Cognitive Discovery: The Next Frontier of R&D

Dr. Alessandro Curioni
IBM Fellow, Vice President, Europe, and
Director IBM Research - Zurich

@Ale_Curioni

IBM Research: World-Class Capability with **3,000** Scientists


Six Nobel Laureates


Ten Medals of Technology


Five National Medals of Science


Three Kavli Prizes


Six Turing Awards

69 Members

123 IEEE Fellows

28 ACM Fellows

98 IBM Fellows

Ireland

Zurich

Almaden

Watson

Austin

Haifa

Tokyo

China

India

Kenya

IBM invested $5.7B on R&D in 2016

Brazil

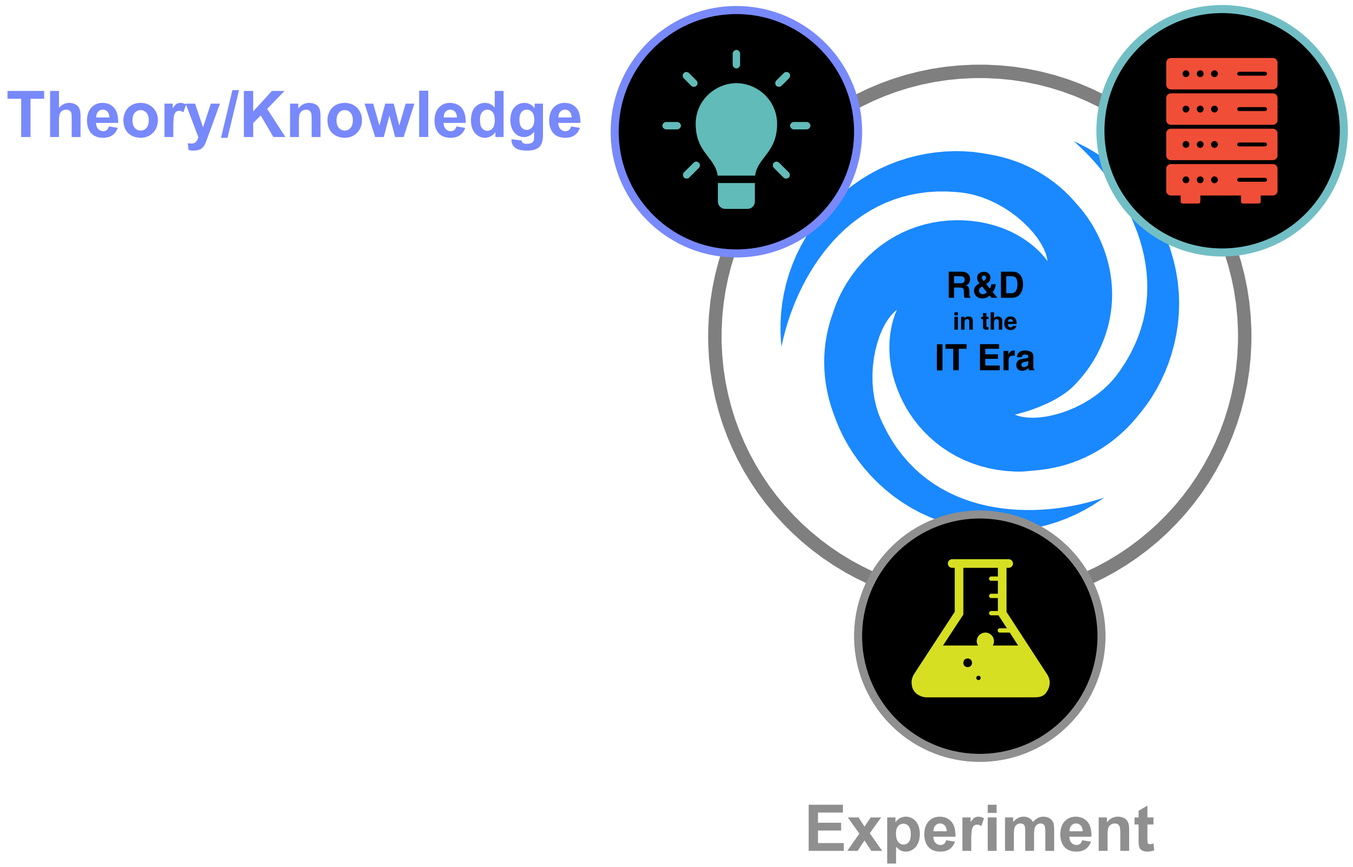
South Africa

Australia

IBM **Research** | Zurich

- Established in 1956
- 45+ different nationalities
- Open Collaboration:
  - Horizon2020: 43 funded projects and 500+ partners
- Two Nobel Prizes:
  - 1986: Nobel Prize in Physics for the invention of the scanning tunneling microscope by Heinrich Rohrer and Gerd K. Binnig
  - 1987: Nobel Prize in Physics for the discovery of high-temperature superconductivity by K. Alex Müller and J. Georg Bednorz
- Binnig and Rohrer Nanotechnology Centre opened in 2011 (Public Private Partnership with ETH Zürich and EMPA)
- 7 European Research Council Grants

Technical R&D today: The three pillars



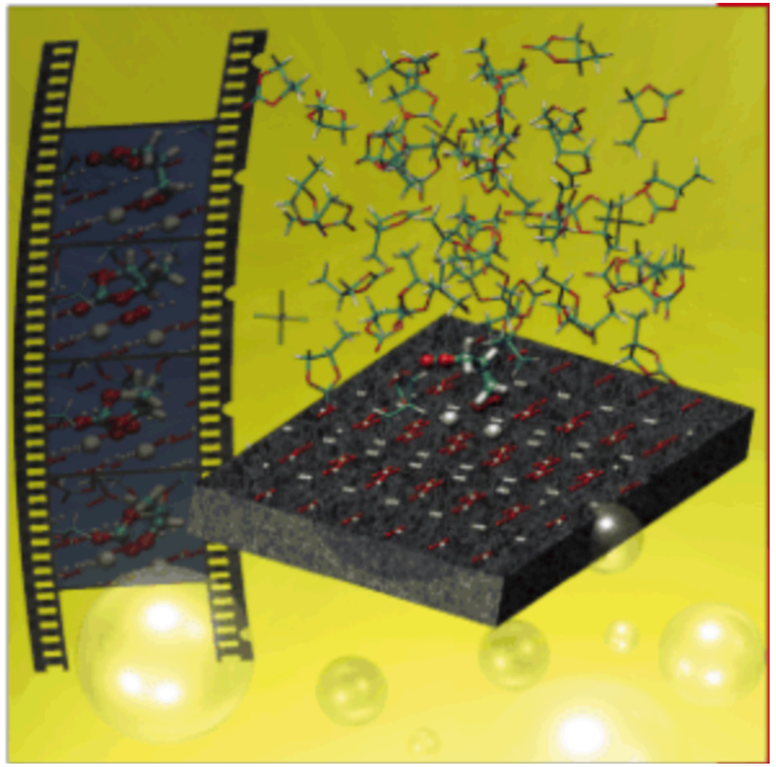
Massive improvements:

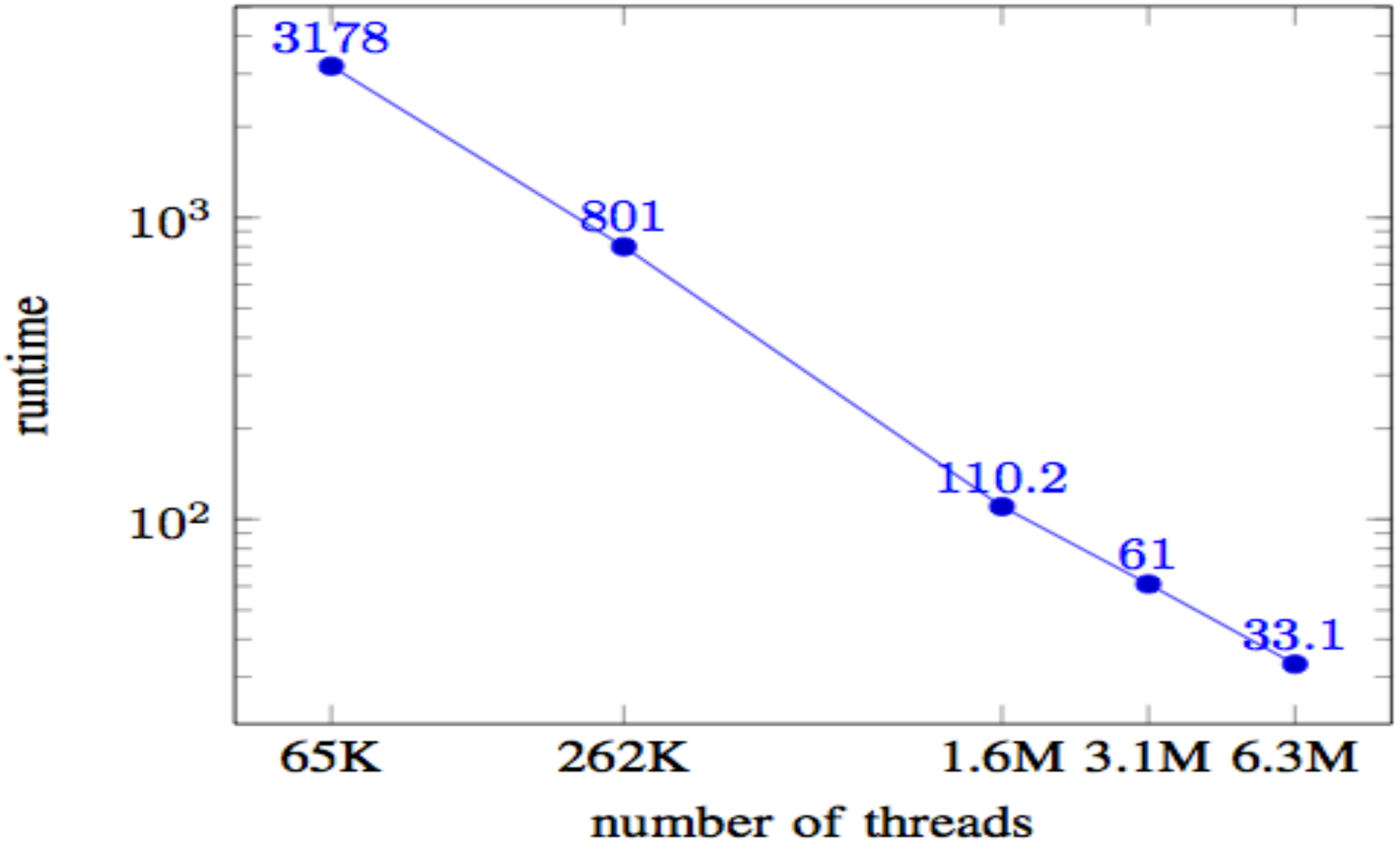
- Applicability & ease of use
- Simulation fidelity
- Scalability
- Throughput

thanks to Supercomputing

Success in Petascale computing: Study what was previously impossible: Discover new electrolytes for LiAir batteries

Implementing Exact-Exchange in CPMD
>99% Parallel Efficiency to over 6.2M threads
Studying Li-Air Batteries, 1736 atoms, 70Rv cuttof





V. Weber, T. Laino, C. Bekas, A. Curioni, A. Bertsch, S. Futral IPDPS 13

Source: Top500.org

Success in Petascale computing: CFD can achieve Linpack like sustained performance

ACM Gordon Bell Prize 2013

14.4 PFLOP/S @73% of peak perf., with I/O

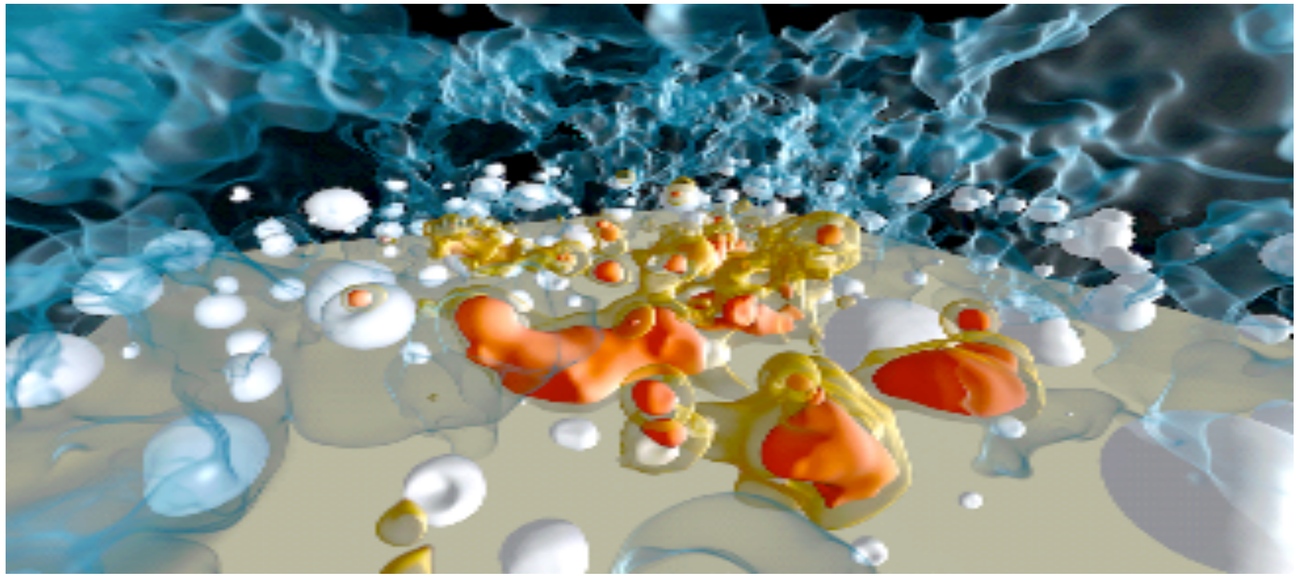
2 orders of magnitude improvement in

- scale of the problem (from 128 to 15K bubbles)
- time to solution

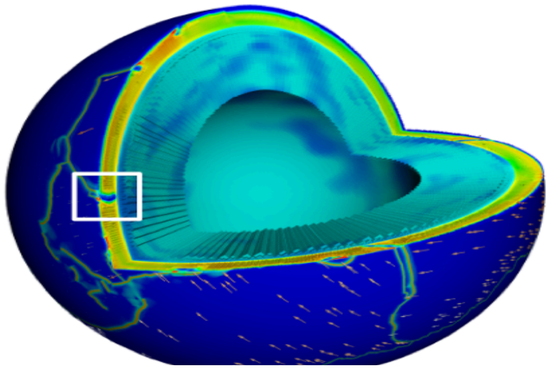
Compute specifics:

