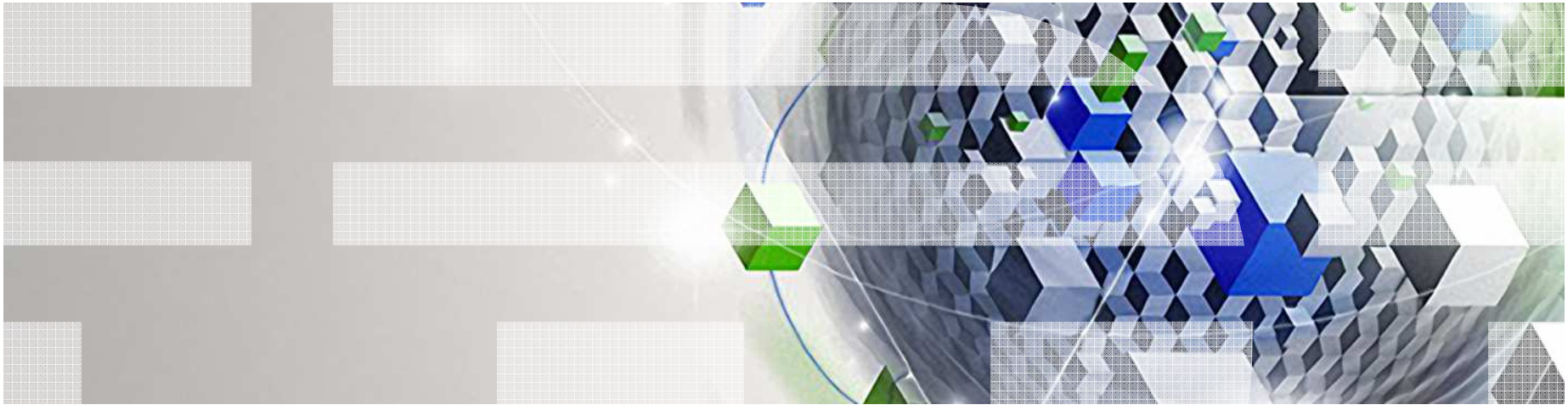
Dr. Axel Koester – Storage Chief Technologist, European Storage Competence Center

IBM Watson technology at work

and the impact on BIG DATA

(or how to compress the English internet in 1TB)



axel.koester@de.ibm.com

Watson Grand Challenge

Current Watson Projects

How to deal with Massive Data

Lab Futures

axel.koester@de.ibm.com

3.14159265358979323846264338327950288419716939937510582097494459230781640628620899862803482534211706798214808651328230664709384460955058223172535940812848111745028410270193852110555964462294895493038196442881097566593344612847564823381286786783165271201909145648566923460348610454326648213393607260249141273724587006606315588174881520920962829254091715364367892590360011330530548820466521384146951941511609433057270365759591953092186117381326879486060651209008240243

Records

Which one is more unbelievable?

Fabrice Bellard's home computer calculated pi to 2.7 trillion digits
(previous record held by a supercomputer)

Lu Chao successfully recited 67.890 digits of pi from memory in 24 hours and 4 minutes Guinness book of records

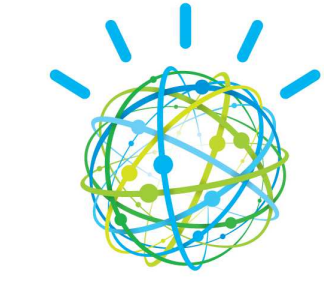
axel.koester@de.ibm.com

Reciting π = Peanuts for a computer.

But how about

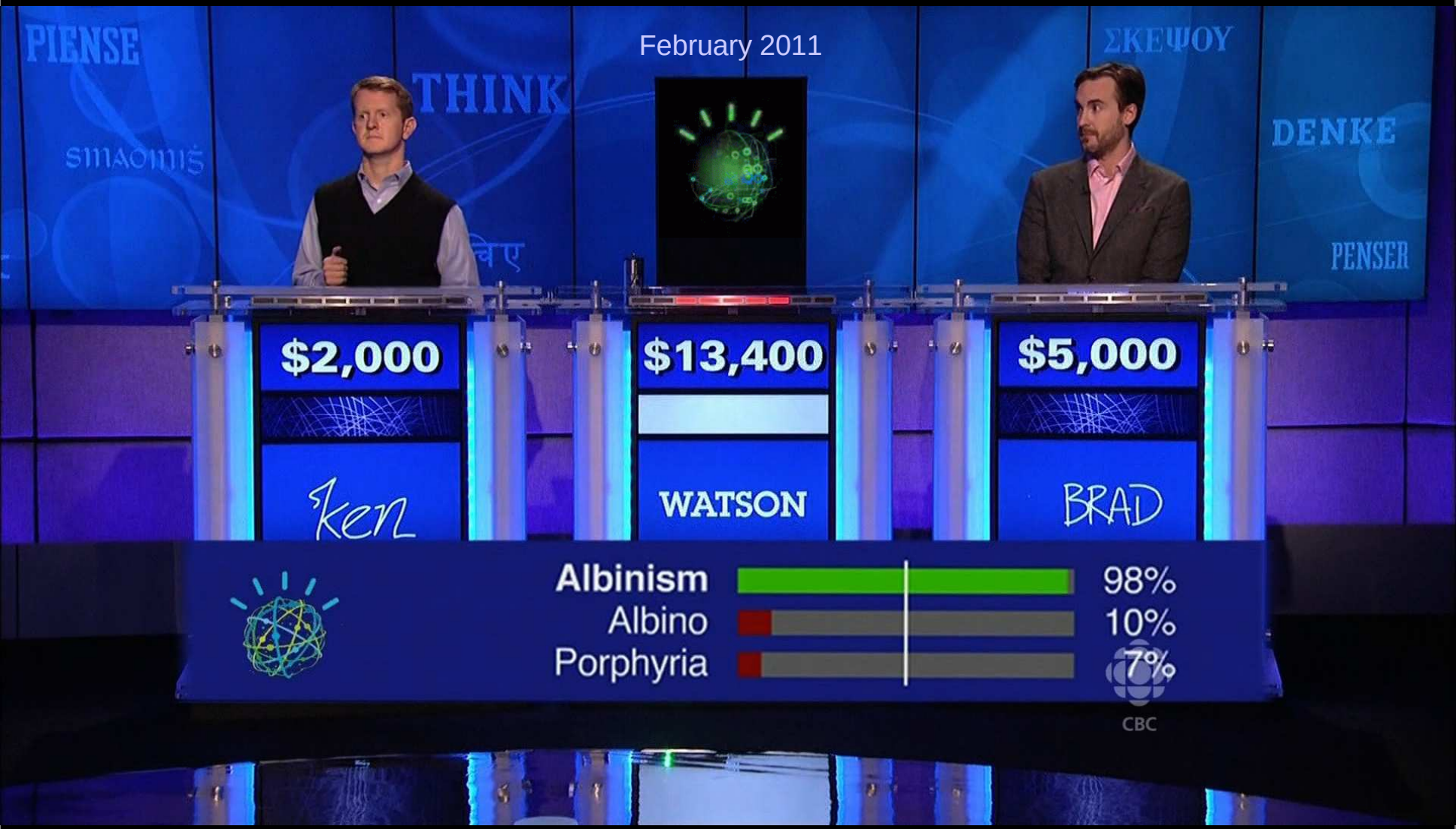
- **Reading & understanding** unstructured documents
- ***Understanding ambiguous human questions and pun*** ?
- *Responding in real time, before being 100% sure*
- Taking a 'good guess'

➔ IBM Grand Challenge



 axel.koester@de.ibm.com

Jeopardy! all-time champions Ken Jennings & Brad Rutter



IBM WATSON accesses its knowledge* in parallel

(*) Multiple millions of complex matching searches per second, within a knowledge base of the English WWW.

