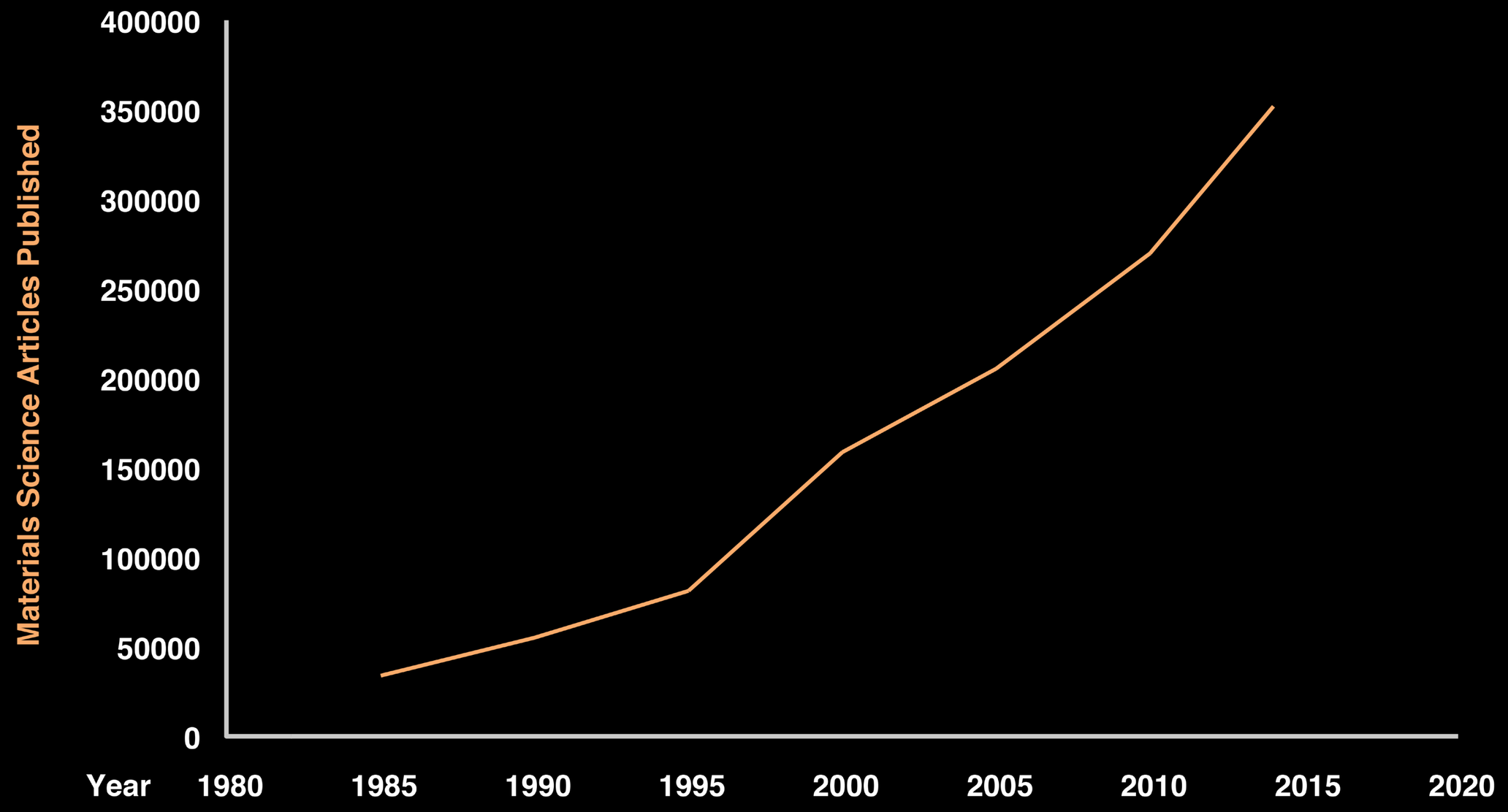


Today &Tomorrow
Cognitive Discovery

Our ability to discover is directly linked to the amount of data available

Cognitive Computing for Discovery

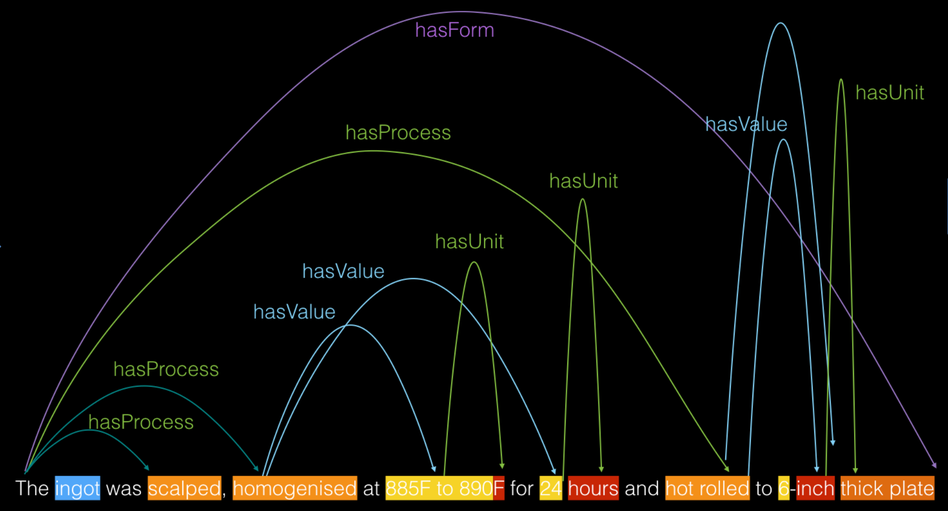
Unstructured Data Deluge in Peer Review Publications



Transforming unstructured data into knowledge

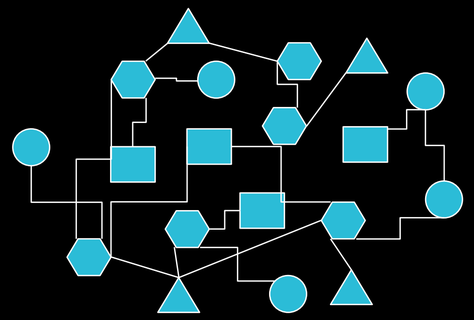
Unstructured data is stored in a complex Knowledge Graph that captures all the knowledge in the text, in the practical experience & from physics/chemistry principles.

NLP
ML



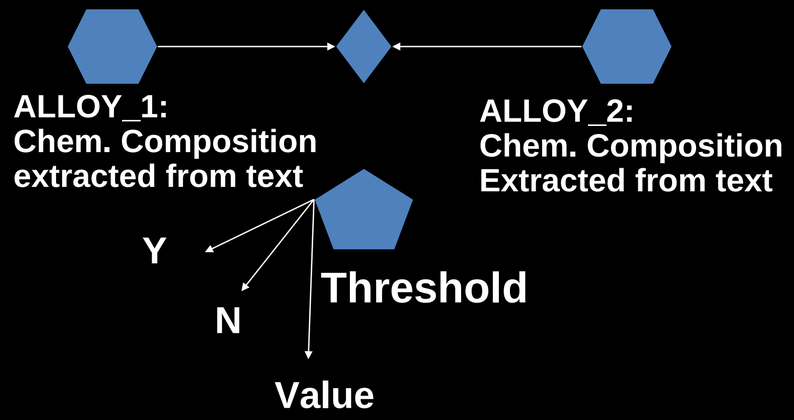
Chemistry
Physics
Simulation

Knowledge Graph



- **ALLOY NODE TYPE**
- **PROCESS NODE TYPE**
- **PRODUCT NODE TYPE**
- **ELEMENT NODE TYPE**

SIMULATION



IBM Watson Discovery Advisor

History Help

Enter some keywords separated by commas.