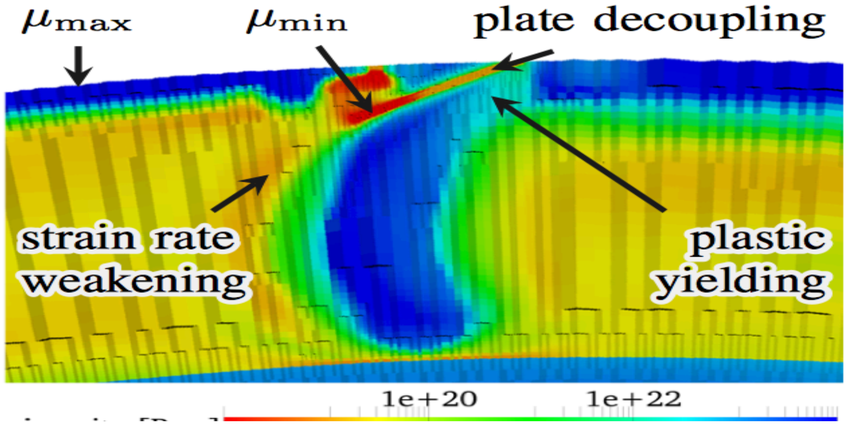
13 Trillion elements, 1.2TBytes compressed I/O per time step, 6.4 M threads

IBM, ETHZ, TUM, LLNL



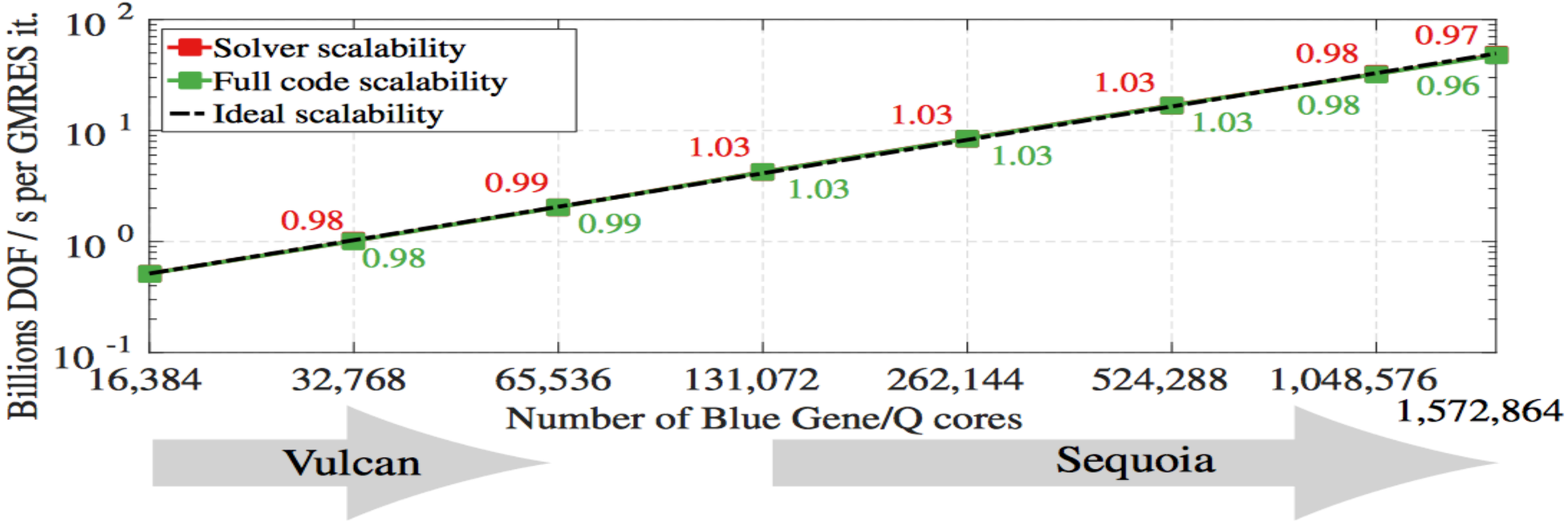
Success in Petascale computing: Implicit linear solvers do scale!



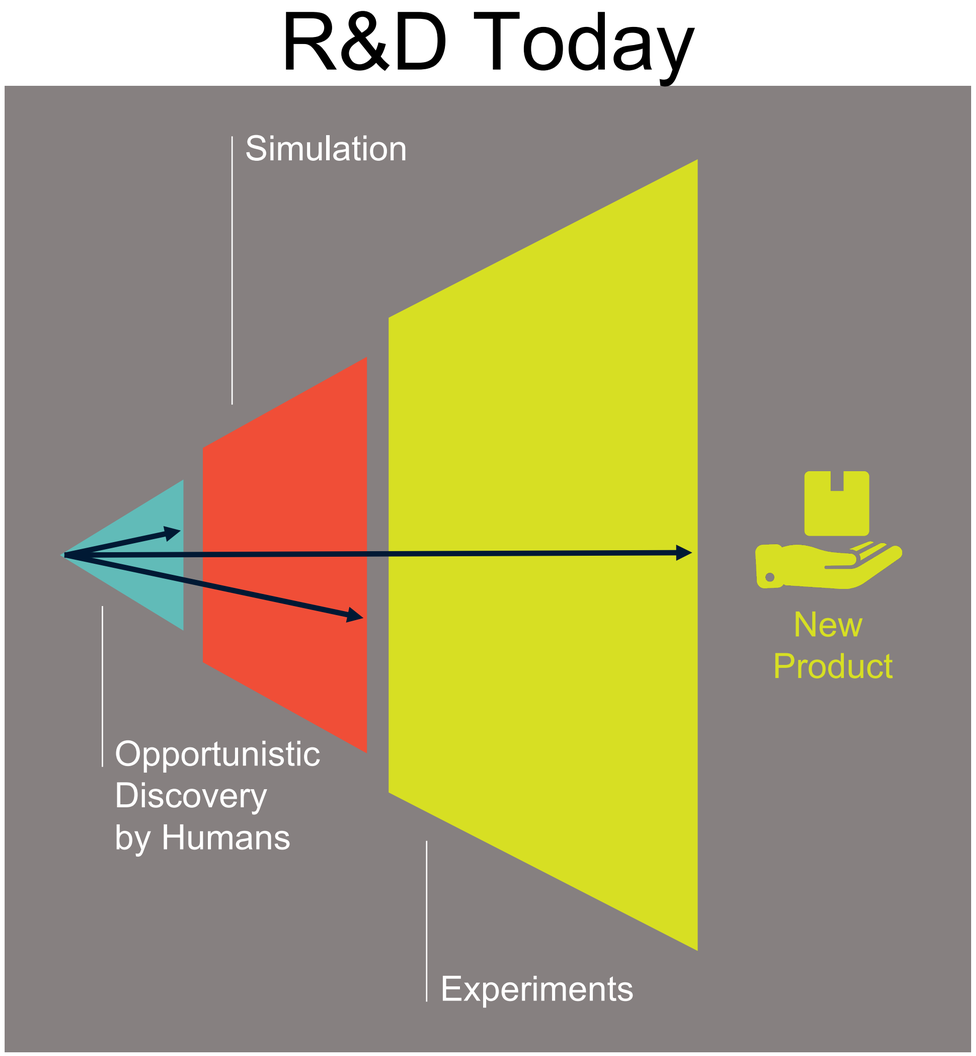


ACM Gordon Bell Prize 2015
97% of sustained scalability for a fully implicit solver. 1.6M cores
3.2M MPI processes
602B DoF,

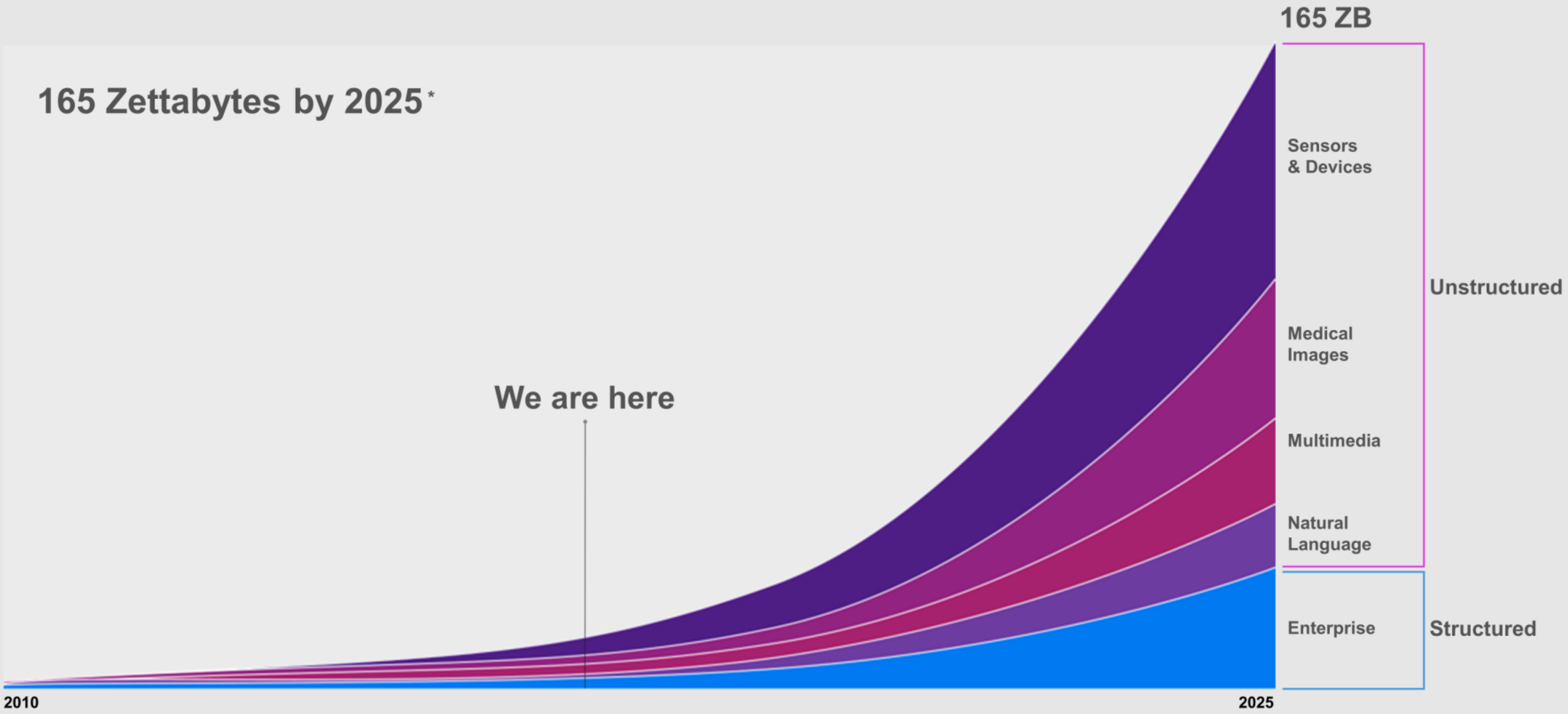
IBM. UT Austin. NYU. CALTECH



But we cannot beat complexity with brute force simulation. Traditional discovery has limits: We need a new, data driven, holistic approach