10 km book stack

200 million memorized book pages,

with Wikipedia alone totaling 2,25 Mio.

takes ~123 years to read

axel.koester@de.ibm.com

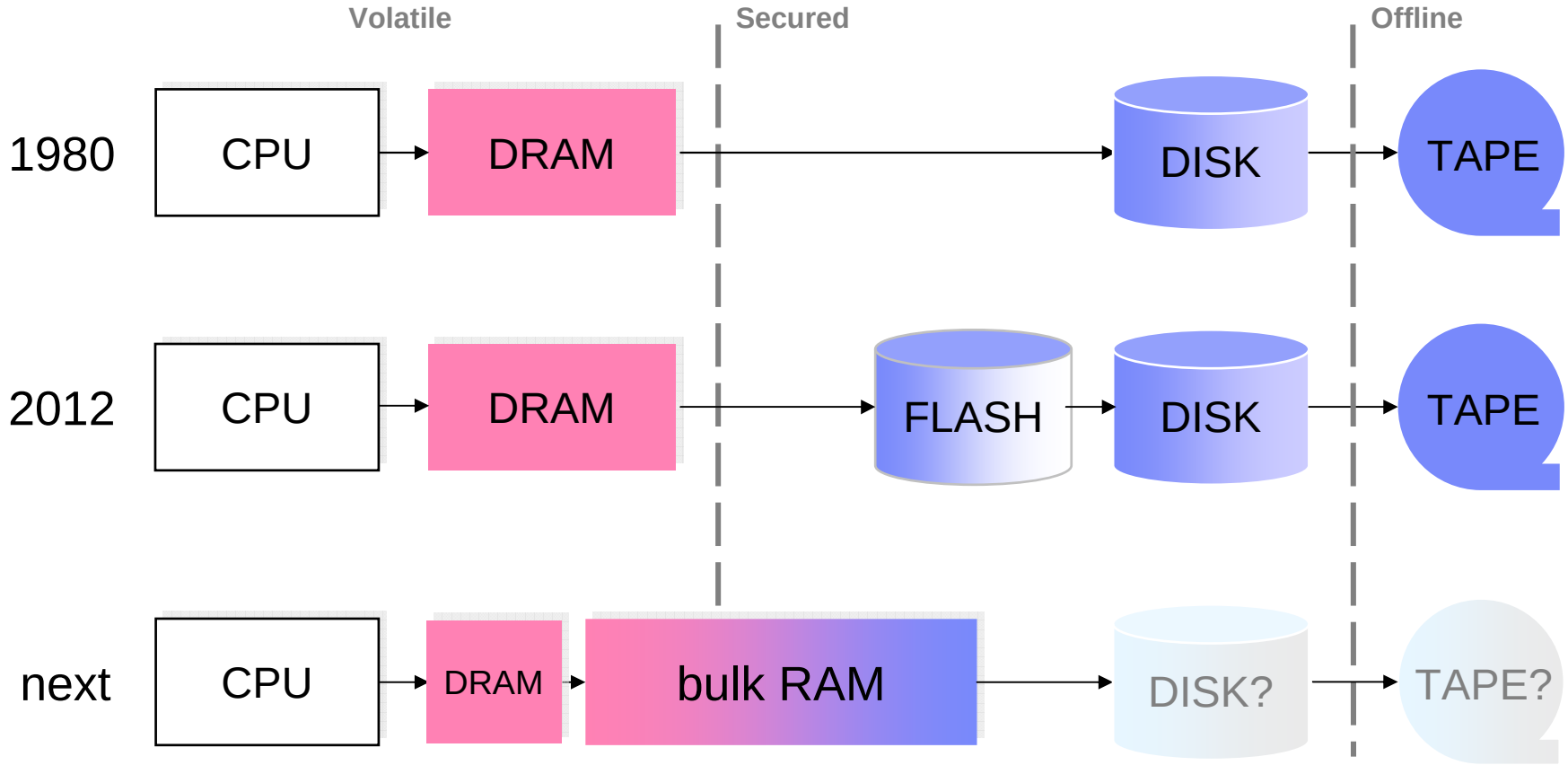
IBM WATSON is an "information aggregator"

90 × 32 core IBM Power®750 / **16 TB RAM, 1TB disk content**

***Parallel access** **to 100% of the data, no Internet access during games**

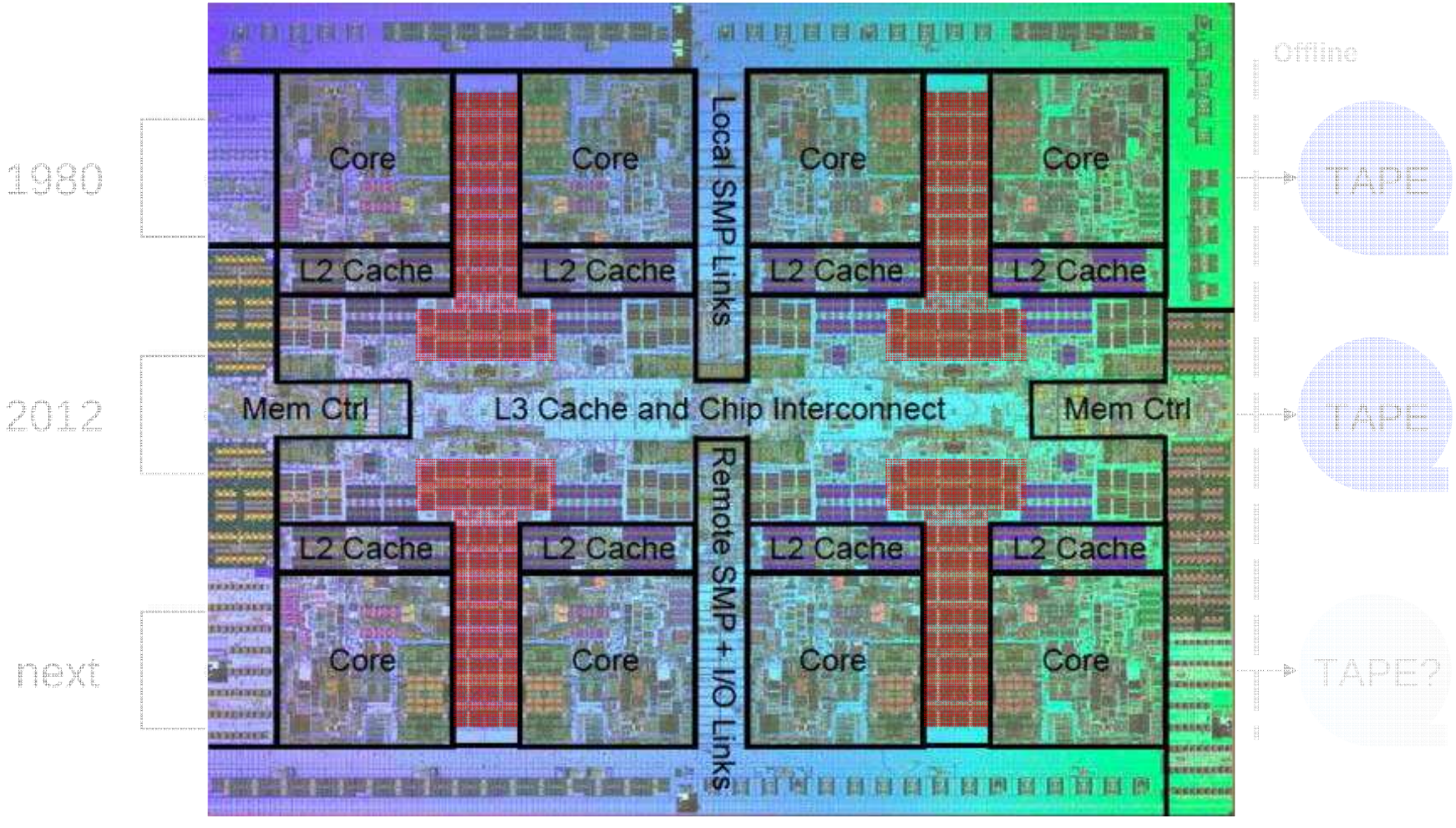
out of 4 TB GPFS disk space, with 16TB = 15 TiB RAM