Ask

Advanced

Query

SELECT ELEMENTS

SELECT PROPERTIES

SELECT FORM

KEYWORDS

Query (Unordered) Query (Ordered) Prediction

Attribute distribution | Attribute correlation | Result in table | Result in Box Plots | Prediction | Document View

Review the distribution of an attribute as found in all documents (the corpus).

Select attribute to plot:

If applicable, limit to range in last query.

Frequency

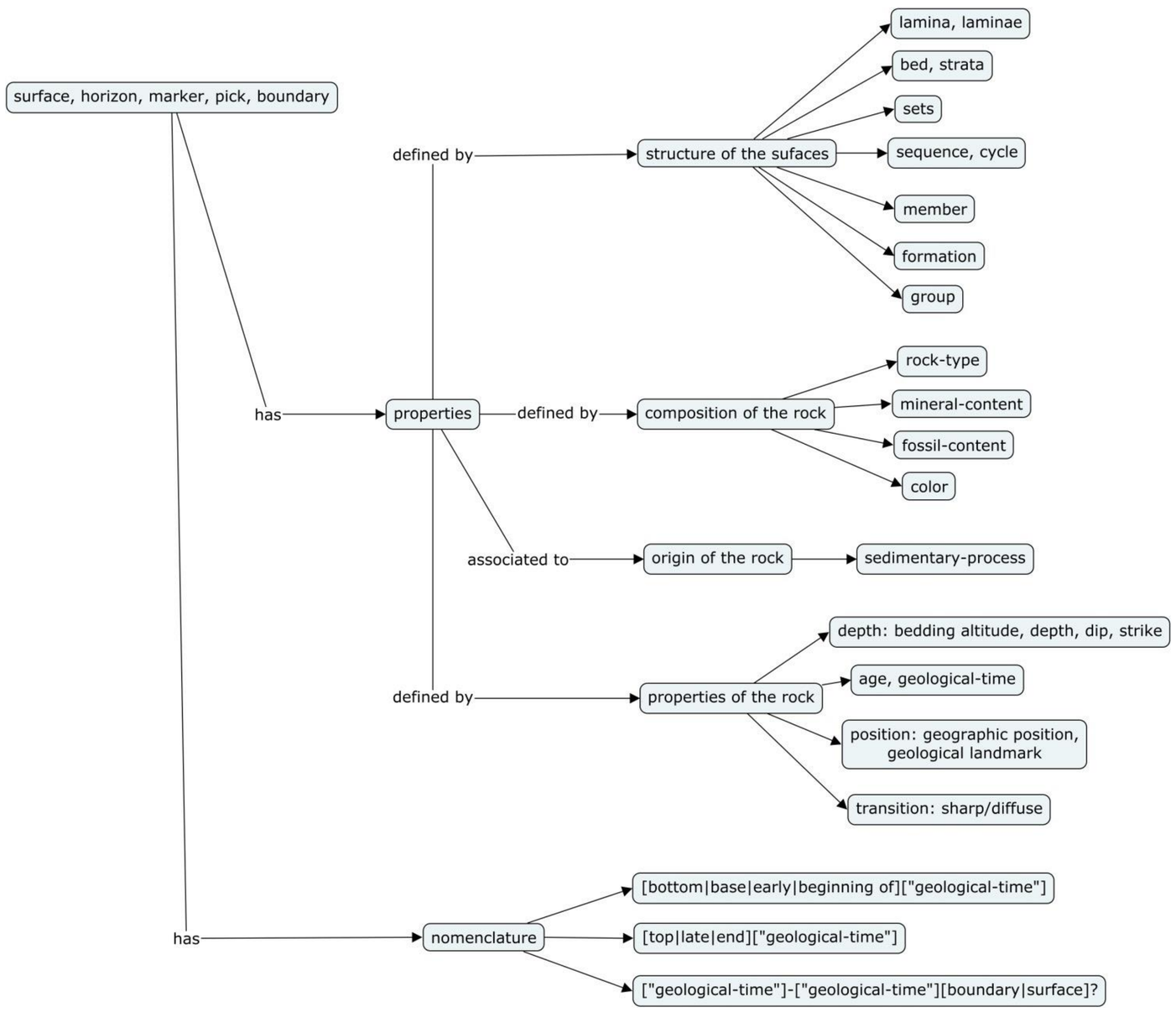
Teaching Geology to Watson

Currently working on the concept of identifying petroleum basin analogues. Complex decision process driven by

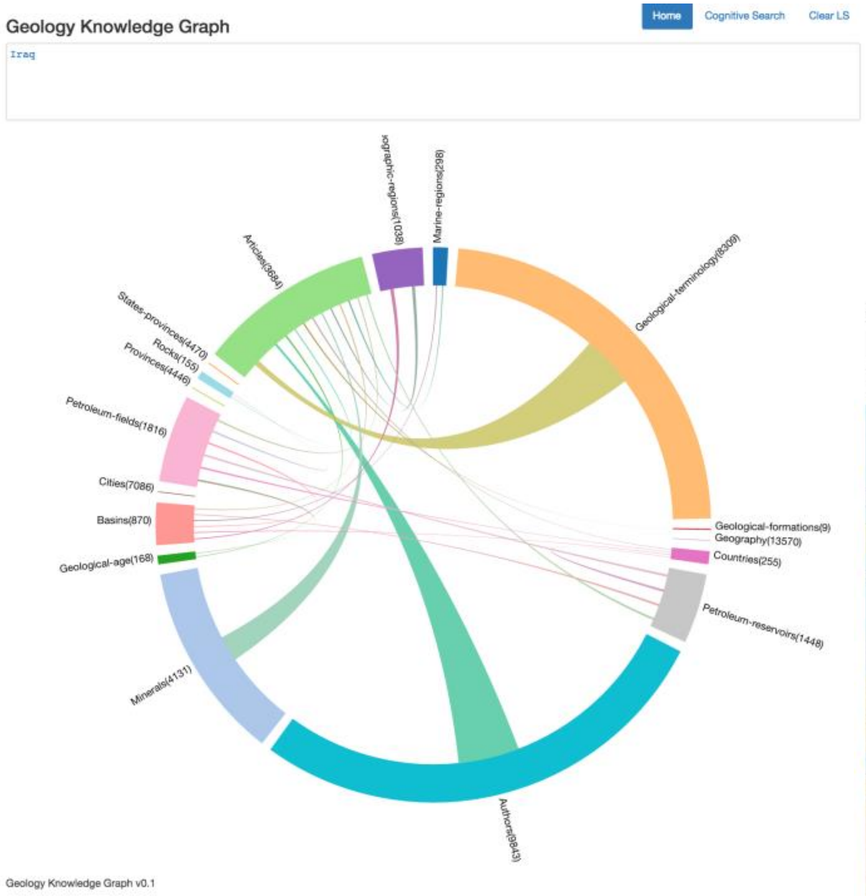
- Structure of rocks
- Composition of formations
- Origin
- Properties