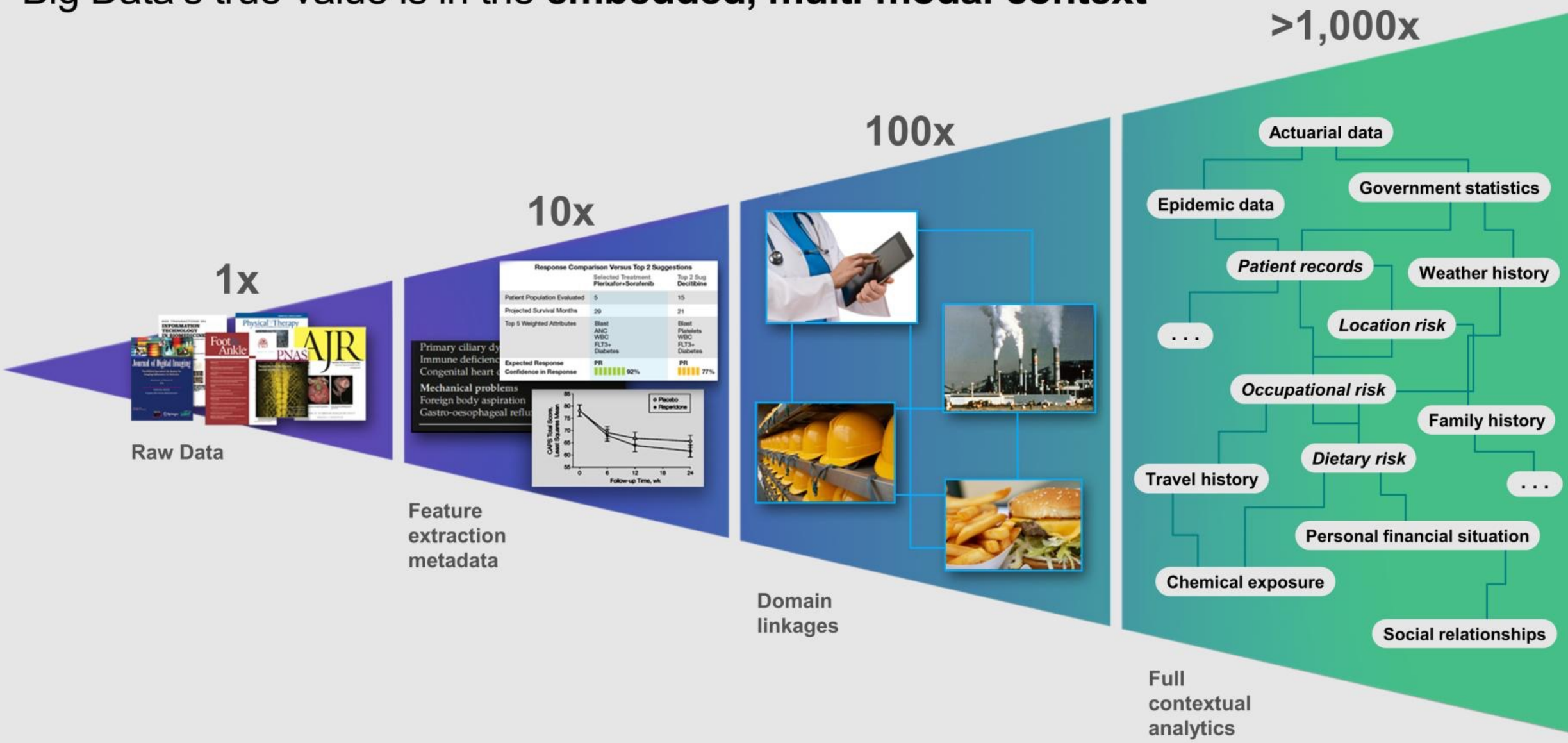


Data is **transforming every industry**



*Source IDC. IBM projections based on analyst report

Big Data's true value is in the **embedded, multi-modal context**



The next revolution is to **augment human cognition**

The Industrial Revolution

Steam and Railways

Steel, Electricity and Heavy Engineering

Oil, Automobiles and Mass Production

Information and Telecom

Cognitive Era

Augmenting Human Intelligence

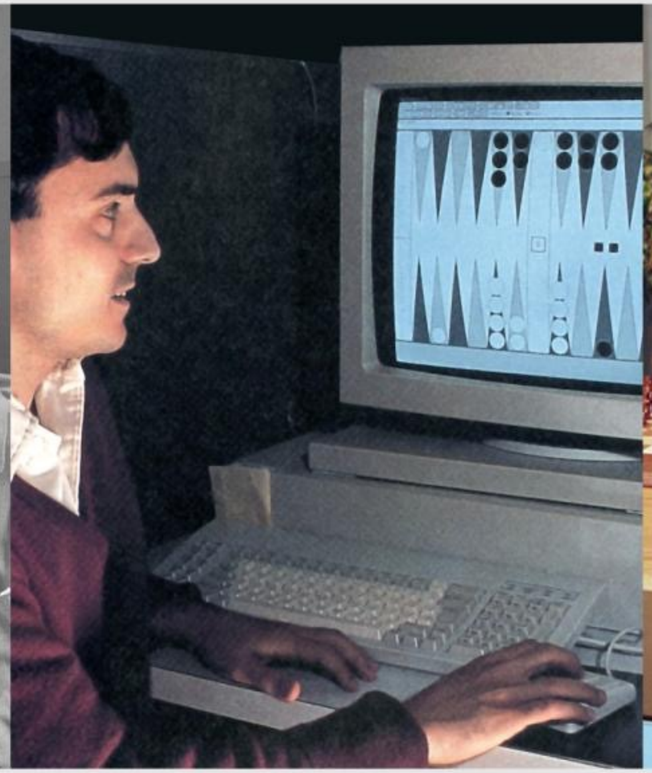
- **Learn at scale**
- **Reason with purpose**
- **Interact naturally with us**

1700 **Today**

Source: Adapted from the work of Carlota Perez

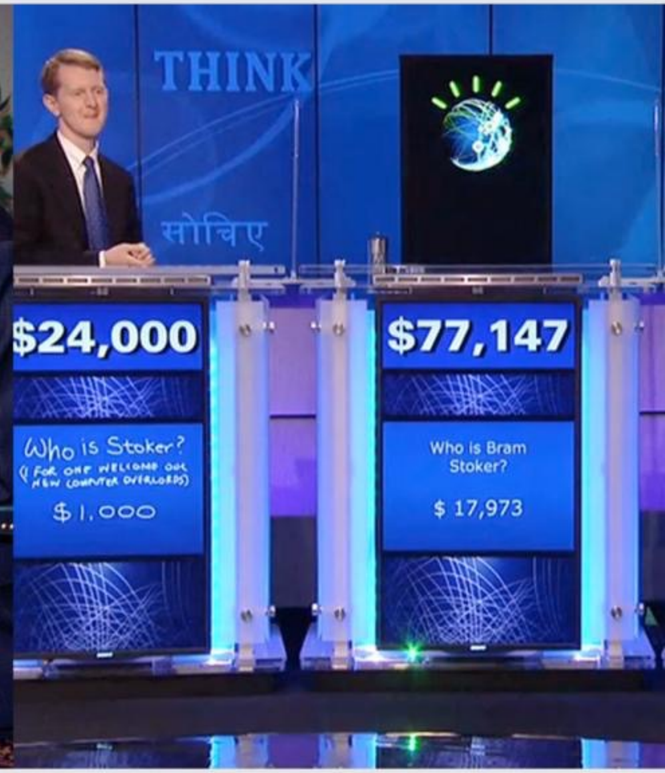
IBM AI advances through **human + machine** grand challenges

Checkers program
Arthur Samuel, 1956



Backgammon
Gerald Tesauro, 1993

Chess / Deep Blue
Murray Campbell, 1997



Jeopardy! / Watson
2011

Watson in **2011**

System Specifications

2880 Processing Cores

90 IBM P750 Servers

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

The Watson that competed on *Jeopardy!* in **2011** comprised what is now a single API—**Q&A**—built on **five underlying technologies**.

Natural Language Processing

Machine Learning

Question Analysis

Feature Engineering

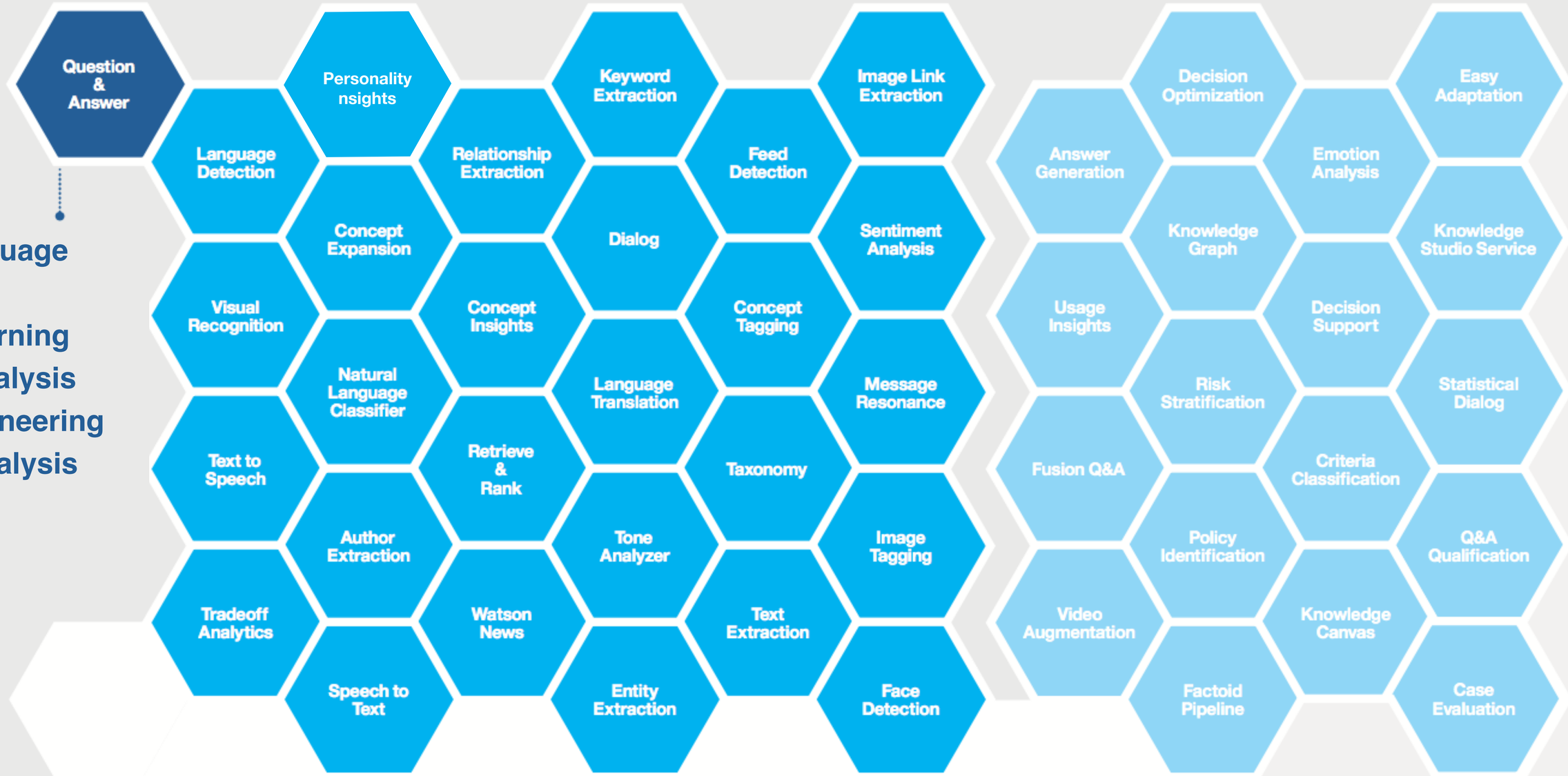
Ontology Analysis

The Watson that competed on *Jeopardy!* in **2011** comprised what is now a single API—**Q&A**—built on **five underlying technologies**.

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

In 2017 there are more than **60 Watson APIs** in the IBM Cloud

Natural Language Processing
Machine Learning
Question Analysis
Feature Engineering
Ontology Analysis



A vision of the **future:**
Everyone who needs expertise will have a **cognitive assistant**

Healthcare
Surface best protocol options for practitioners

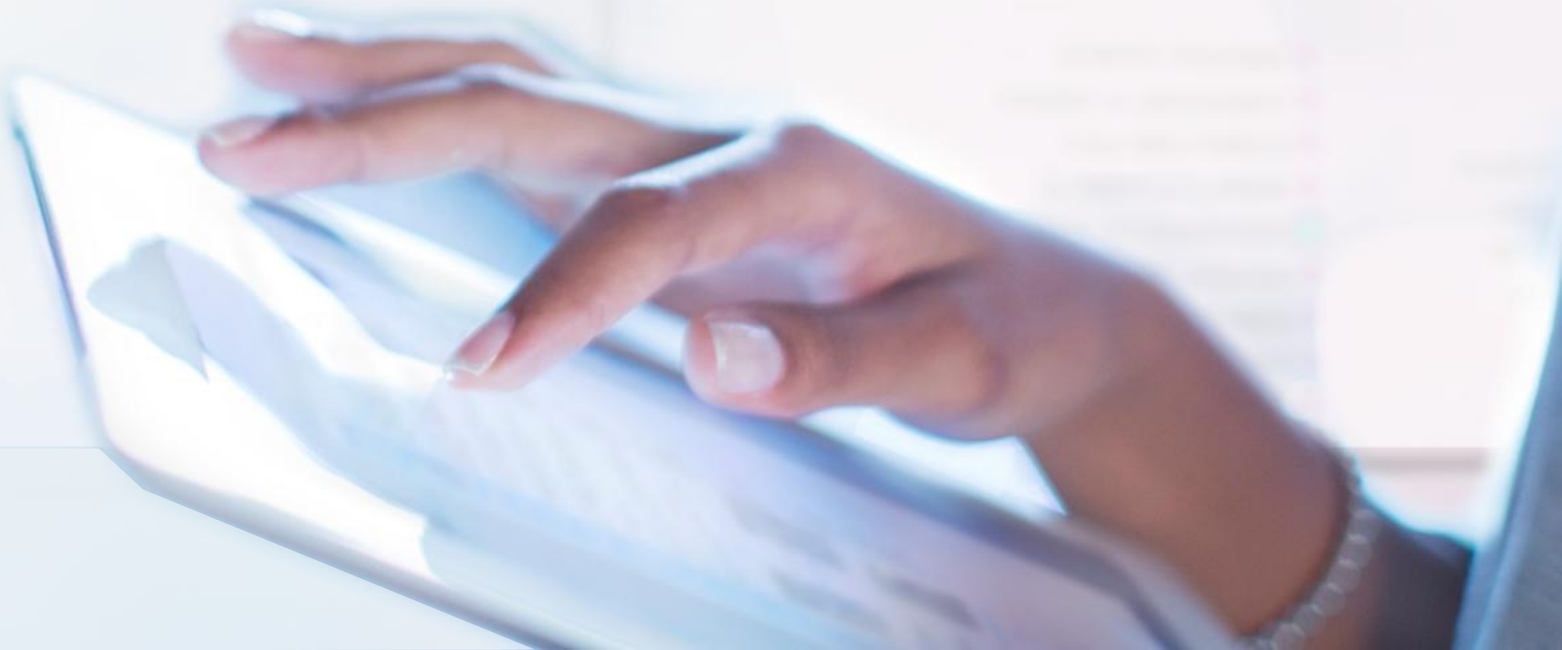
Finance
Enhance portfolio analysis and risk management

Education
Deliver personalized programs for students & teachers

Business Decisions
Analyze complex scenarios and support strategic decisions

Research & Development
Change ages old practices

"Before I recalculate the findings, would you like to hear about the other important factors that may impact your decision?"



A vision of the **future:**
Everyone who needs expertise will have a **cognitive assistant**

Healthcare
Surface best protocol options for practitioners

Finance
Enhance portfolio analysis and risk management

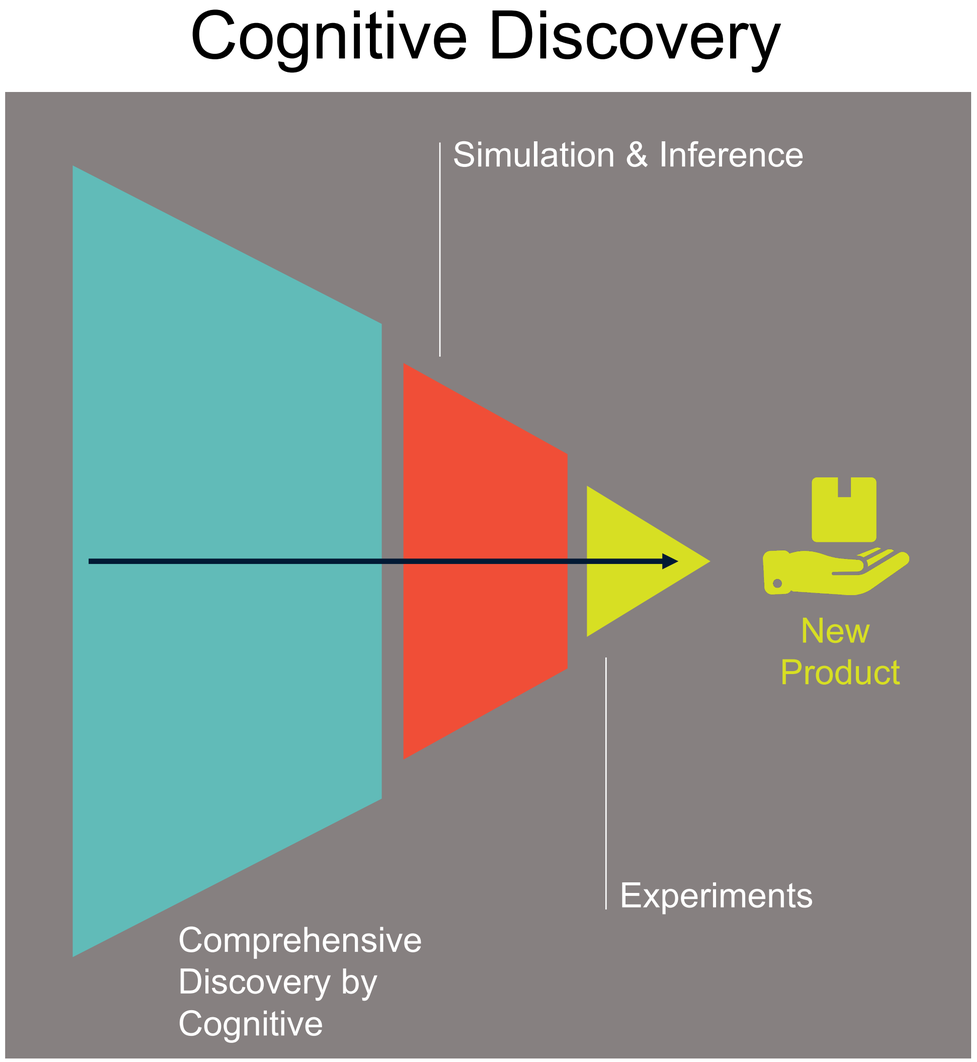
Education
Deliver personalized programs for students & teachers

Business Decisions
Analyze complex scenarios and support strategic decisions

Research & Development
Change ages old practices

"Before I recalculate the findings, would you like to hear about the other important factors that may impact your decision?"