How to achieve full parallel access to data



axel.koester@de.ibm.com

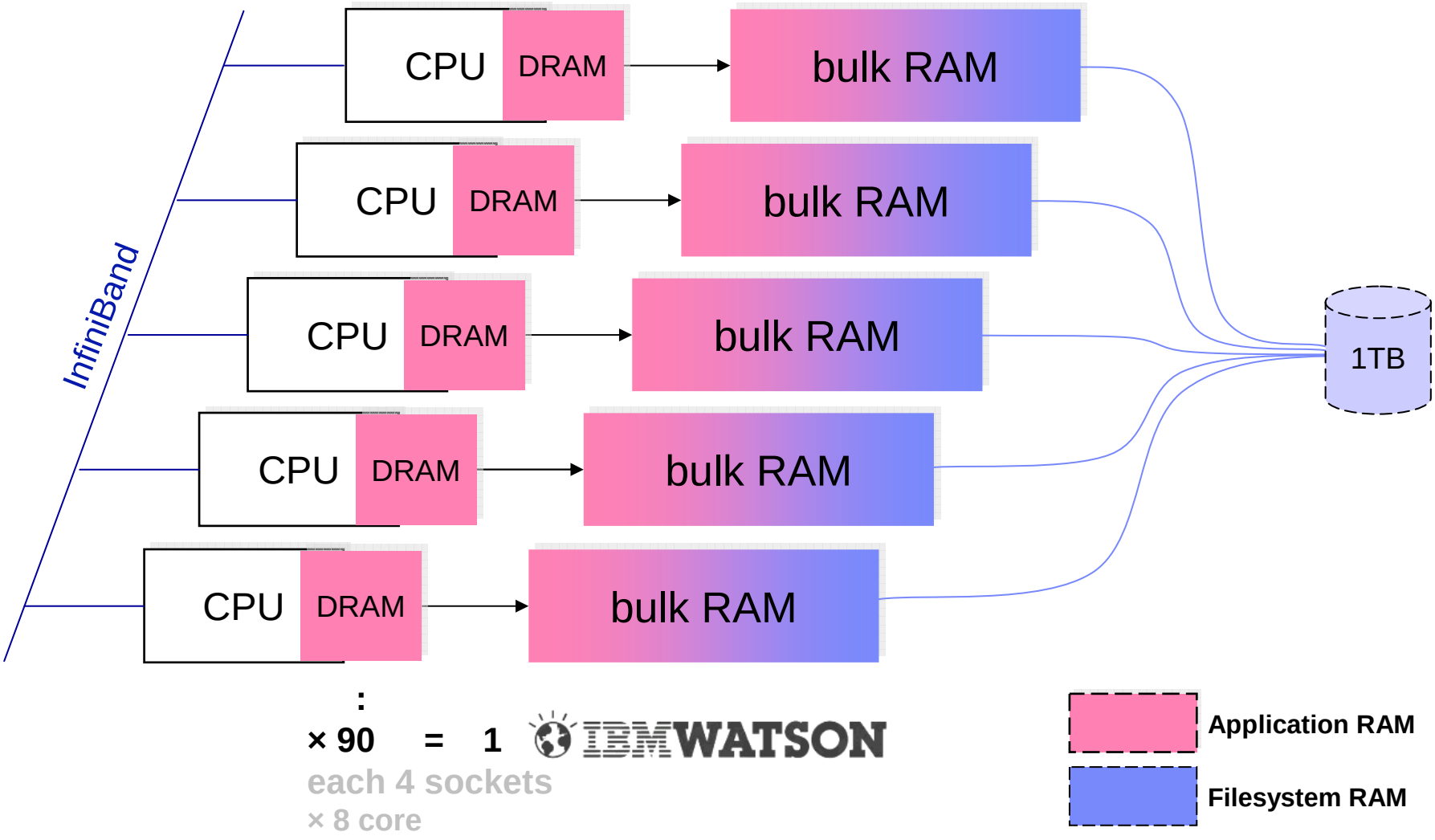
POWER7 with integrated DRAM 32MB



Challenge: CPU = Transistors, SRAM = Transistors, **DRAM = Capacitors (denser!)**