Work based on:

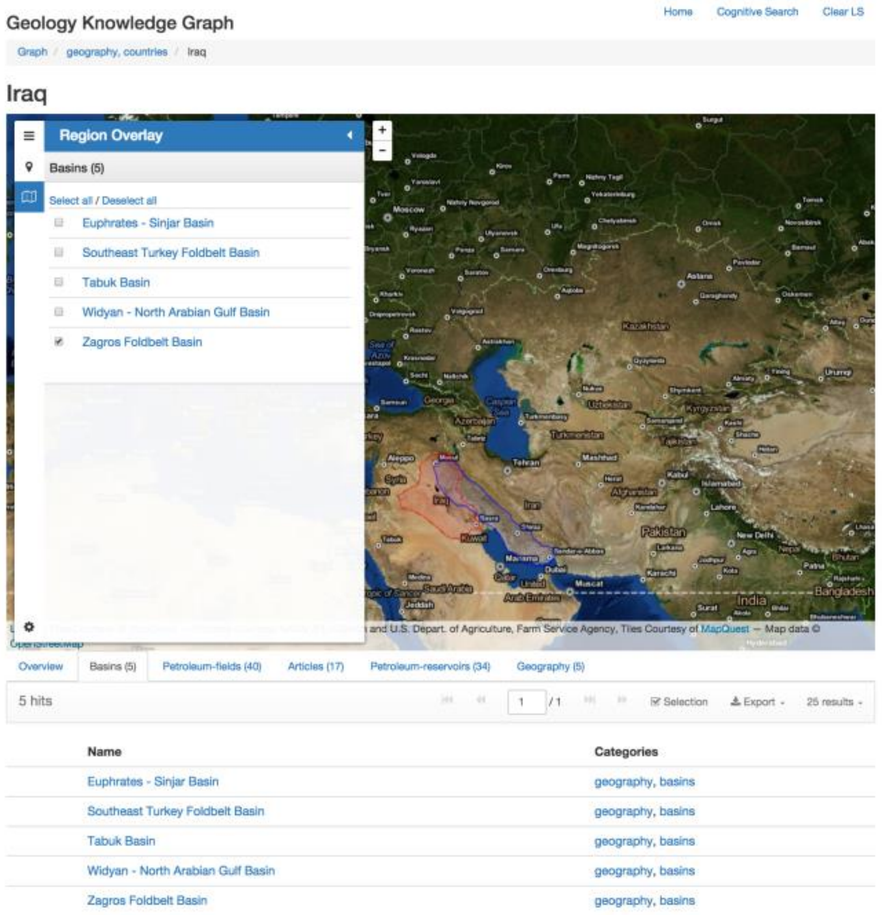
- Advanced semantic extraction form PDF documents
- Cognitive representation of the decision processes of Oil & Gas geologists



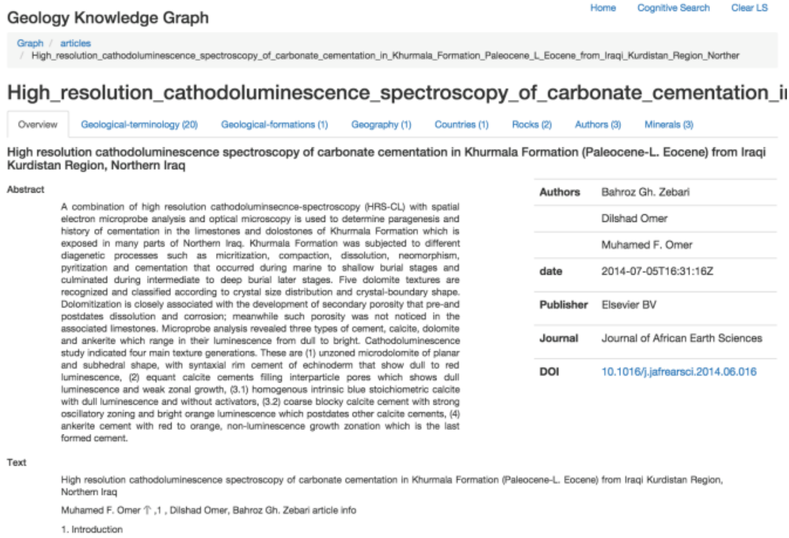
Teaching Geology to Watson



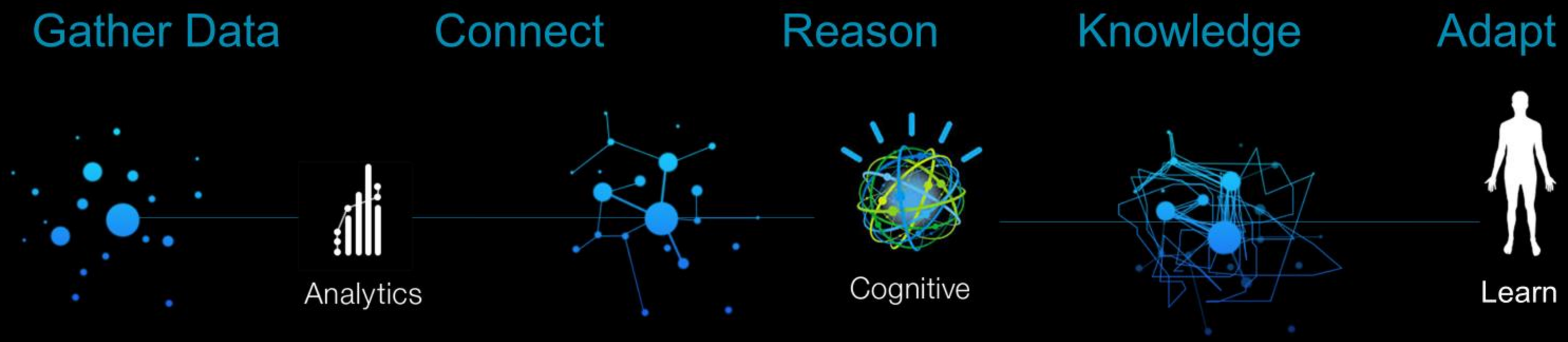
User Interface for exploring Knowledge Graph



Advanced Exploration



Data Driven Knowledge Discovery Pipeline



Cognitive Computing for Healthcare

Healthcare Industry is dealing with data overload

Exogenous data
(Behavior, Socio-economic, Environmental, ...)