Cognitive Discovery: Knowledge based. Uses A&I to target which simulations to run and the best way to execute them



Technical R&D today: Disruption opportunity

Knowledge

Companies need to easily access quickly growing and widely diverse information sources.

- Highly unstructured/dark
- Current human based approach not scalable

Inference & Simulation

Domain related inference is largely missing. Setting up and deploying the right simulations is very hard.

- Human capital intensive, non scalable

Evidence & Experiments

Internal evidence and experiments are driven primarily empirically, often brute force, and their results are isolated from wider knowledge space.

The way forward: Cognitive Discovery

Knowledge

Create technical area specific knowledge space from all relevant sources. Link with company data.

Science, Products & Economics

Simulations

Experimental Results

Inference & Simulation

Use inference on the knowledge space & simulation on the models

- To augment the knowledge space
- Sharpen simulation models
- Make precise decisions

Cognitive Discovery

Drastically accelerate pace of systematic discovery and maximize ROI for R&D

Evidence & Experiments

Use knowledge space to

- Drastically augment internal know-how & modeling
- Focus on which experiment is relevant
- Embed results in knowledge base

Rapid and Precise Materials R&D drives new value for our clients

Pharma

Materials

Engineering & Manufacturing

The Materials Discovery Case Study

- Atomistic simulations
- Deep Learning based property prediction

Simulation & Inference

DAYS

Lab tests experiments data

Scientific literature & internal reports

Weeks

Knowledge space

Deep Search

Design alloys to avoid catastrophic failure that can lead to huge liabilities

- Corrosion
- Cracks
- Special environmental and deployment conditions

Ingesting PDF

Billions of documents
Millions of concurrent users

Pdf-parser:

- Parses the pdf-code and presents the raw data of the pdf (text-cells, embedded images and vector-graphics in consumable format)

Pdf-interpreter:

- Captures ground truth by massive Crowd-sourcing big Data system
- Uses HPC for ML-techniques (Deep Leaning), to train automatic annotation models

Semantic-representation:

- Uses HPC & Big Data systems to to obtain a semantic representation in JSON-format of the original text

Journal of Structural Geology 32 (2010) 1349–1362

Contents lists available at ScienceDirect

Journal of Structural Geology

journal homepage: www.elsevier.com/locate/jsg

Analysis of the internal structure of a carbonate damage zone: Implications for the mechanisms of fault breccia formation and fluid flow

Stefan Hausegger [a], Walter Kurz [b,*], Robert Rabitsch [a], Eva Kiechl [a], Franz-Josef Brosch [a]

[a] Institut für Angewandte Geowissenschaften, Technische Universität Graz, Rechbauerstr. 12, A-8010 Graz, Austria
[b] Institut für Erdwissenschaften, Universität Graz, Heinrichstr. 12, A-8010 Graz, Austria

ARTICLE INFO

Article history:
Received 11 November 2008
Received in revised form
17 April 2009
Accepted 27 April 2009
Available online 7 May 2009

Keywords:
Fault zone
Carbonates
High-angle fractures
Joint-bounded slices
Fault rocks
Fluid–rock interaction

ABSTRACT

A segment of the Salzach–Ennstal fault zone (Talhof fault, Eastern Alps) shows evidence for joint nucleation by layer-parallel shear, causing the formation of joint-bounded slices oriented at high angles (65–85°) with respect to the shear zone boundary (SZB). Subsequent slice rotation resulted in joint reactivation as antithetic shears, slice kinking, and breaking-up of the individual slices into smaller fragments. The latter process, due to the longitudinal constraint of slices with impeded shear zone widening, marked the transition to cataclasite formation and fault core evolution during shear localization. Cataclasites were subsequently cemented and underwent continuous shear deformation by re-fracturing. Cement precipitation from fluids therefore played a fundamental role in the evolution of the fault zone, with a cyclic change between an open and a closed permeability system during fault evolution. Stable isotope compositions ($\delta^{13}C$, $\delta^{18}O$) of fault rock cements indicate a continuous equilibration between protolith-derived fragments and cements precipitated from those fluids. This points to limited fluid amounts, only temporally replenished by meteoric water, and a hydraulic gradient that directed fluid flow from the damage zone towards the fault core.

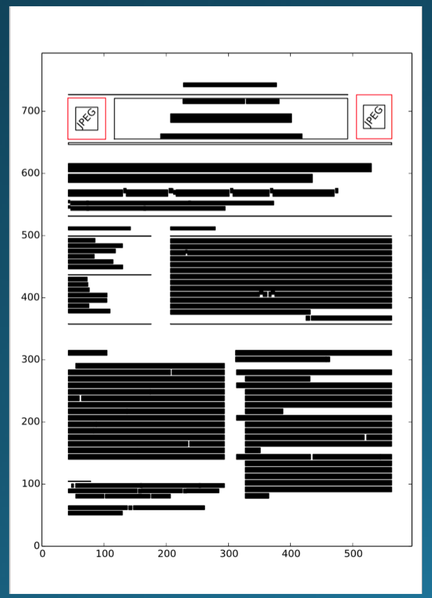
1. Introduction

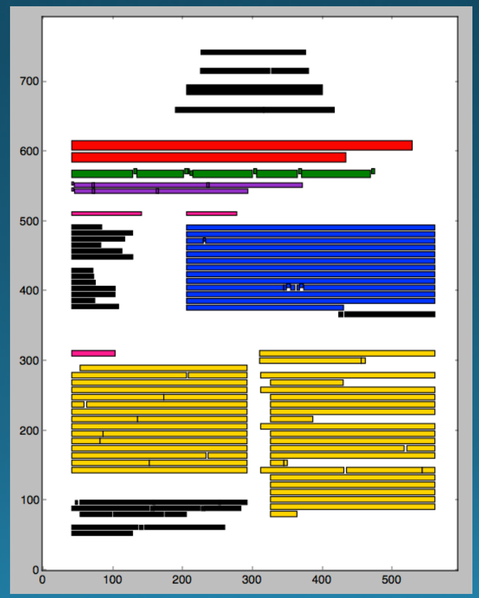
Knowledge of the internal structure of brittle fault zones has been gathered from both field studies (e.g., Anderson, 1951; Tchalenko, 1970; Sibson, 1986; Chester and Logan, 1987; Chester et al., 1993; Caine et al., 1996; Billi et al., 2003a; Faulkner et al., 2003; Wibberley and Shimamoto, 2003), and laboratory experiments (e.g., Riedel, 1929; Reches, 1978; 1983; Logan et al., 1979, 1992; Reches and Dieterich, 1983; Chester and Logan, 1990; Sagy et al., 2001; Katz et al., 2003). Fluid infiltration into faults, and the subsequent fluid–rock interaction, influence the fault mechanical behaviour (Hubbert and Rubey, 1959; Janssen et al., 1998; Kurz et al., 2008). Characterization of the internal structure of fault zones is an essential pre-requisite to understanding and predicting their mechanical, hydraulic and seismic properties (e.g., Faulkner et al., 2003; Woodcock et al., 2007). Generally, the following structural elements may be discriminated across brittle fault zones (following Chester and Logan, 1986; Caine et al., 1996; Caine and Foster, 1999; Billi et al., 2003a; Faulkner et al., 2003):

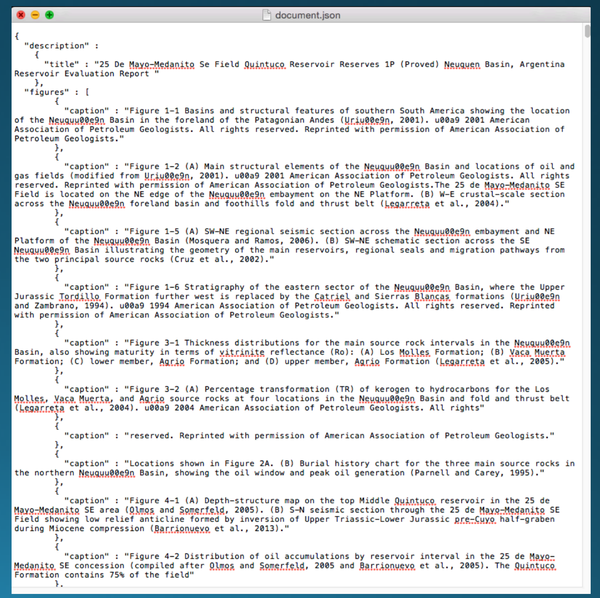
(1) The host rock, or protolith, consists of the rock mass bounding the fault-related structures.
(2) The damage zone is characterized by secondary faults of small displacement, veins, and networks of shear and extensional fractures generally related to the processes of fault growth. Generally, the transition from the host rock to the damage zone is quite gradual.
(3) The fault core is where shear displacement is localized. The core is associated with the development of fault rocks by bulk crushing, particle rotation, abrasion and grain size diminution that obliterate the original host rock fabric (e.g., Billi et al., 2003a,b; Storti et al., 2003; Billi and Storti, 2004; Billi, 2005, 2007).
(4) Following the definition of Vermilye and Scholz (1998), the process zone comprises those features that result directly from propagation of the fault tip. As the damage zone, the process zone is characterized by secondary faults of minor displacement, veins, and networks of shear and extensional fractures. It may overlap with the damage zone as well as with parts of the fault core.

* Corresponding author at: Institut für Erdwissenschaften, Universität Graz, Heinrichstr. 26, A-8010 Graz, Austria. Tel.: +43 316 380 5588; fax: +43 316 380 9870.
E-mail address: walter.kurz@uni-graz.at (W. Kurz).

0191-8141/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsg.2009.04.014







document.json


```
{
  "description" :
    {
      "title" : "25 De Mayo-Medanito Se Field Quintuco Reservoir Reserves 1P (Proved) Neuquen Basin, Argentina Reservoir Evaluation Report "
    },
  "figures" : [
        {
          "caption" : "Figure 1-1 Basins and structural features of southern South America showing the location of the Neuquu00e9n Basin in the foreland of the Patagonian Andes (Uriu00e9n, 2001). u00a9 2001 American Association of Petroleum Geologists. All rights reserved. Reprinted with permission of American Association of Petroleum Geologists."
        },
        {
          "caption" : "Figure 1-2 (A) Main structural elements of the Neuquu00e9n Basin and locations of oil and gas fields (modified from Uriu00e9n, 2001). u00a9 2001 American Association of Petroleum Geologists. All rights reserved. Reprinted with permission of American Association of Petroleum Geologists.The 25 de Mayo-Medanito SE Field is located on the NE edge of the Neuquu00e9n embayment on the NE Platform. (B) W-E crustal-scale section across the Neuquu00e9n foreland basin and foothills fold and thrust belt (Legarreta et al., 2004)."
        },
        {
          "caption" : "Figure 1-5 (A) SW-NE regional seismic section across the Neuquu00e9n embayment and NE Platform of the Neuquu00e9n Basin (Mosquera and Ramos, 2006). (B) SW-NE schematic section across the SE Neuquu00e9n Basin illustrating the geometry of the main reservoirs, regional seals and migration pathways from the two principal source rocks (Cruz et al., 2002)."
        },
        {
          "caption" : "Figure 1-6 Stratigraphy of the eastern sector of the Neuquu00e9n Basin, where the Upper Jurassic Tordillo Formation further west is replaced by the Catriel and Sierras Blancas formations (Uriu00e9n and Zambrano, 1994). u00a9 1994 American Association of Petroleum Geologists. All rights reserved. Reprinted with permission of American Association of Petroleum Geologists."
        },
        {
          "caption" : "Figure 3-1 Thickness distributions for the main source rock intervals in the Neuquu00e9n Basin, also showing maturity in terms of vitrinite reflectance (Ro): (A) Los Molles Formation; (B) Vaca Muerta Formation; (C) lower member, Agrio Formation; and (D) upper member, Agrio Formation (Legarreta et al., 2005)."
        },
        {
          "caption" : "Figure 3-2 (A) Percentage transformation (TR) of kerogen to hydrocarbons for the Los Molles, Vaca Muerta, and Agrio source rocks at four locations in the Neuquu00e9n Basin and fold and thrust belt (Legarreta et al., 2004). u00a9 2004 American Association of Petroleum Geologists. All rights"
        },
          "caption" : "reserved. Reprinted with permission of American Association of Petroleum Geologists."
        },
        {
          "caption" : "Locations shown in Figure 2A. (B) Burial history chart for the three main source rocks in the northern Neuquu00e9n Basin, showing the oil window and peak oil generation (Parnell and Carey, 1995)."
        },
        {
          "caption" : "Figure 4-1 (A) Depth-structure map on the top Middle Quintuco reservoir in the 25 de Mayo-Medanito SE area (Olmos and Somerfeld, 2005). (B) S-N seismic section through the 25 de Mayo-Medanito SE Field showing low relief anticline formed by inversion of Upper Triassic–Lower Jurassic pre-Cuyo half-graben during Miocene compression (Barrionuevo et al., 2013)."
        },
        {
          "caption" : "Figure 4-2 Distribution of oil accumulations by reservoir interval in the 25 de Mayo-Medanito SE concession (compiled after Olmos and Somerfeld, 2005 and Barrionuevo et al., 2005). The Quintuco Formation contains 75% of the field"
        },
```