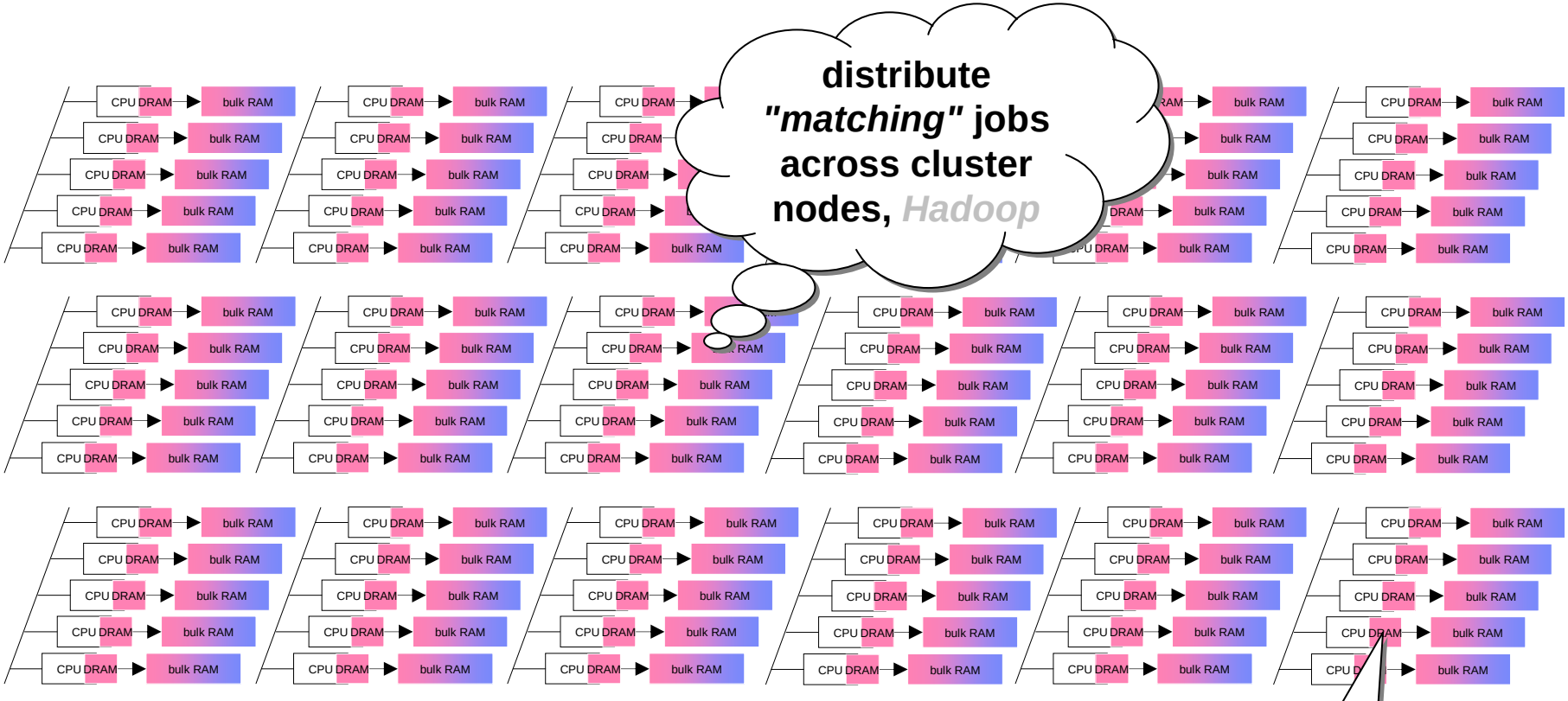
 axel.koester@de.ibm.com

How to achieve full parallel access to data



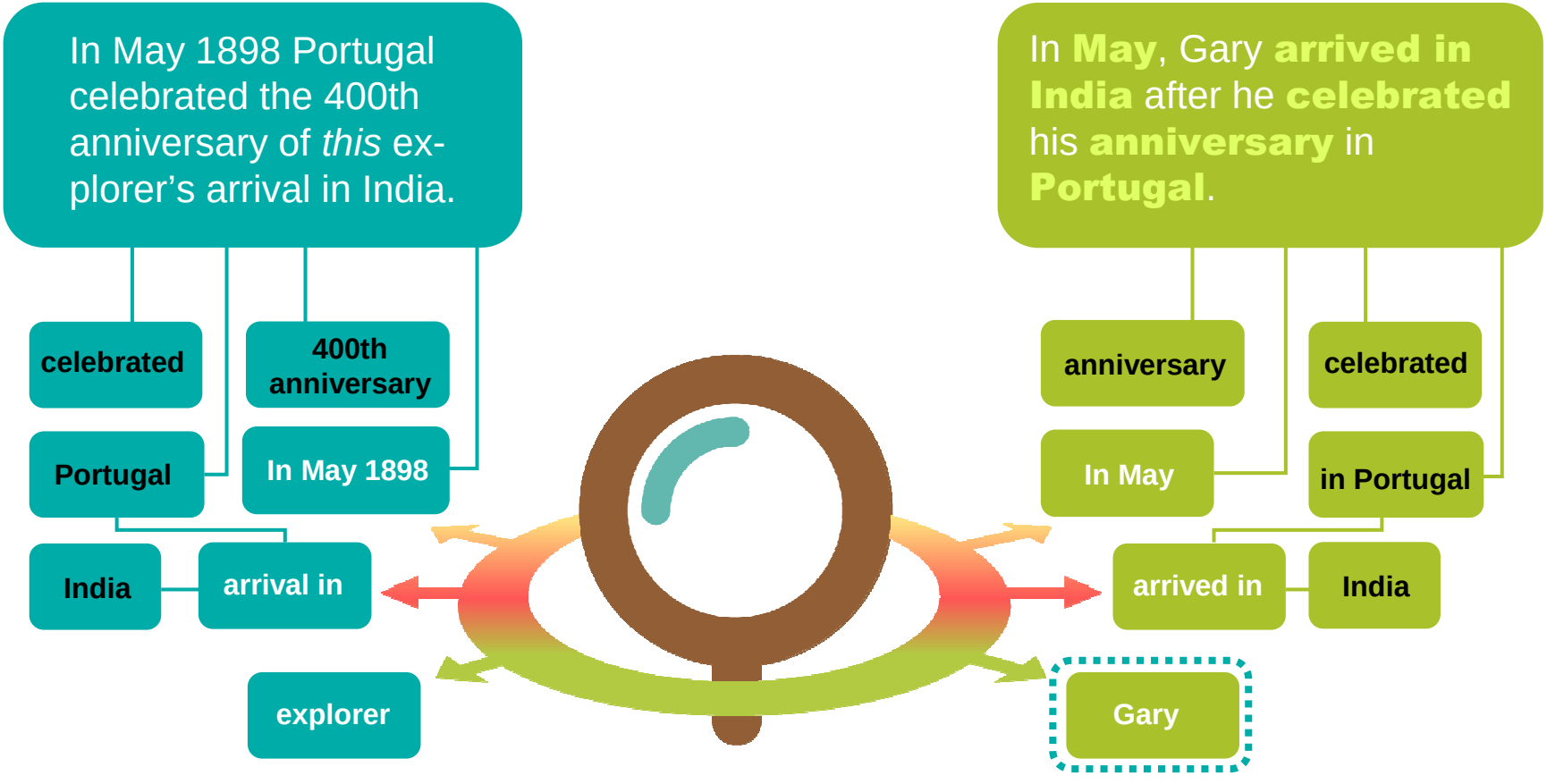
 axel.koester@de.ibm.com

What is IBM Watson doing with these massive resources?



How is Watson different from Google?

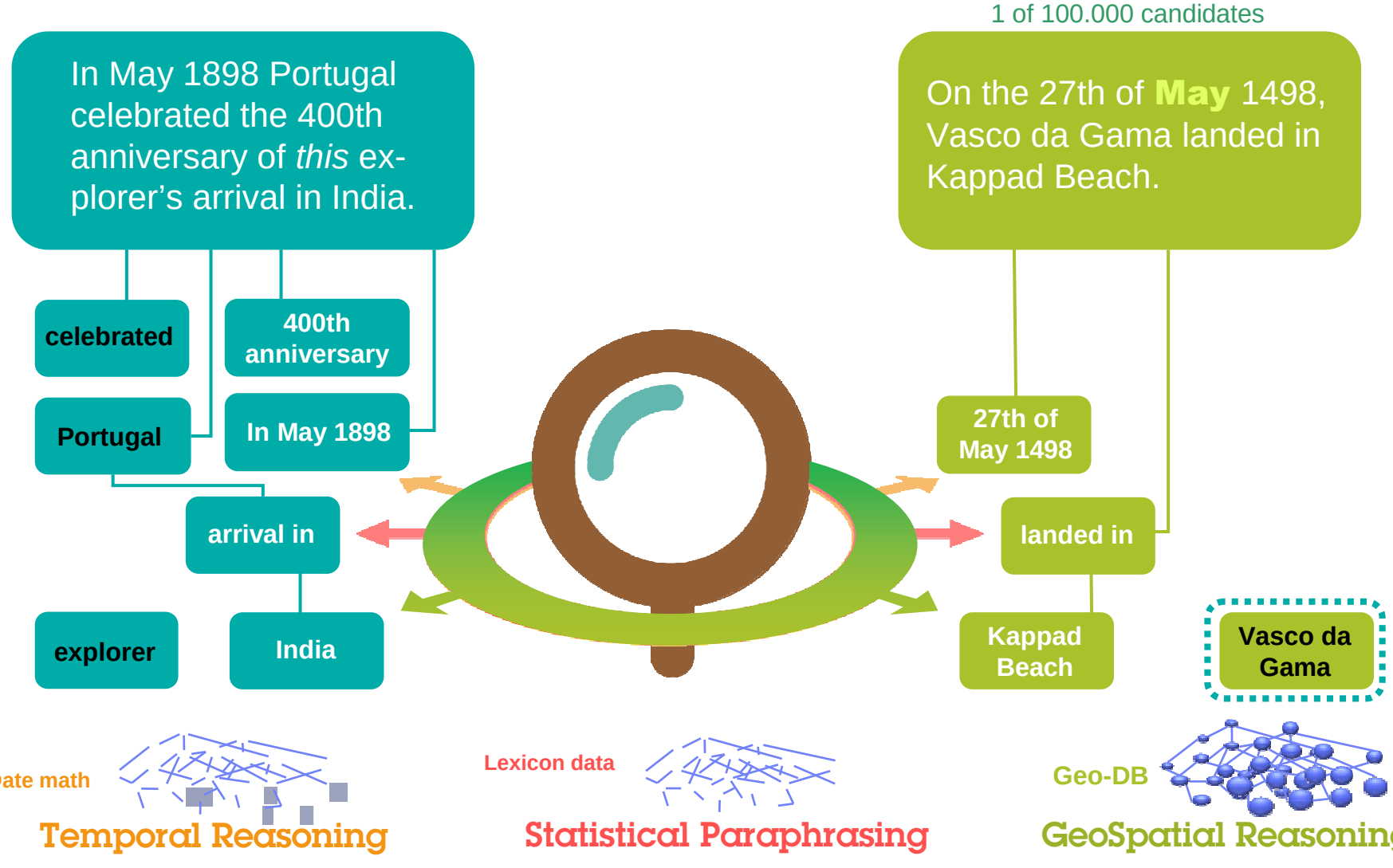
Here is Google:



Keyword Matching
(+page ranking)

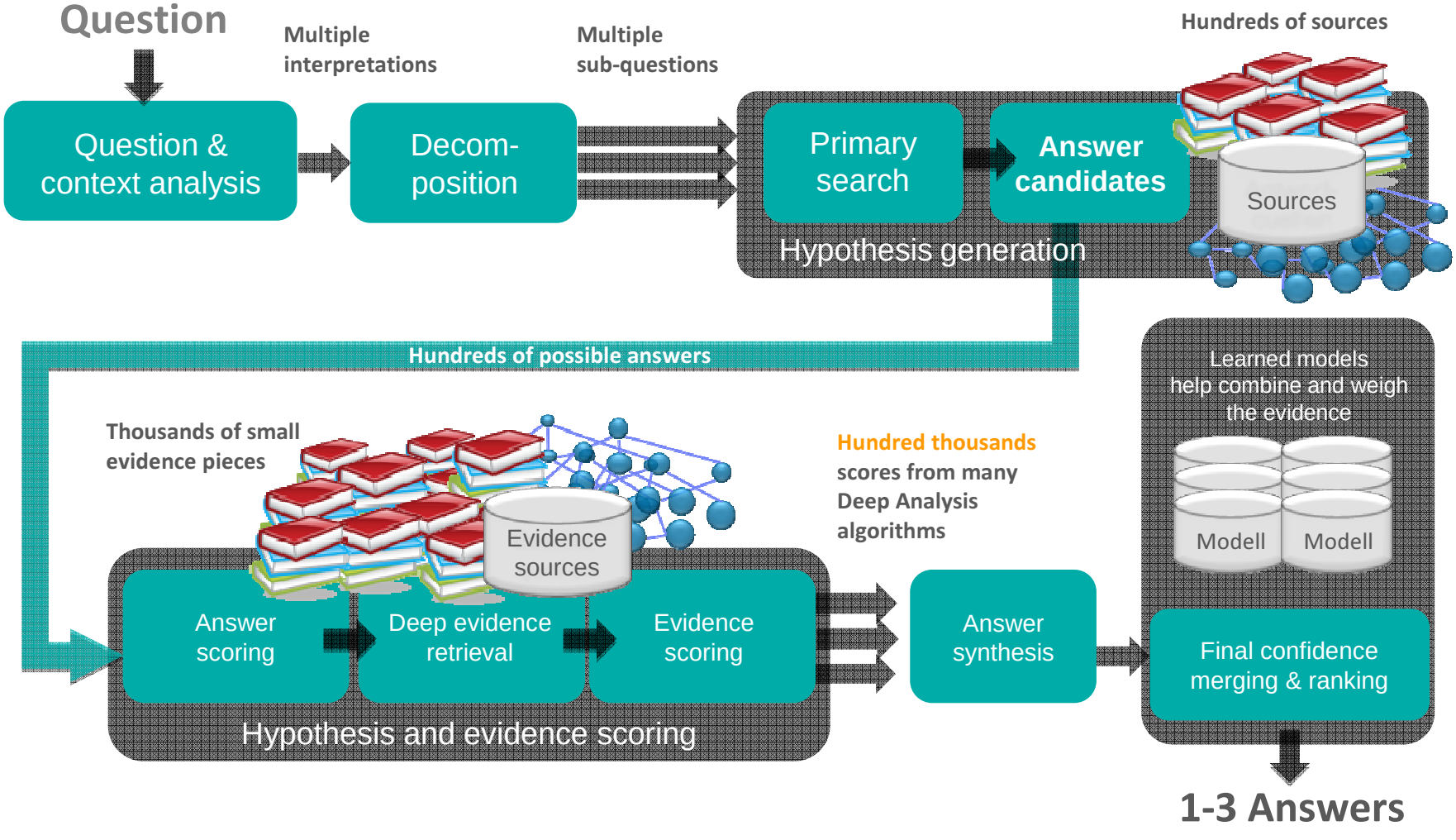
 axel.koester@de.ibm.com

How is Watson different from Google?
Here is **Watson**:



 axel.koester@de.ibm.com

Software Architecture



 axel.koester@de.ibm.com

Watson's inventors surprised by amazing correct answers

Dr. Jennifer Chu-Carrol,
Watson algorithms

axel.koester@de.ibm.com

A precursor of new informatics – much less *deterministic*

IBM WATSON

Watson @ work
Recent Projects

axel.koester@de.ibm.com

In mid 2011, first **Watson** projects were started

WELLPOINT

CEDARS-SINAI Hospital, L.A.

Healthcare

Watson reads all medical publications, reports, journals using Nuance Clinical Language Understanding CLU and respond to nurse's questions with DeepQA®

NUANCE

80% of patient data is unstructured

Call center project

Reduced average call time from 5 to 4 min. by avoiding full traversal of yes/no questions tree: 20% reduced cost

More "Smarter Planet" ideas were born: fraud detection, justice antecedent case search, traffic flow …

axel.koester@de.ibm.com

Since late 2011, Watson projects are multiplying

Banking

a Micro-Watson reads all *minutes of consultation* and establishes trends and marketing directives based on written client requests.
It also finds omissions & inconsistencies.

Insurance

a Micro-Watson reads relevant websites in order to quickly capture client opinion and competitive threats.
Counselors can "ask Watson" for related information.

Airline

a Micro-Watson reads competitors' websites & publications in order to identify potential niches in flight plans / destinations.

Defense

Watson scans non-classified repositories for text with classified content / keywords.

axel.koester@de.ibm.com

German car makers started various **Watson** projects

Maintenance & Repair

"There's a rattling noise in the rear when I…"

Watson reads garage reports and helps identifying defects quickly.