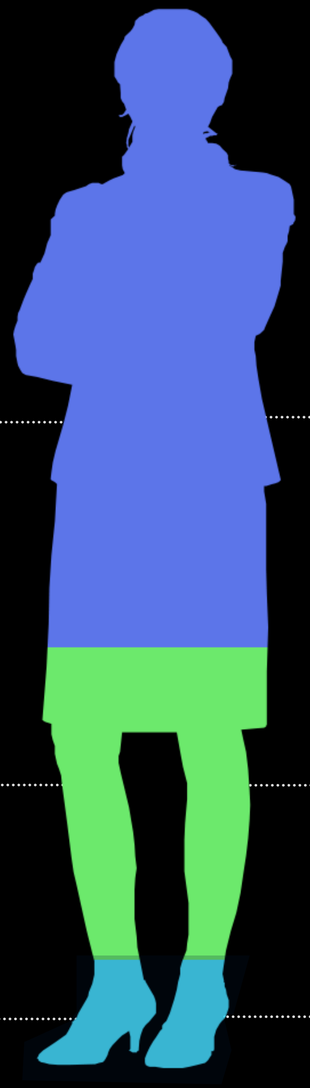
60% of determinants of health
Volume, Variety, Velocity, Veracity

Genomics data

30% of determinants of health
Volume

Clinical data

10% of determinants of health
Variety



1100 Terabytes
Generated per lifetime

6 TB
Per lifetime

0.4 TB
Per lifetime

Source: "The Relative Contribution of Multiple Determinants to Health Outcomes", Lauren McGover et al., Health Affairs, 33, no.2 (2014)

IBM Watson for Oncology | Dr. Stone | Feedback | Notices | IBM Watson

New Patient

Sample Patient Cases for Evaluation

Filter

Sort by:

Range: 1-10 Total: 49 | 1 2 3 ... 5 | 10 | 25 | 50 | All

Bryan, Mr. Eric
Diagnosis: Colon Cancer, Gender: M, Age: 73, Patient ID#: Example 91505762-75fb-4a13-a336-25d641db57a6
Last Updated: 1/22/2015, 1:46:45 PM

Brown, Mrs. Christine
Diagnosis: Breast Cancer, Gender: F, Age: 64, Patient ID#: Example-gh36q7-g45h-ahe1-4562sk7a
Last Updated: 1/22/2015, 1:46:37 PM

Cary, Mrs. Olivia
Diagnosis: Breast Cancer, Gender: F, Age: 67, Patient ID#: Example-gf56f5-hy67-a45b-aeh1-cvgw89a
Last Updated: 1/22/2015, 1:46:18 PM

Jackson, Mr. Larry
Diagnosis: Colon Cancer, Gender: M, Age: 47, Patient ID#: Example fc8bf916-3486-4d81-9af8-dc411e898f47
Last Updated: 1/21/2015, 7:53:16 AM

Lewis, Mrs. Samantha
Diagnosis: Rectal Cancer, Gender: F, Age: 68, Patient ID#: Example c10833e4-f4de-45ed-ae3a-941b36c93df9
Last Updated: 1/21/2015, 7:52:29 AM

Wu, Mrs. Alice
Diagnosis: Breast Cancer, Gender: F, Age: 66, Patient ID#: Example d04b4081-c268-43fb-9d69-c77e7d20daad
Last Updated: 1/20/2015, 1:40:41 PM

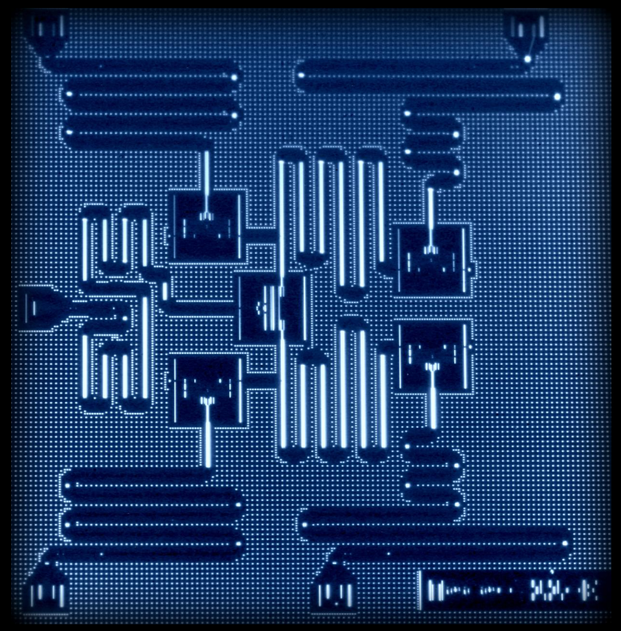
Bryan, Mr. Eric
Diagnosis: Colon Cancer, Gender: M, Age: 73, Patient ID#: Example 91505762-75fb-4a13-a336-25d641db57a6
Last Updated: 1/20/2015, 1:40:27 PM