Putting it all together. Greatly expedite discovery of new materials: From years to weeks

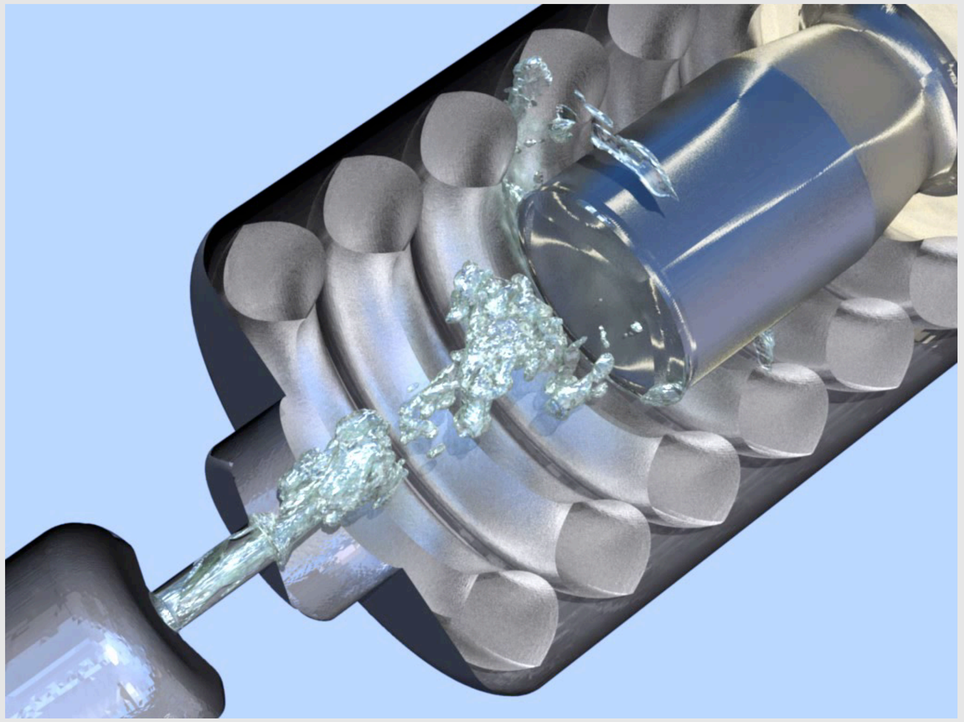
Ingestion
- Tens of thousands of patents ingested in days

Deep Search
- Search for properties, compositions, processes

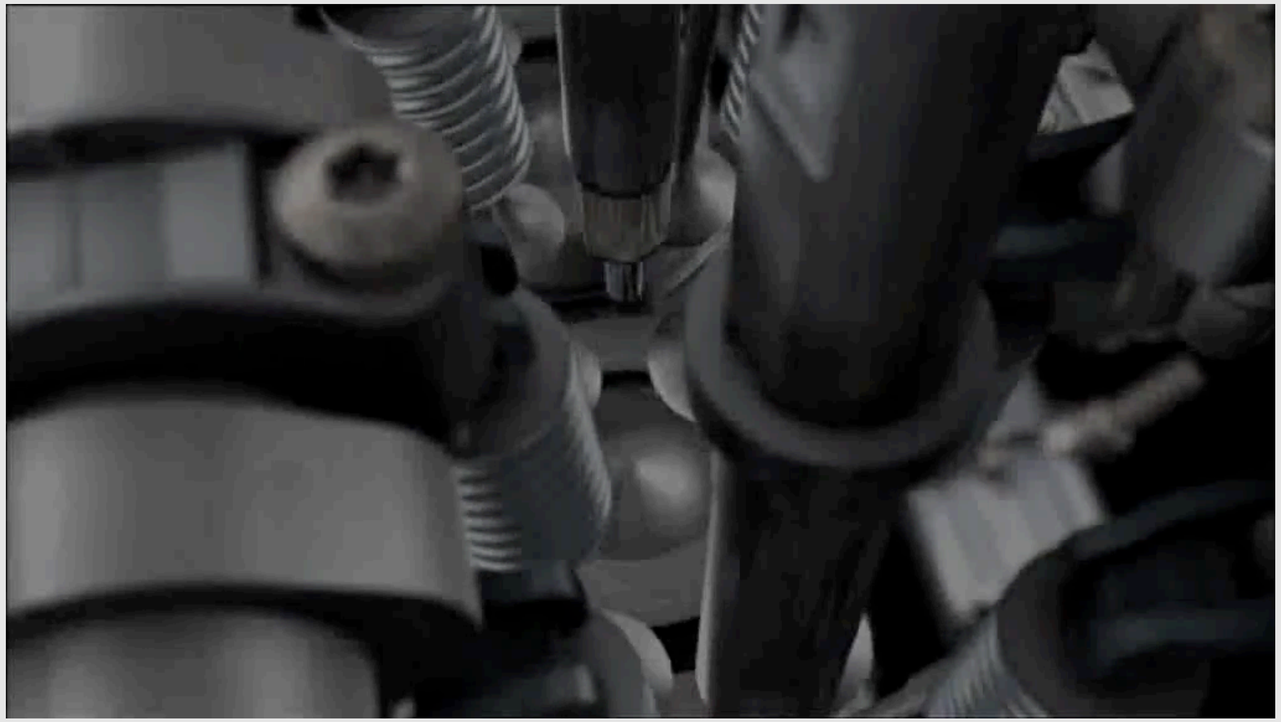
Inference
- Detect salient trends in the data

Simulation
- Fill in knowledge gaps with simulation

An Engineering Case Study: Enriching knowledge space with simulation



Credits: Technical University of Munich



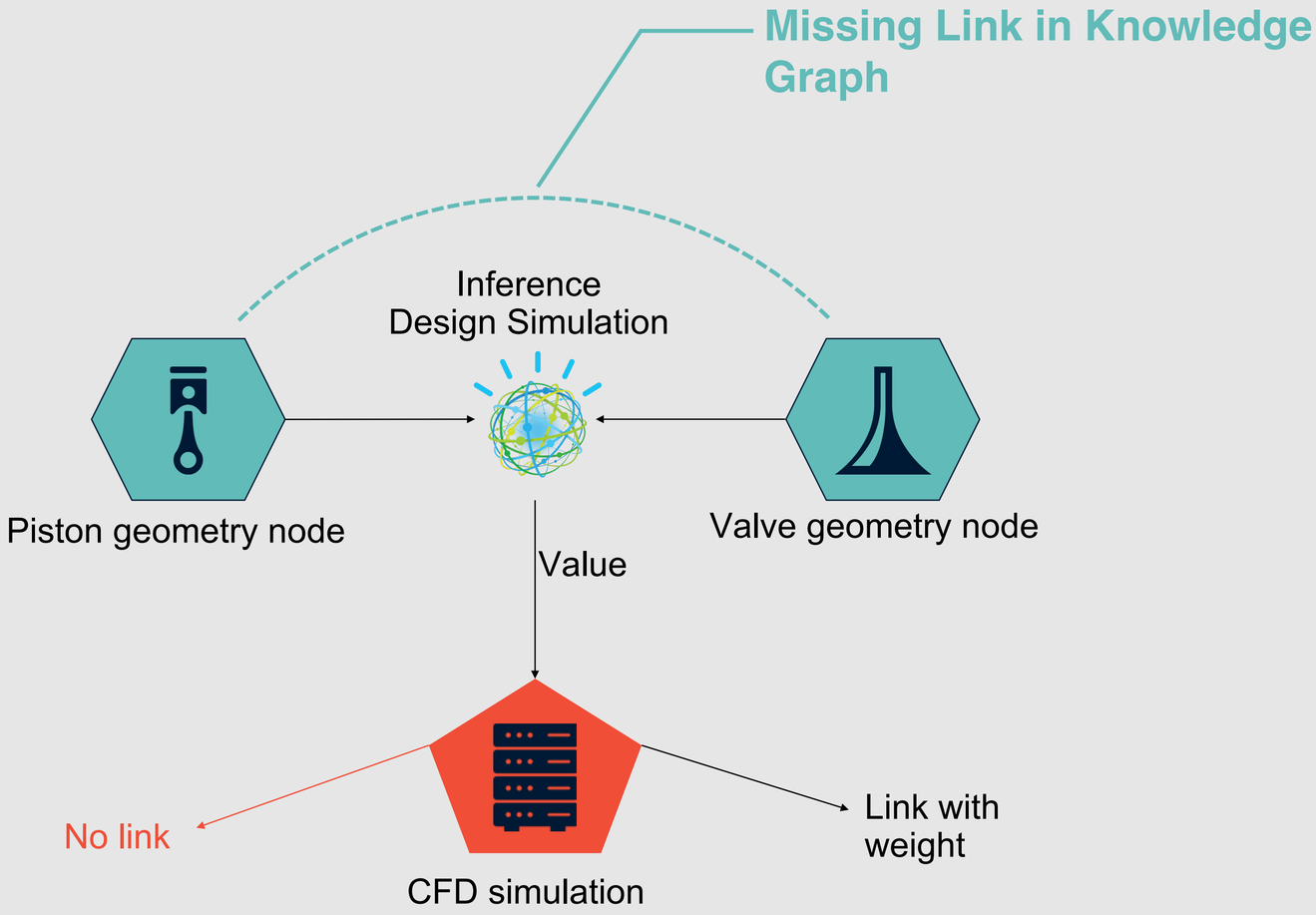
Credits: Mercedez-Benz

An Engineering Case Study: Enriching knowledge space with simulation

Knowledge Graph
- Academic literature & information on internal combustion engines
- Links lab data with literature: fuel combustion + piston geometry + …

New cylinder/piston/injector geometries
- Use Knowledge Graph to quickly rule out non-viable design directions
- Augment missing information & perform validation with advanced CFD

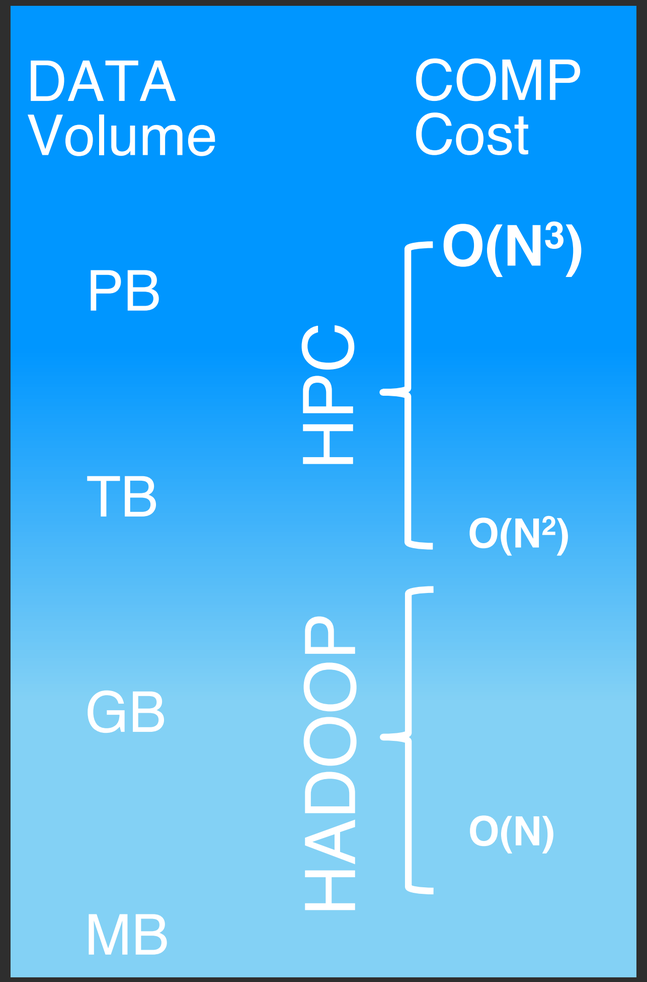
Where to augment: Knowledge analytics
- Adding link if significantly changing the quality of the knowledge graph
- Only specific and well defined simulations need to be done

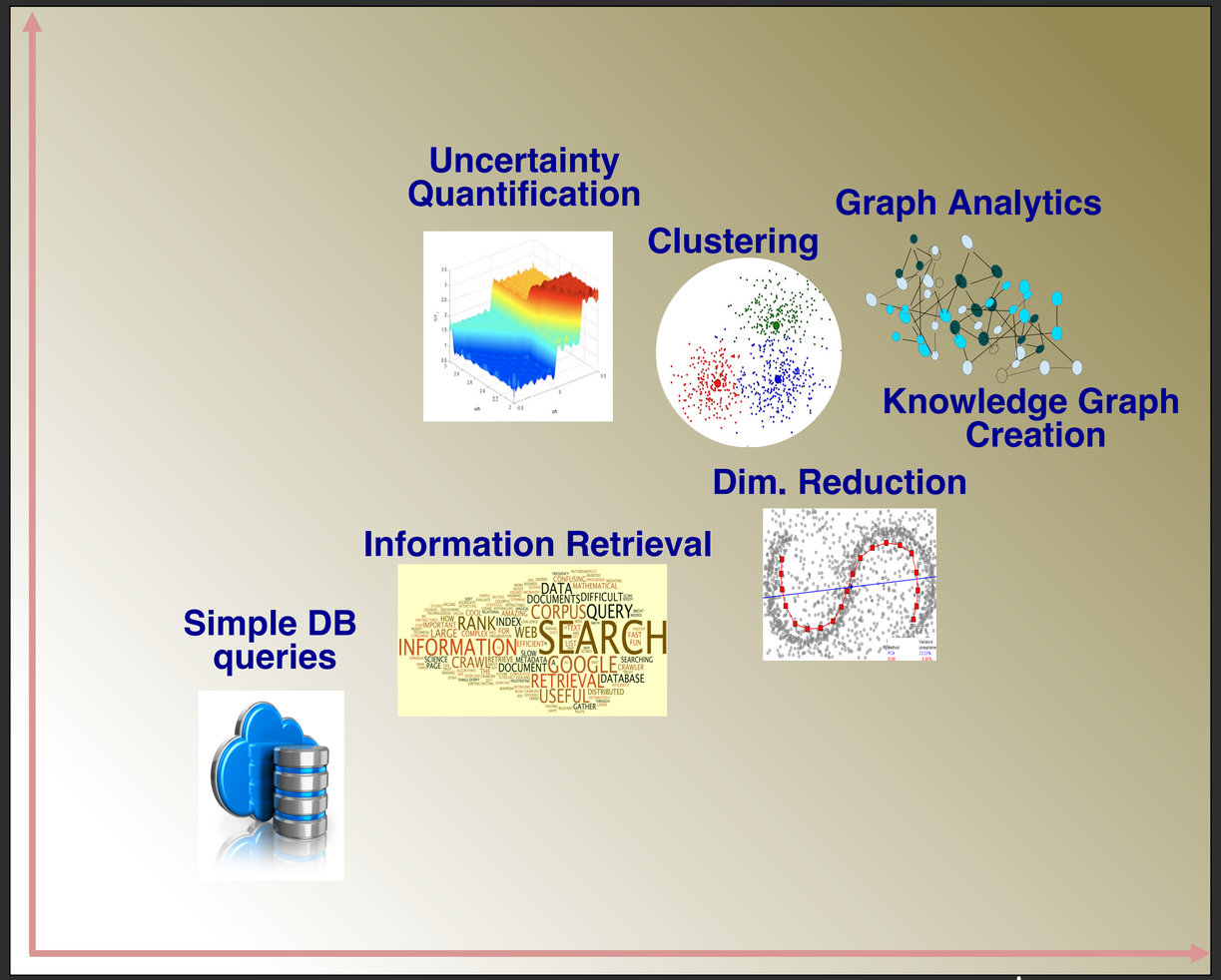


WHAT WAS TAKING YEARS CAN NOW BE DELIVERED IN DAYS

Beating complexity:

We need both new collaboration & new computing paradigms





TOWARDS COGNITIVE COMPUTING

OpenPOWER

OpenPOWER is an open development community using the POWER Architecture

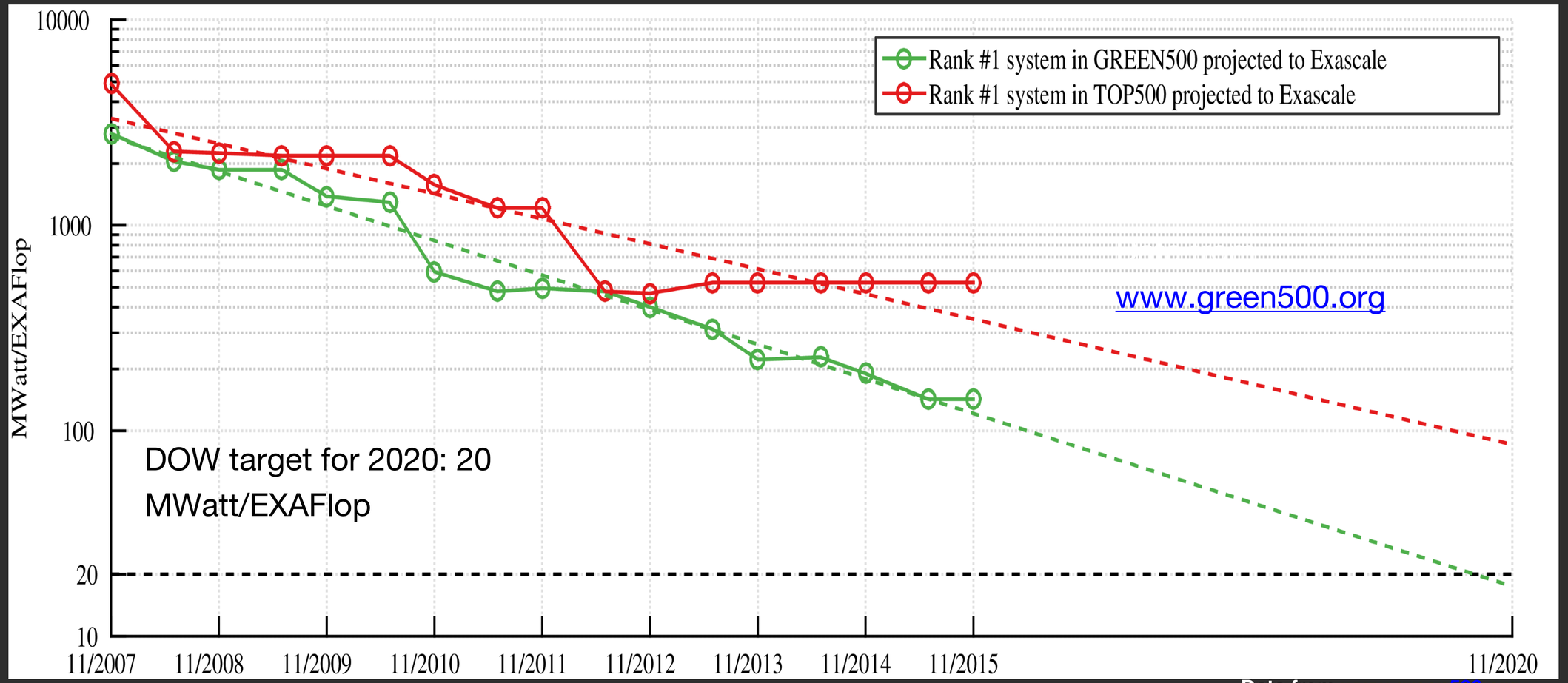
Vibrant ecosystem through open development

Accelerated innovation through collaboration of partners

Driving industry performance leadership

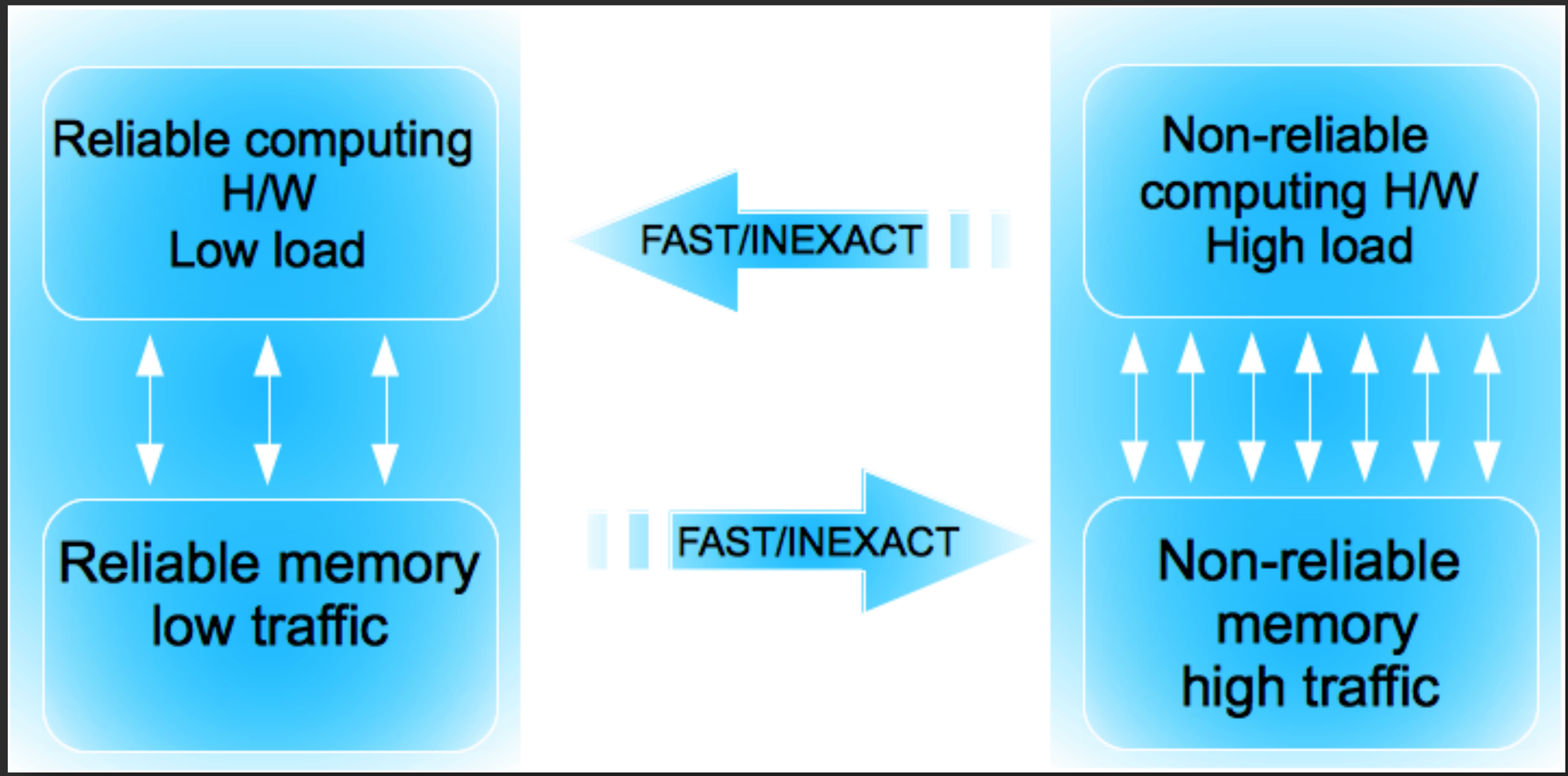
OpenPOWER

- We start to see an exponential behavior in the Green500. But is this really affecting the top line?
  5 years ago: 2.1 GF/W, now 1.9 GF/W
- OpenPOWER brings more than 3x improvement in M/W performance wrt top line systems



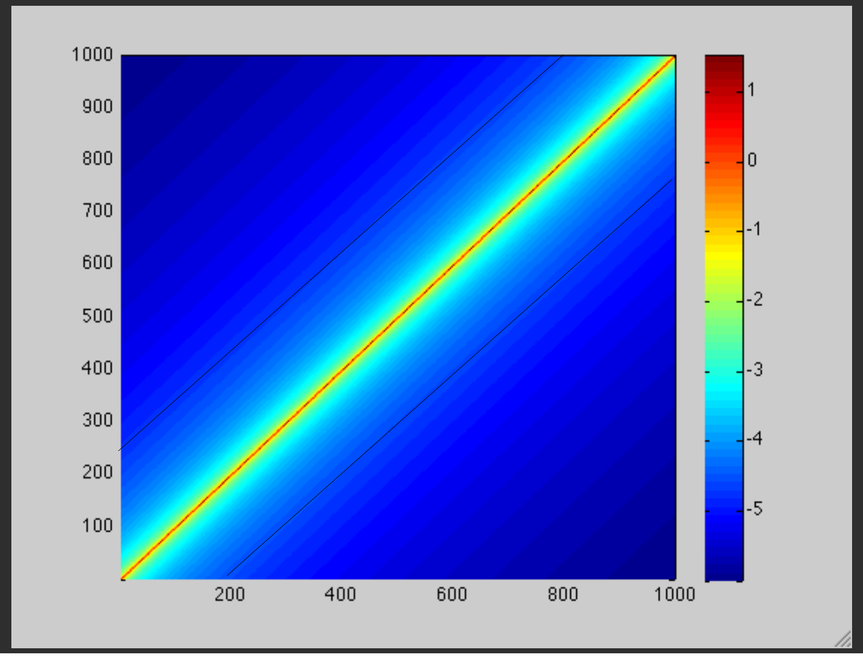
Data from www.green500.org

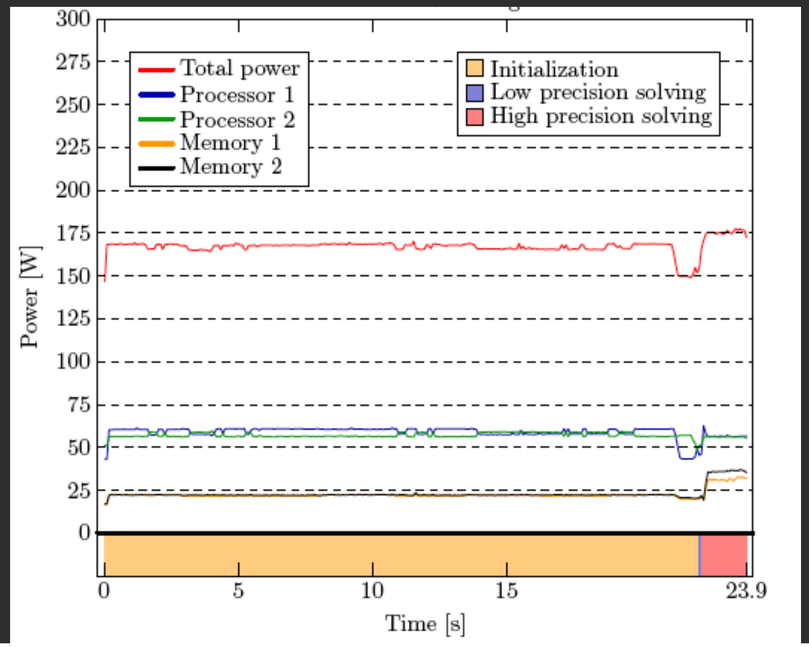
New Horizon: OpenPOWER + INEXACT ACCELERATION:
Transprecision Computing Architecture



Actual implementations on OpenPOWER

Inference: 103KJ to 0.3KJ: 300x improvement





Method	Time	Average power	Energy	GFlops	GFlops/W
banded CG 1 RHS	1.8 s	174.1 W, s.e. 4.9 W	0.3 kW·s	5.5	0.03
banded CG 32 RHS's	8.4 s	172.6 W, s.e. 14.2 W	1.5 kW·s	37.8	0.22
CG 1 RHS	53.8 s	179.0 W, s.e. 1.8 W	9.6 kW·s	15.7	0.09
CG 32 RHS's	125.5 s	195.0 W, s.e. 10.8 W	24.6 kW·s	222.2	1.13
Cholesky	546.0 s	190.0 W, s.e. 13.5 W	103.7 kW·s	214.4	1.11

Research Roadmap: Accelerating Big Data workloads

OpenPOWER + INEXACT ACCELERATION
Transprecision Computing Architecture

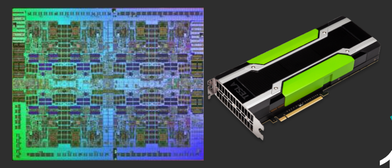
SPEEDUP / ENERGY GAINS

NEW TECH.