More car makers' **Watson** projects

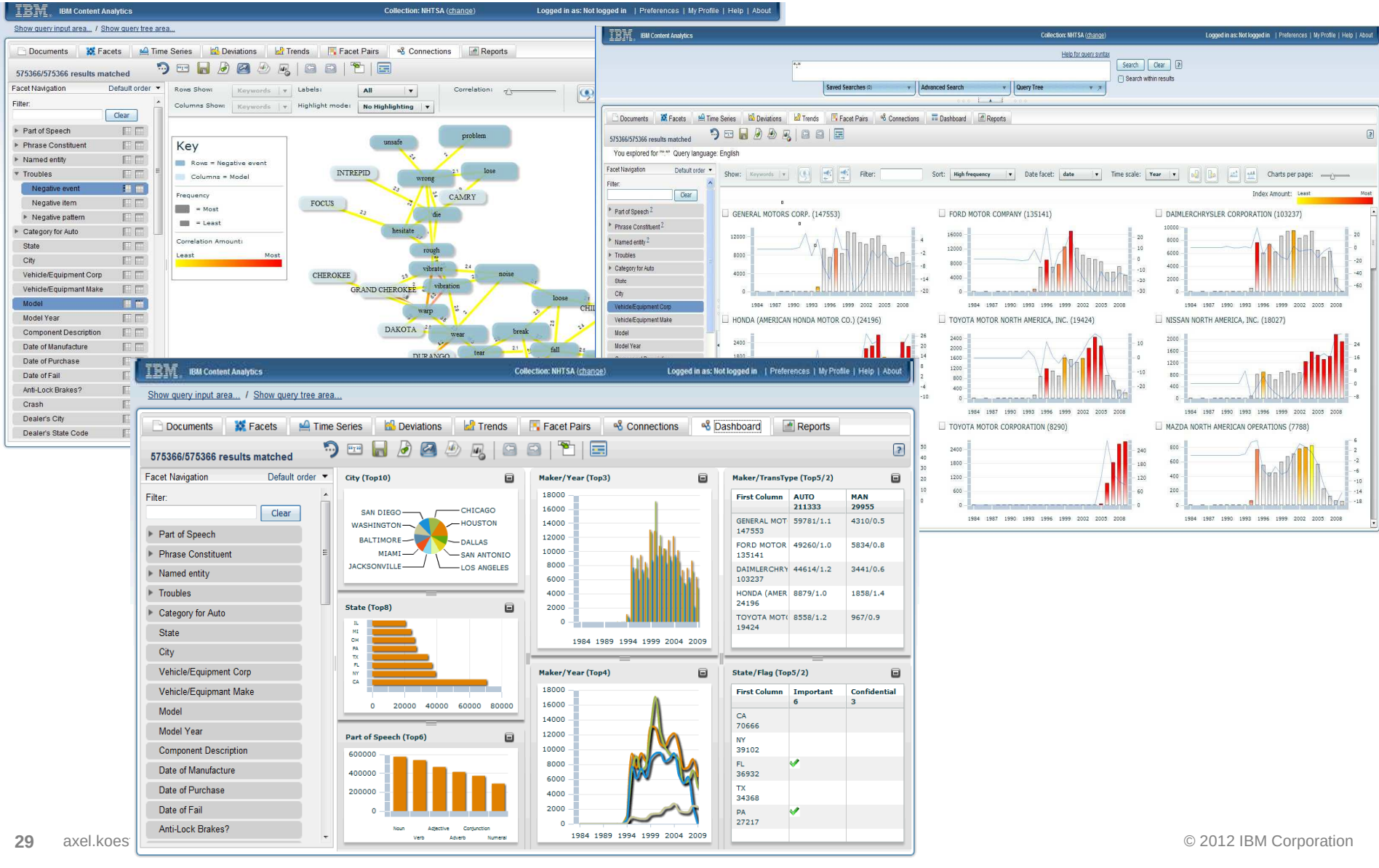
1. Help for complex problems with ambiguous error codes (or no codes)
2. Quick decision support for call centers on red warning light during journey
3. Car owners chat forum scan, public opinion, defect reports, …
4. Traffic flow prediction for navigation systems
5. …

axel.koester@de.ibm.com

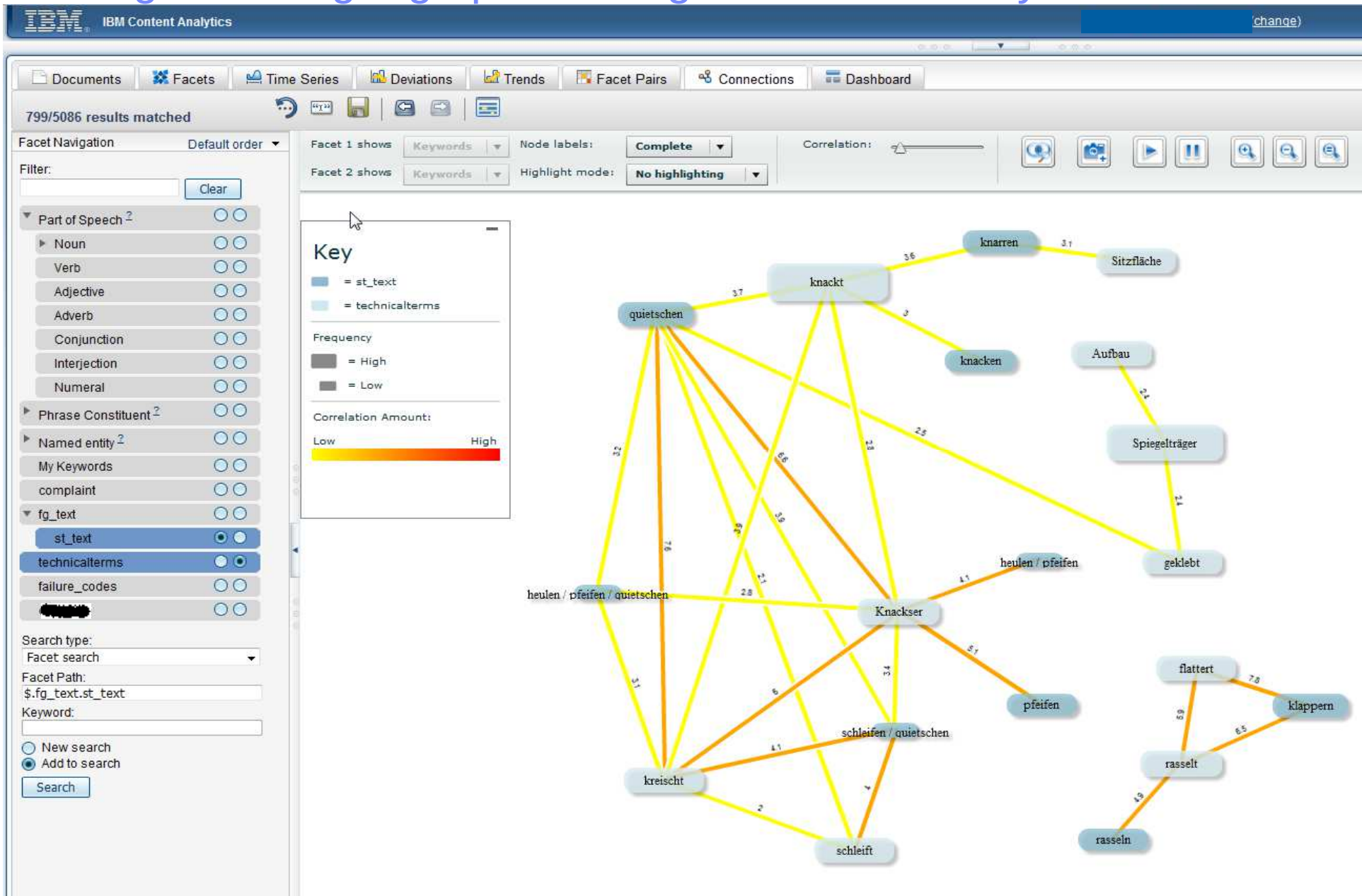
What do Watson projects 'look like'?