Range: 1-10 Total: 49 | 1 2 3 ... 5 | 10 | 25 | 50 | All

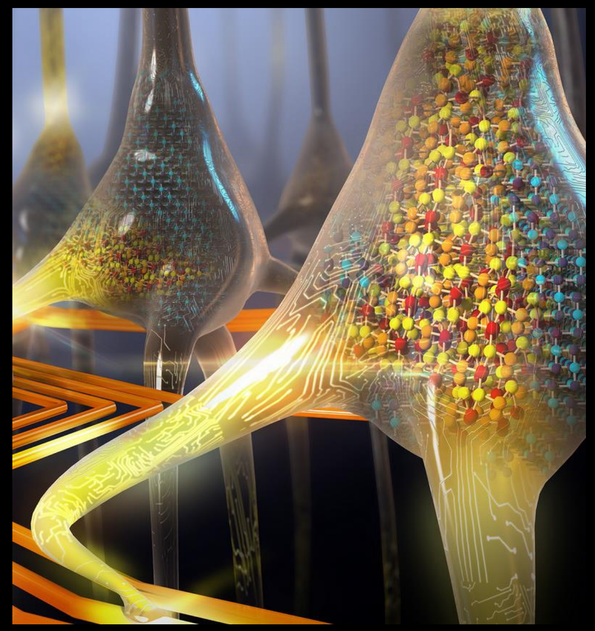

POLITICS
DOMANI SERA ORE 21.10
Rai 3 HD
PRESA DIRETTA

POLITICS
DOMANI SERA ORE 21.10

Watson in the Future:
Non-Von Neuman for Next Generation Cognitive Applications



Quantum Computing



Neuromorphic Computing

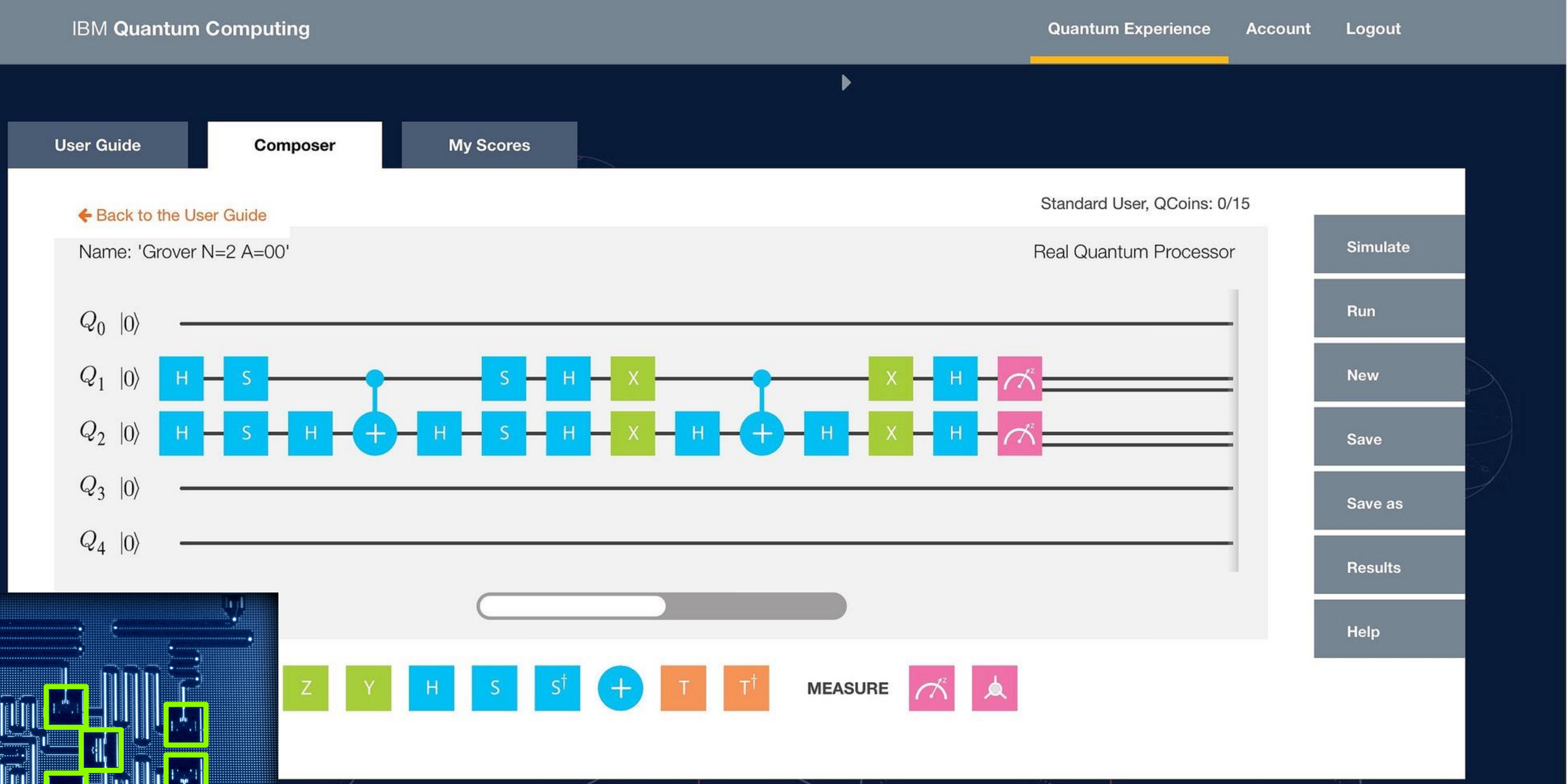
Classic vs Quantum Computer

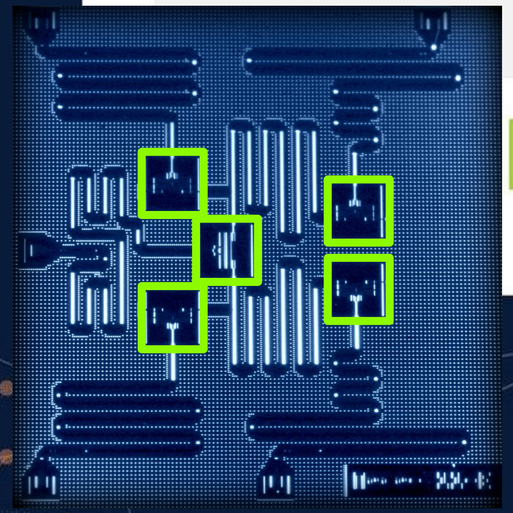
- A classical computer makes use of bits to process information, where each bit represents either a 1 or a 0.

A quantum bit (qubit) can represent a 1, a 0, or both at once, which is known as superposition.

This property along with other quantum effects enable quantum computers to perform certain calculations vastly faster than is possible with classical computers.

Classic Bit Qubit





Program a Qubit in the IBM Cloud Today!

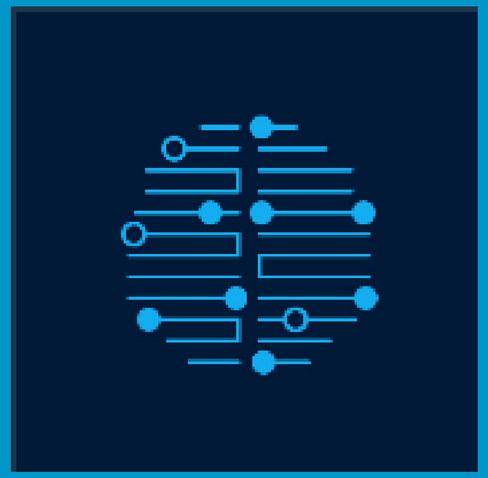
https://quantumexperience.mybluemix.net

Applications for Quantum

- With Moore's Law running out of steam, quantum computing will be among the technologies that IBM believes will usher in a new era of innovation across industries.
- This leap forward in computing could revolutionize:

Cryptography

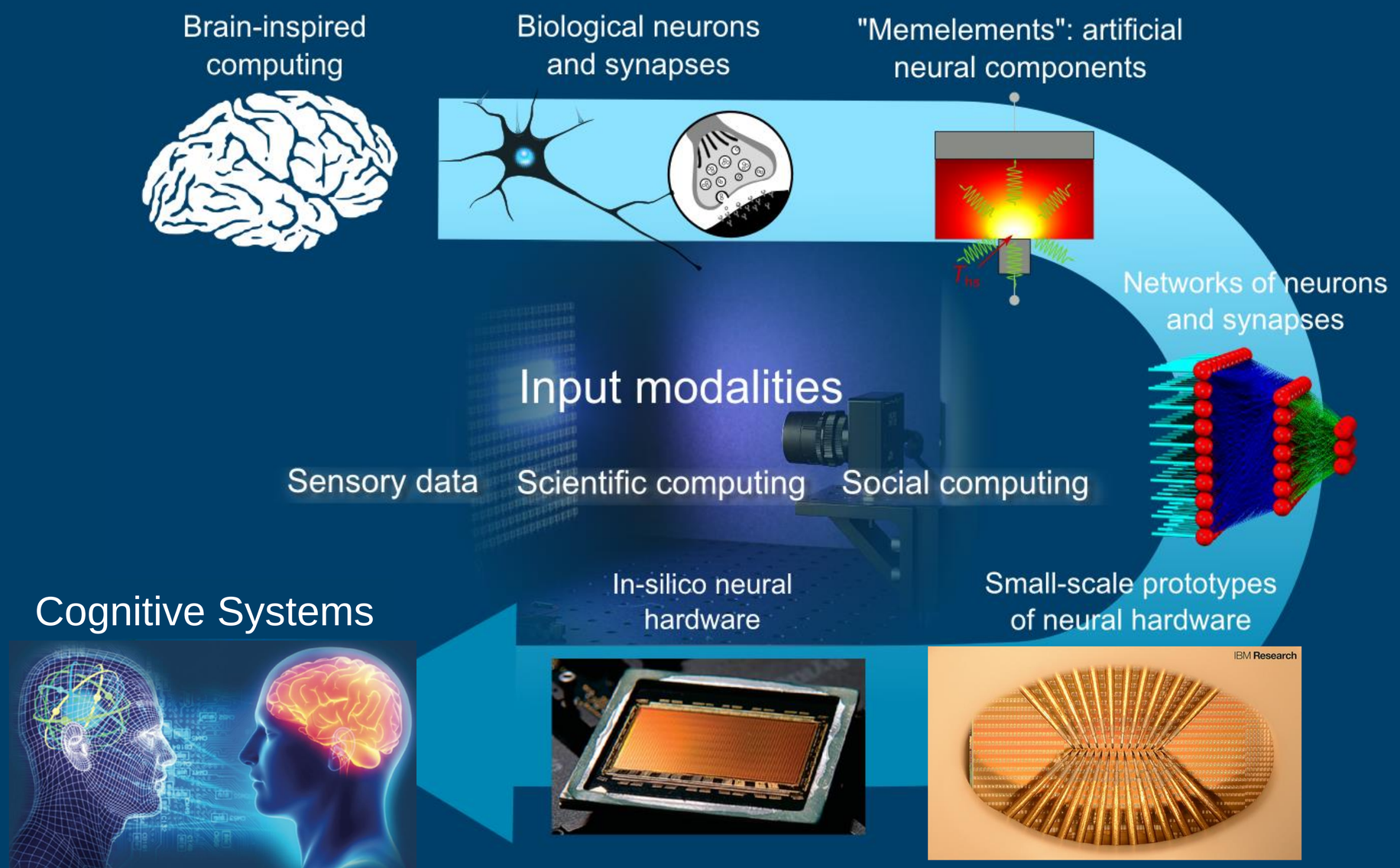
Medicine & Materials



Machine Learning

Searching Big Data, Pattern Recognition, IoT

Neuromorphic Computing



TrueNorth Chip (SyNAPSE)