LIBRARIES & APPLIANCE

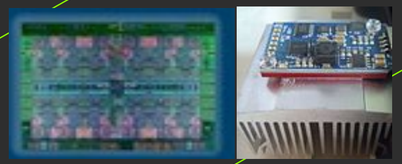
LIBRARIES

100x



POWER8 + K80
TRANSPRECISION COMPUTING
SINGLE-DOUBLE PRECISION
~5 secs

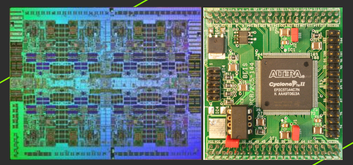
1000x / 2000x



POWER8'-9 + ASIC
FIXED POINT/8b/16b/32b PREC.
REDUCED DATA ACCURACY
STOCHASTIC DATA TRANSFER
Projection: 0.5-1 sec / 0.05 KJ

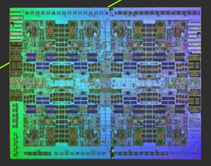
CUSTOM
STOCHASTIC ARITH. & DATA MOVEMENT & STORAGE
Projection: 0.1 sec / 0.001 KJ

50x / 500x

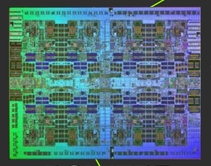


POWER7 + FPGA
FIXED POINT/8b/16b/32b PREC.
REDUCED DATA ACCURACY
SENSOR POWER MEASUR.
10 secs / 0.2 KJ
FPGA: 19 W peak / 15 W idle

10x



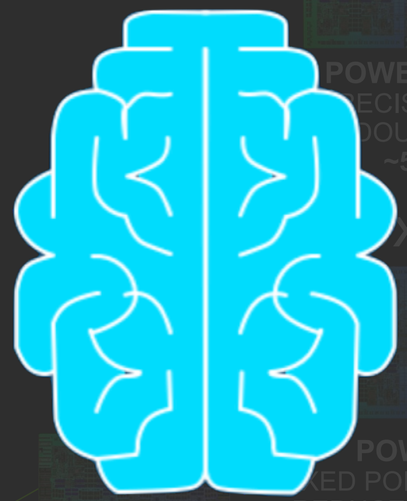
POWER7
TRANSPRECISION COMPUTING
SENSOR POWER MEASUR.
54 secs / 9.6 KJ
170 W peak 140 W idle



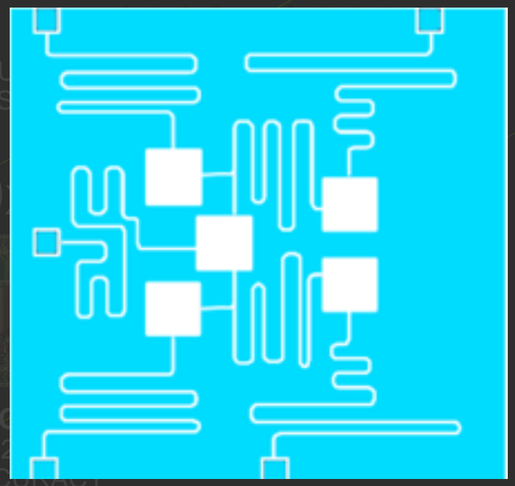
POWER7
BASELINE BLAS3
SINGLE/DOUBLE PREC.
SENSOR POWER MEASUR.
546 secs / 104 KJ
210 W peak / 140 W idle

2011 2012 2015 2016/2018 2020+

Research Roadmap: Accelerating Big Data workloads

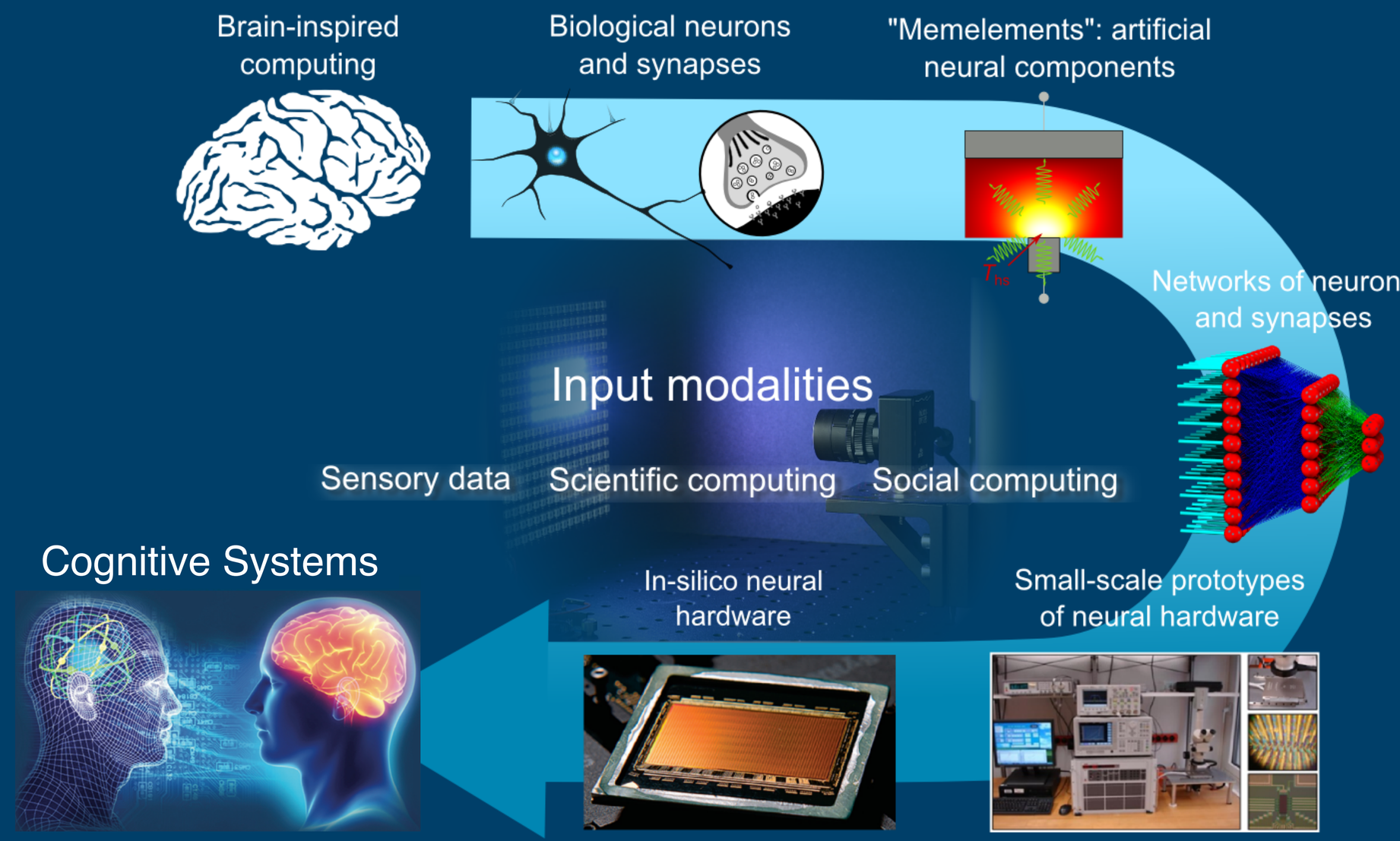


Brain Inspired Computing



Quantum Computing

Neuromorphic Computing



TrueNorth Chip (SyNAPSE)

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

Saliency

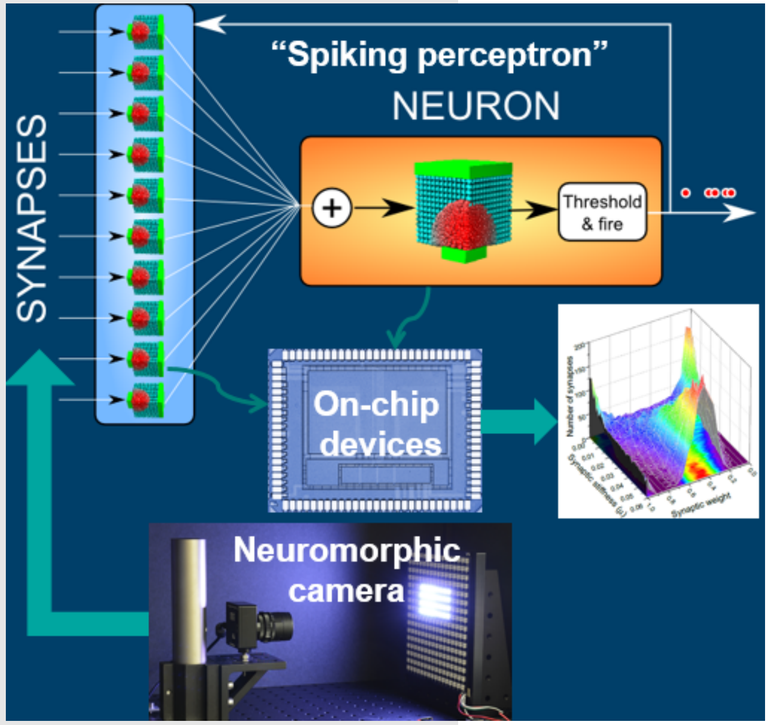
Saliency + Classification

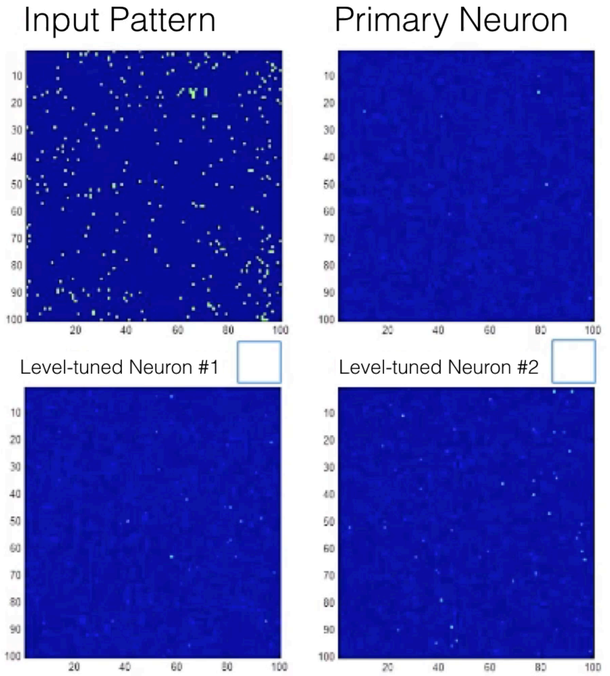
Object Centers

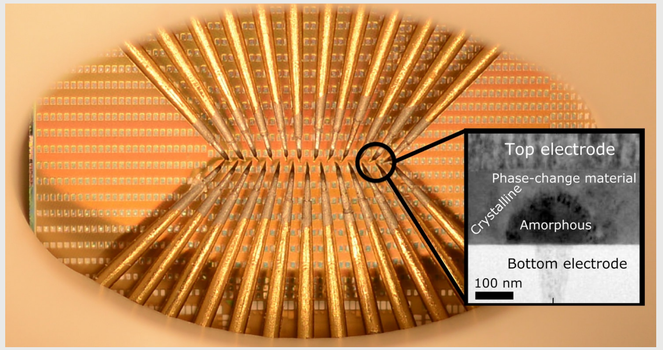
Output

Detecting Correlations with a Spiking Neural Network

Searching Big Data, Pattern Recognition, IoT Data



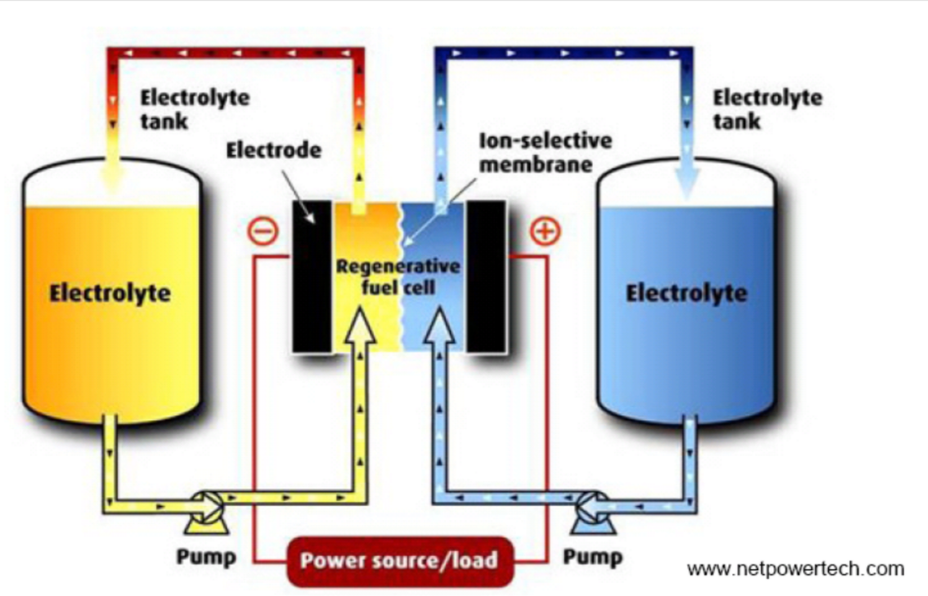


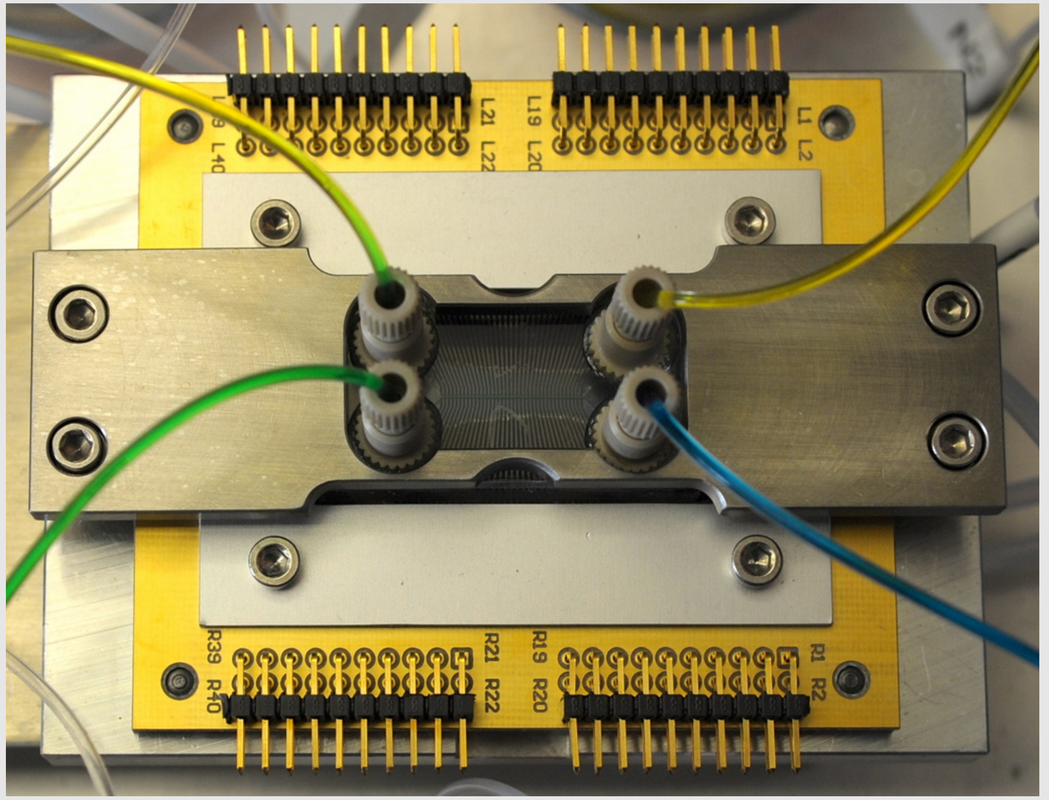


Use **only unsupervised learning** and consume very low power

Brain Inspired Computing: Electronic Blood

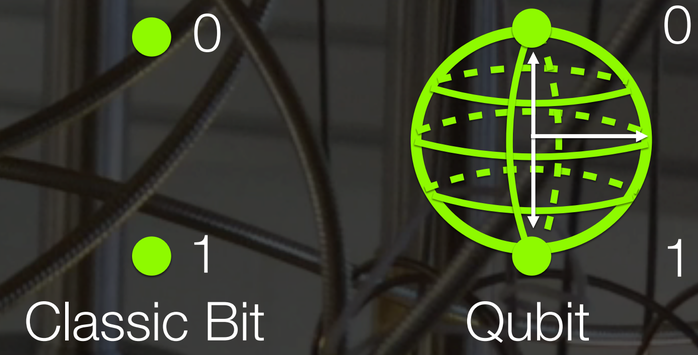
- 98% of the energy of a computer is for cooling
- Liquid removes heat 4000x more efficiently than air
- The brain is powered and cooled using liquid, can we do the same for computers?
- The result: a 1 PetaFlop supercomputer in 10 liters