axel.koester@de.ibm.com

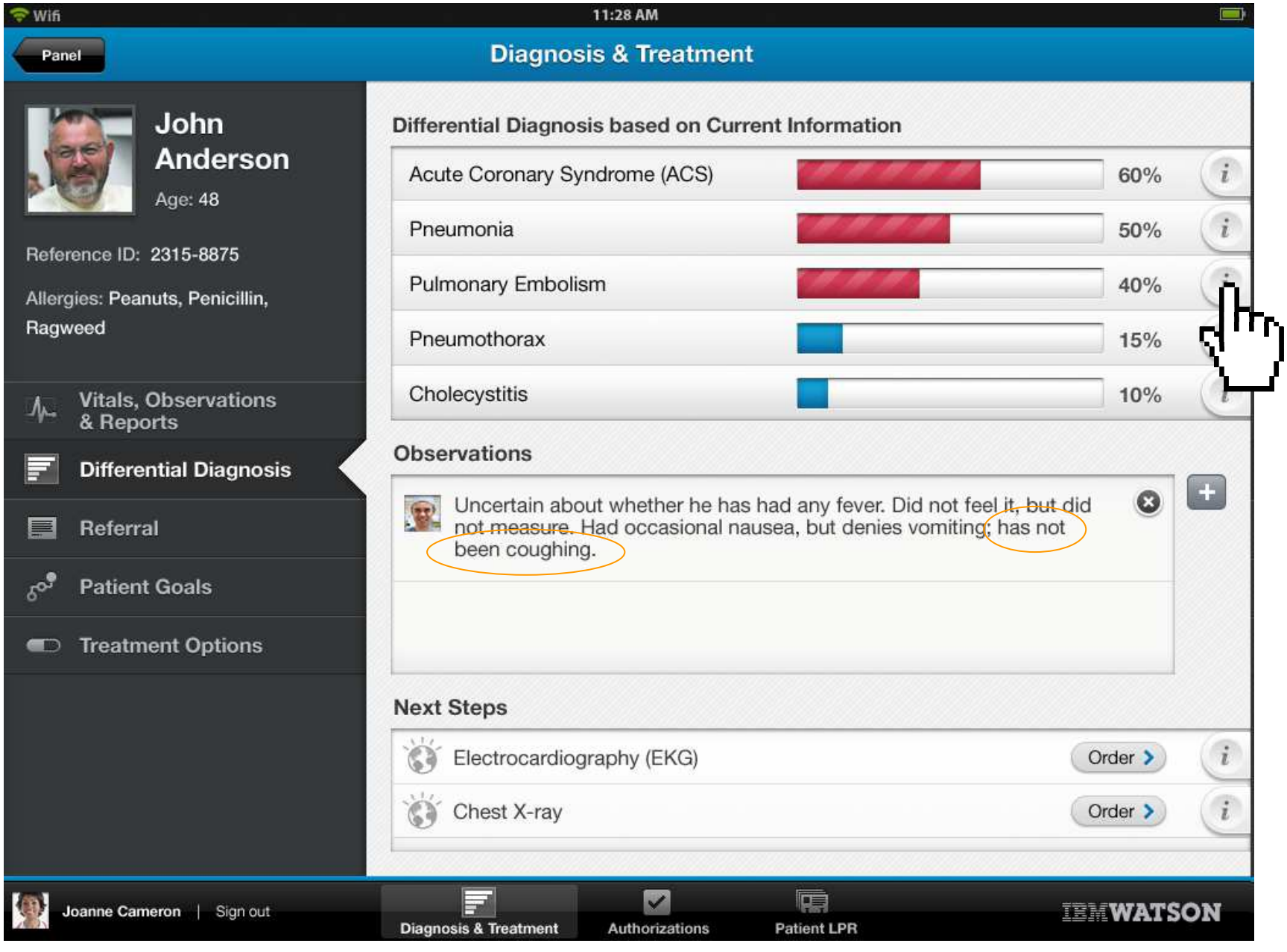
Mini-Watson = IBM Content Analytics

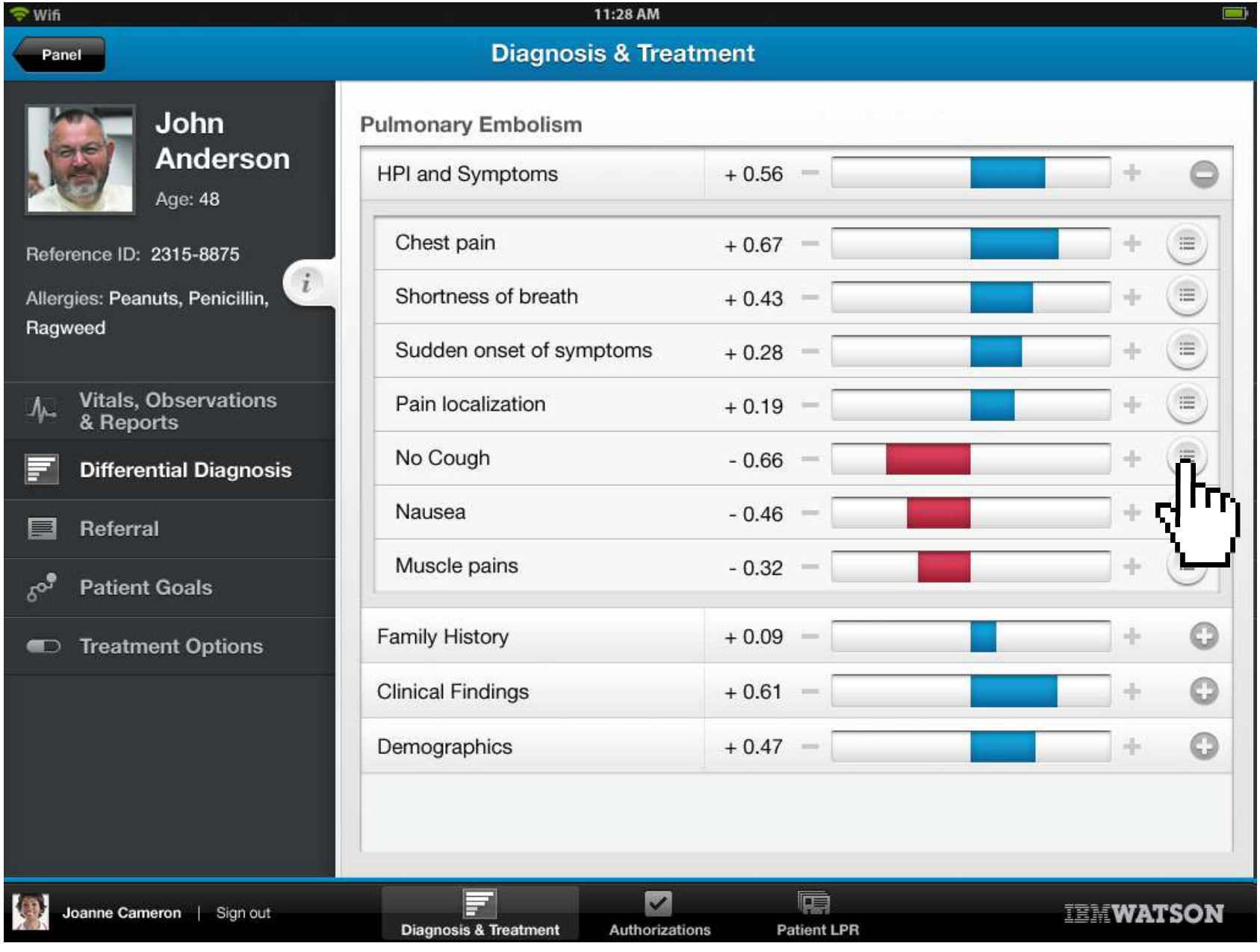


Ambiguous language processing in Content Analytics

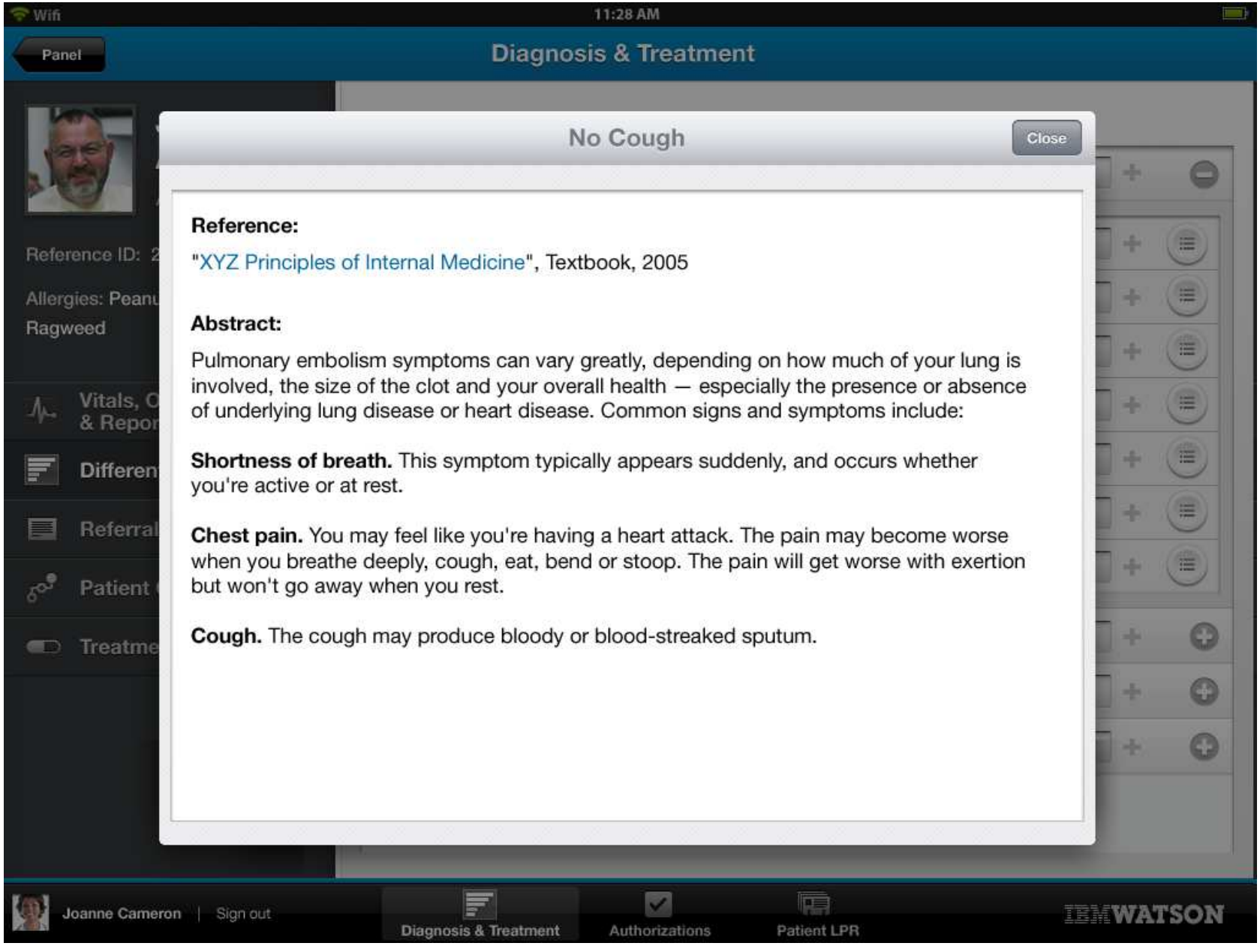


"Dr. med. Watson"





 axel.koester@de.ibm.com



 axel.koester@de.ibm.com

Financial rating:
Automatic evaluation of related news ("optimism" ranking)

Wind turbine planning at *Vestas* (InfoSphere BigInsights)