	2011	Now
Programmable Neurons	256	1 million
Programmable Synapses	262,144	256 million
Neurosynaptic Cores	1	4096

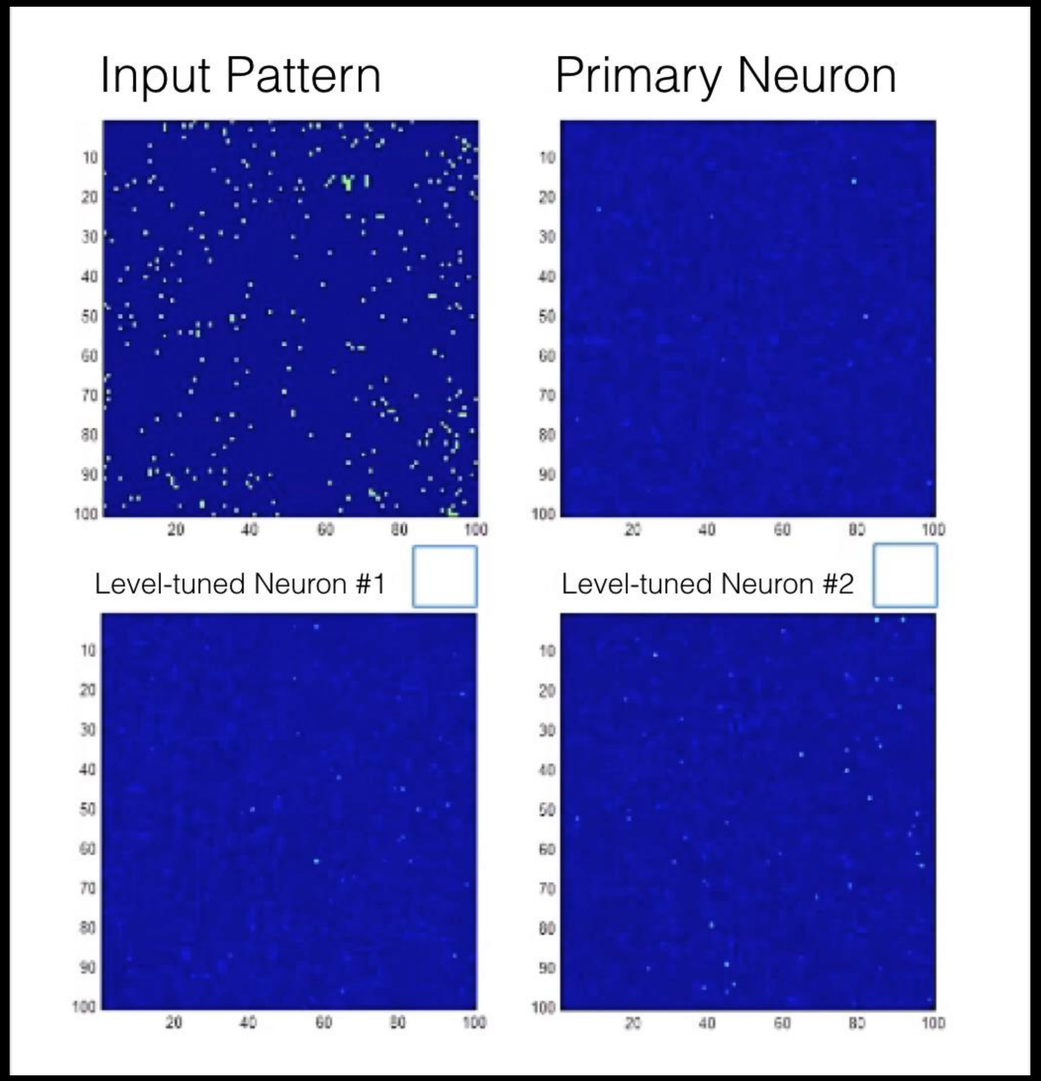
Saliency

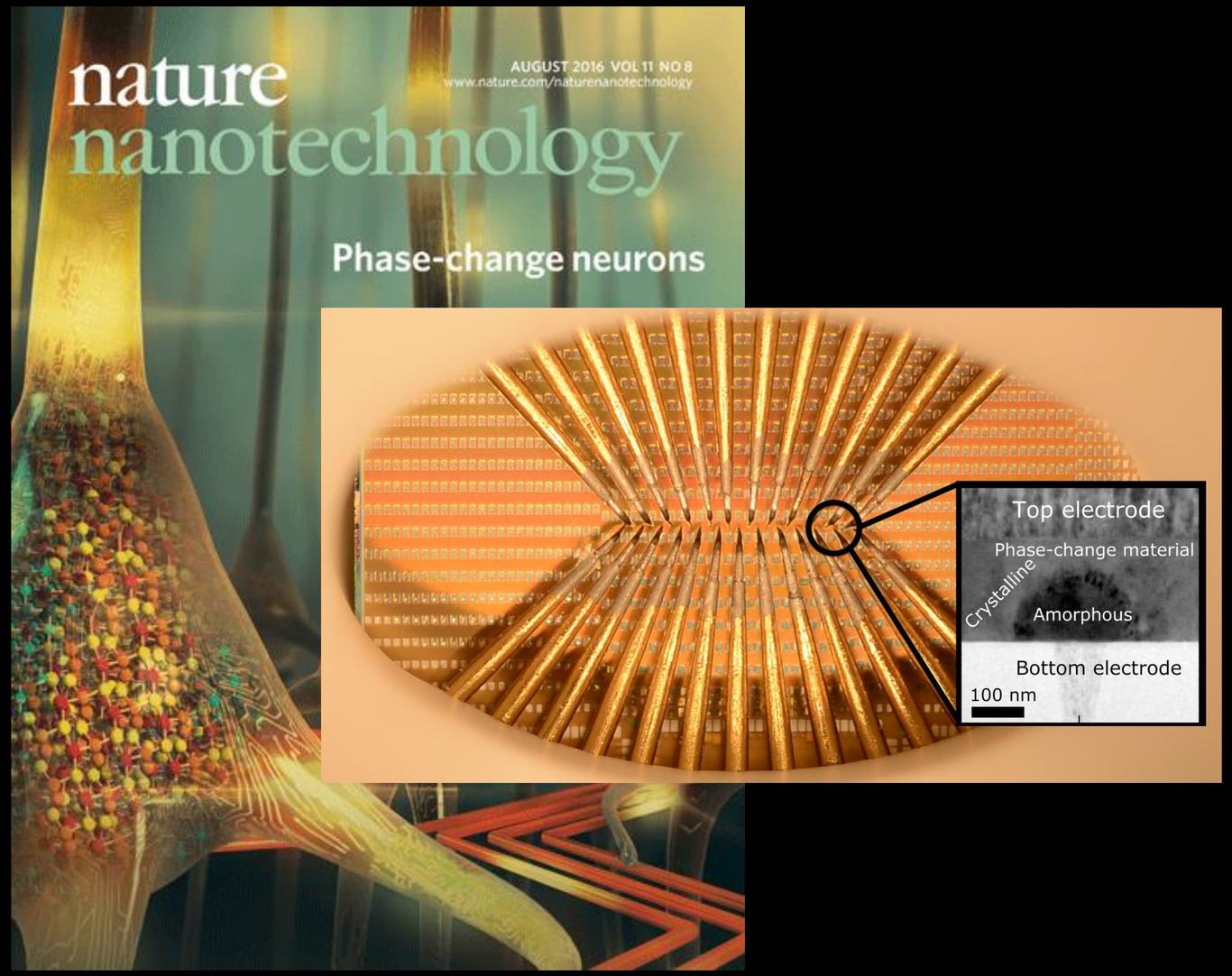
Saliency + Classification

Object Centers

Output

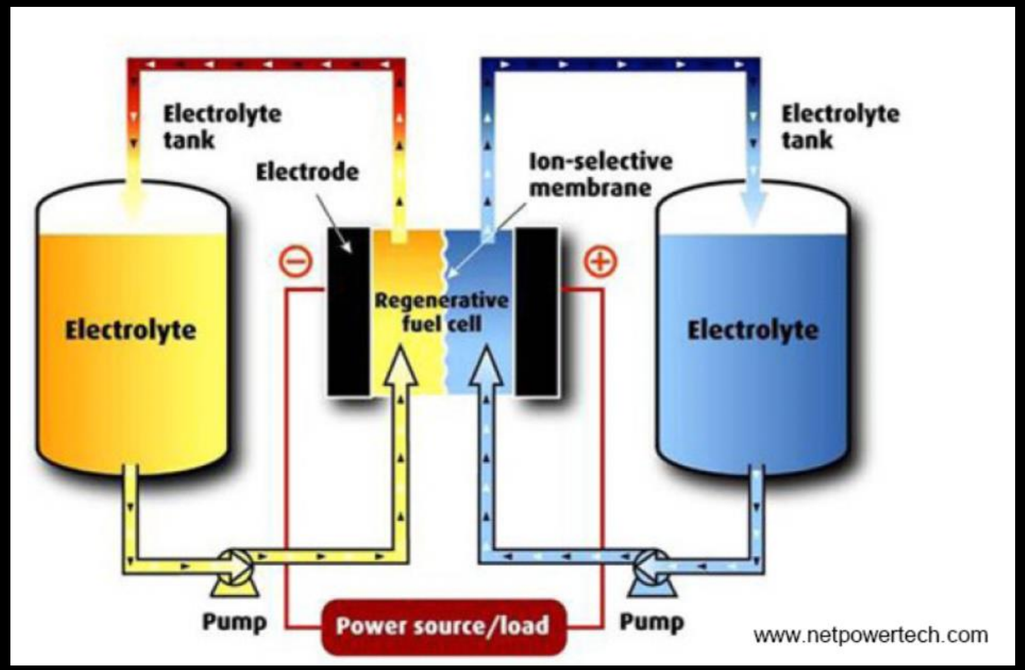
Detecting Correlations with a Spiking Neural Network