Building a Universal Quantum Computer

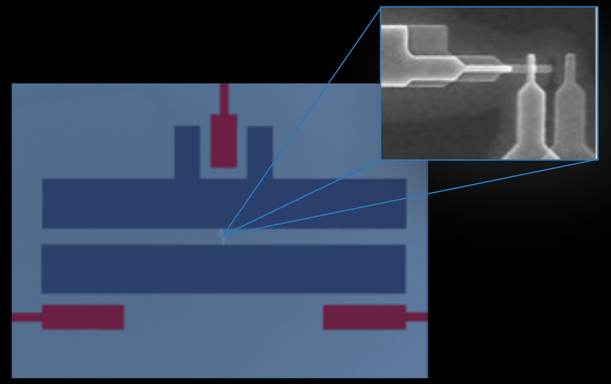
- IBM is building the first Universal Quantum Computers for business and science - IBM Q systems with ~50 qubits in the next few years
- IBM Quantum Experience allows the public to program qubits in the cloud
- The release of an upgraded simulator on the IBM Quantum Experience can model circuits with up to 20 qubits.
- Applications: Drug and Materials Discovery, Supply Chain, Financial Services, Artificial Intelligences and Cloud Security

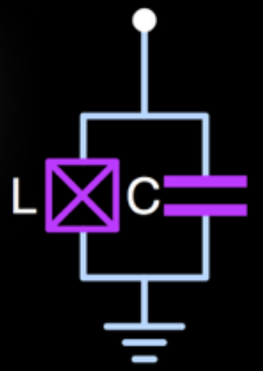


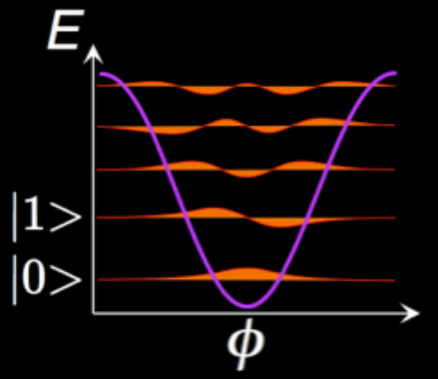
Superconducting Qubit Processor – A Closer Look

Superconducting Qubit:

- non-linear Josephson Junction (Inductance)
- anharmonic energy spectrum => qubit
- nearly dissipationless => T_1, T_2 ~ 70 μs



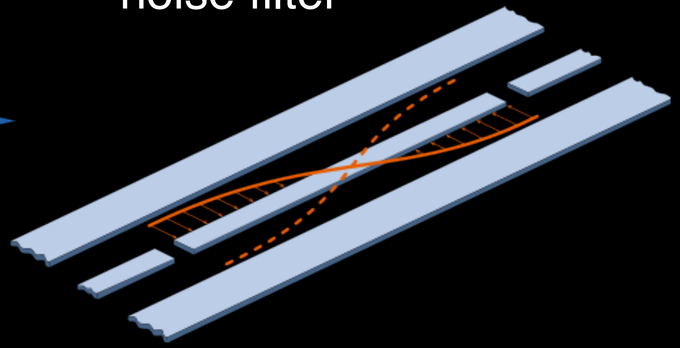


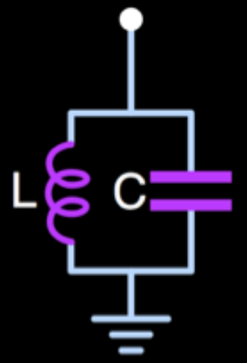


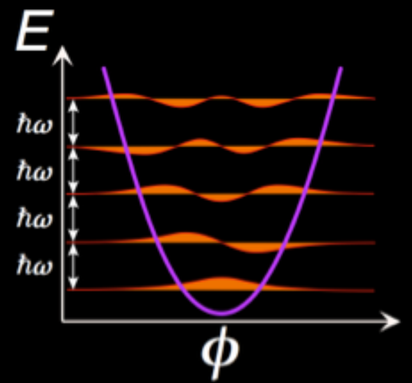
$E_{01} \approx 5\ GHz \approx 240\ mK$

Microwave Resonator as:

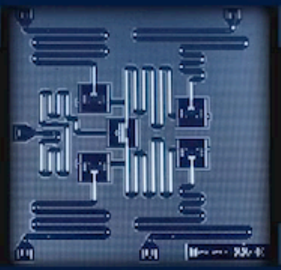
- read-out of qubit states
- multi-qubit quantum bus
- noise filter







Directions



Qubit 0 properties

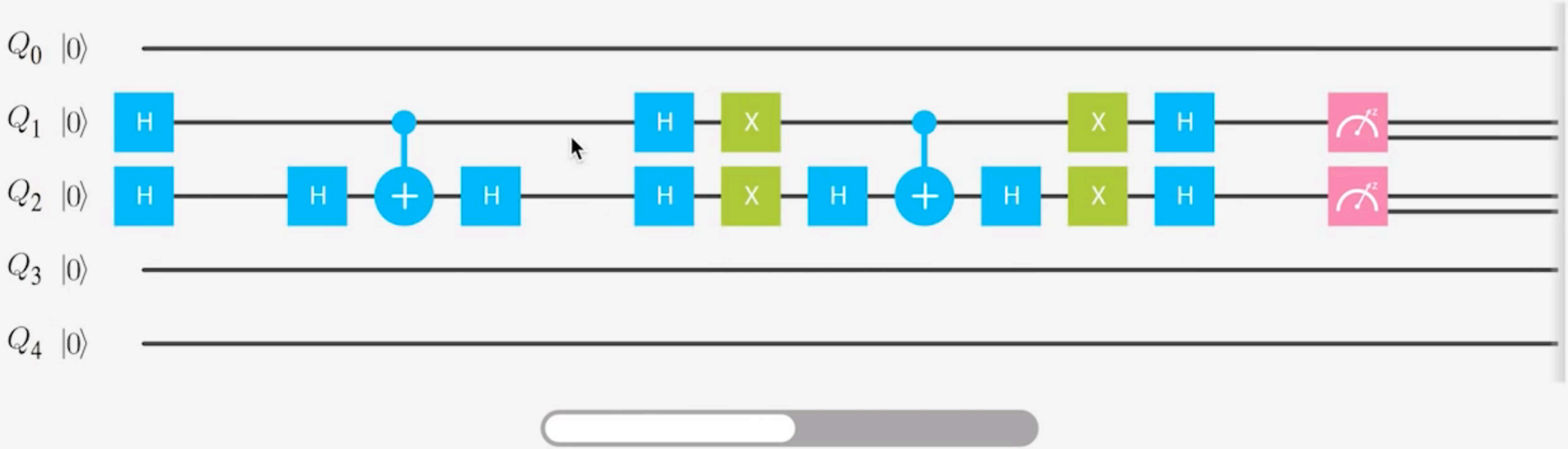
f	5.35 GHz
T_1	54 μs
T_2	74.3 μs
ϵ_g	2.6×10^{-3}

2016-04-27 02:47

User Guide | **Composer** | My Scores

Name: 'Grover's Search Algorithm, 11'

Real Quantum Processor



Simulate

Run

New

Save

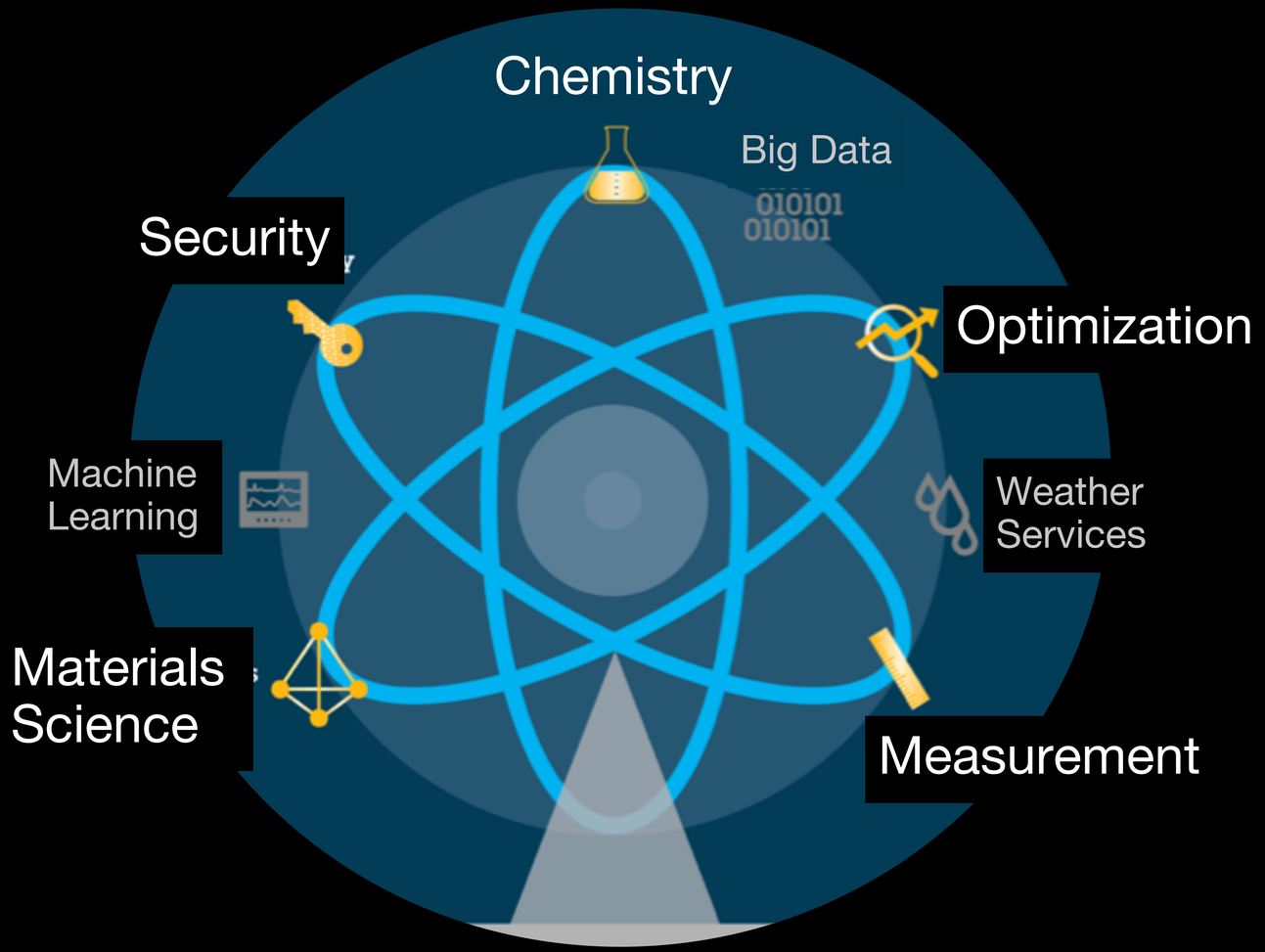
Save as

Results

Help

GATES Id X Z Y H S $S^\dagger$ + T $T^\dagger$ MEASURE

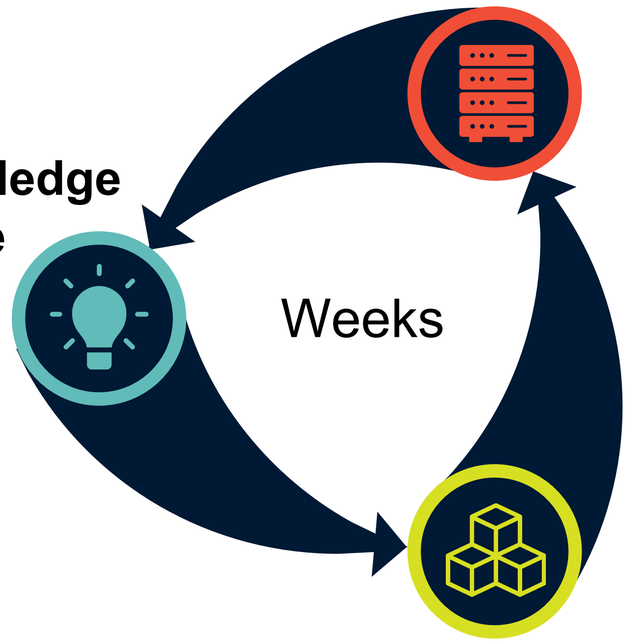
Industry Applications in Quantum Computing



- **Chemistry,** e.g. for catalyst design
- **Material Science**, e.g. for energy efficient devices
- **Life Sciences,** e.g. for drug development
- **Optimization**, e.g. for cognitive computing and business processes
- **Cryptography,** e.g. for secure communication and information processing
- **Education**, e.g. to train engineers for the future quantum industry

Cognitive Discovery: A new way to do R&D

Ingest data and create massive knowledge spaces



- Supercomputing
- Quantum and new computing paradigms
- Inference (ML)

Link evidence with knowledge spaces. Drive deep search

THANK YOU

Q&A