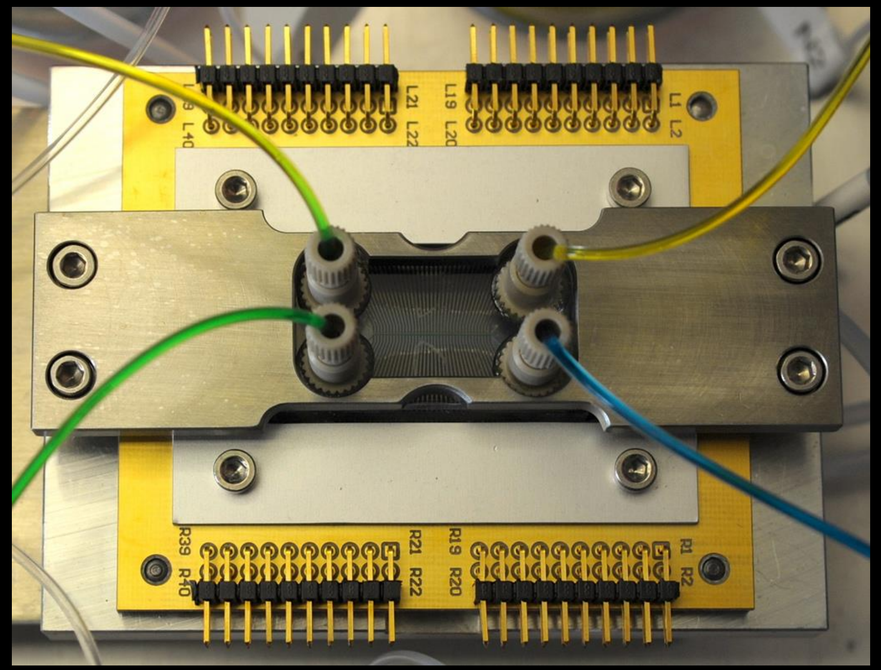


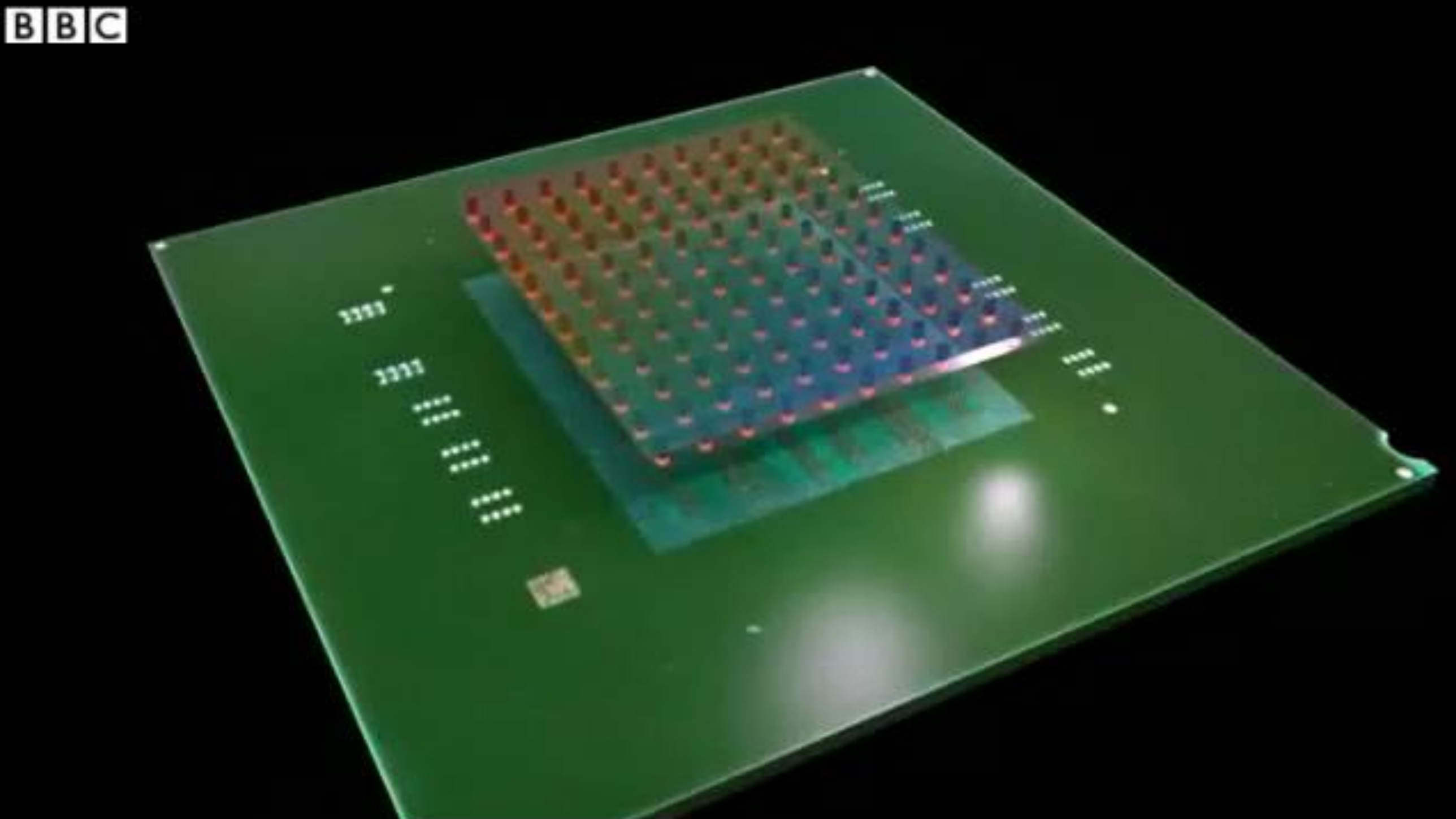


Brain Inspired Computing: Electronic Blood

- 98% of the energy of a computer is for cooling
- Liquid removes heat 4000x more efficiently than air
- The brain is powered & cooled using liquid, can we do the same for computers?
- The result: a 1 PetaFlop supercomputer in 10 liters






BBC

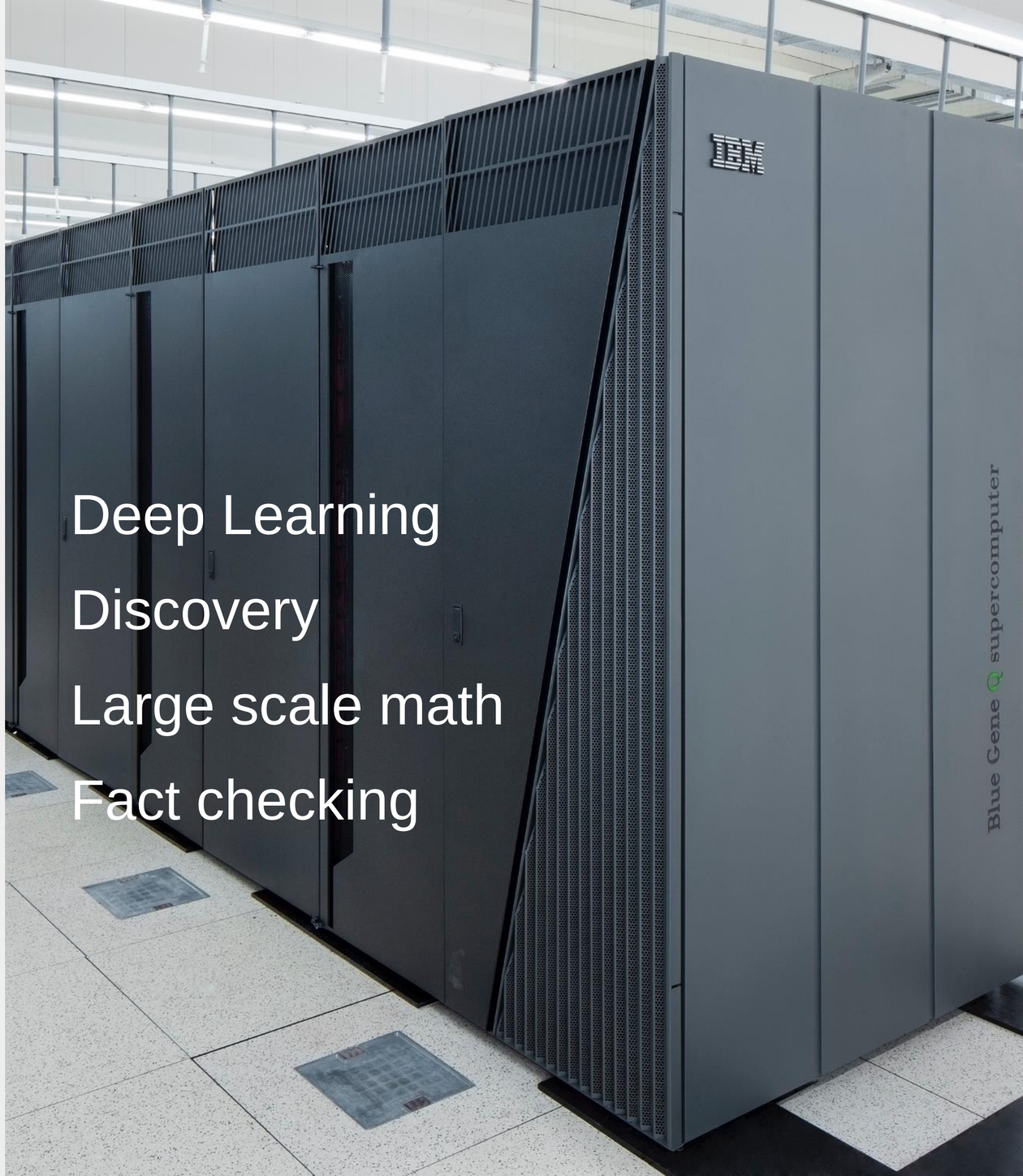
Compassion

Intuition

Design

Value judgements

Common sense



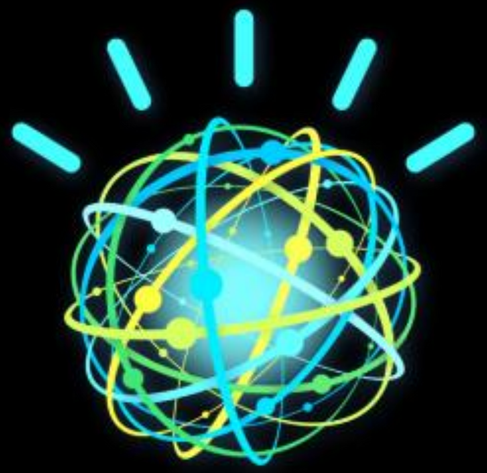
Deep Learning

Discovery

Large scale math

Fact checking

Human + Machine



Questions?