→ Analyze **petabytes of data** including weather reports, tidal phases, geospatial and sensor data, satellite images and deforestation maps to find the ideal placement of wind turbines.

1.222 node "SuperWatson" at Vestas Wind Systems (DK)

axel.koester@de.ibm.com
http://windsector.tumblr.com/post/11924311917/ibm-vestas-start-a-firestorm-for-wind-energy

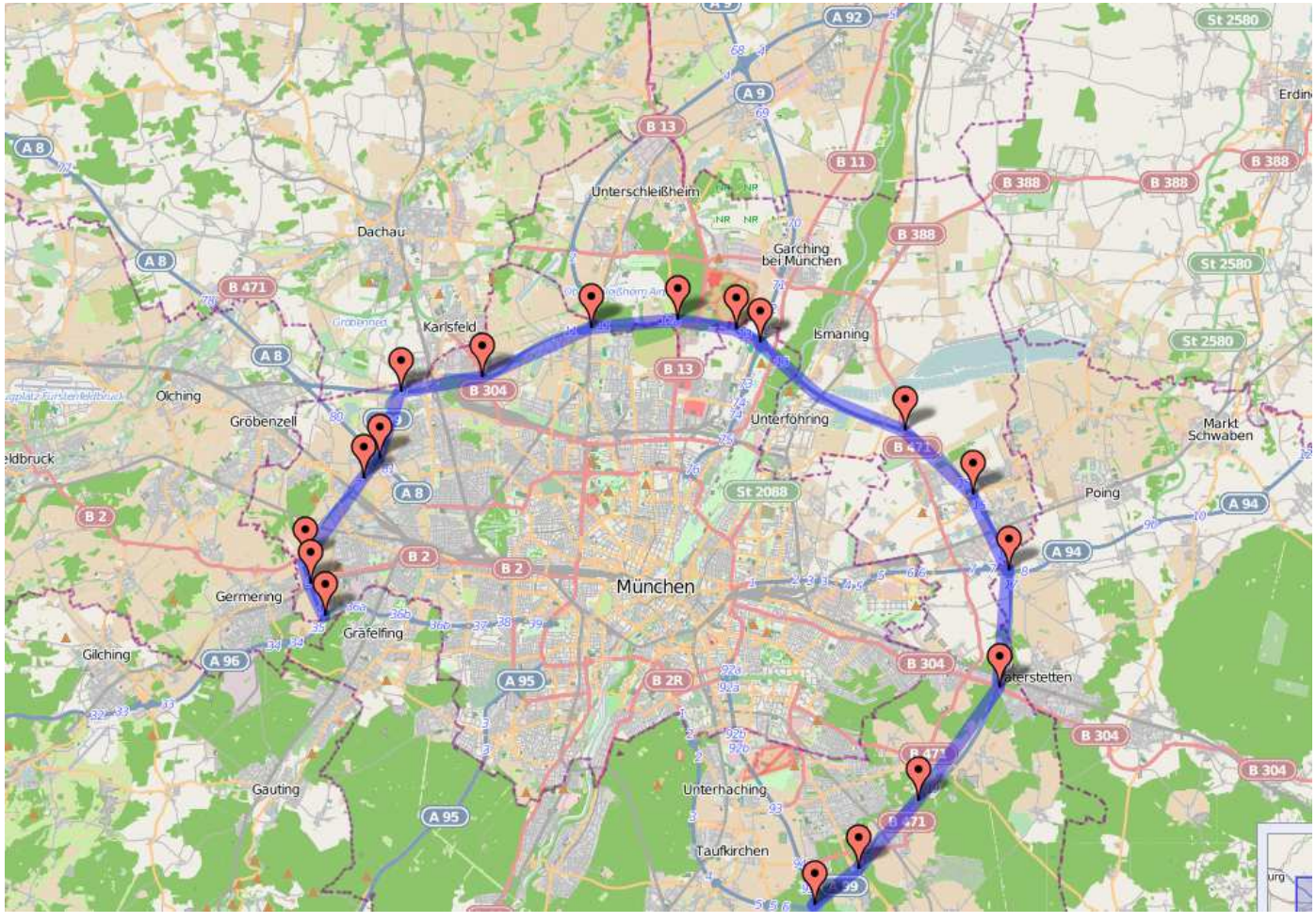
What "normal" Watsons look like (IBM pre-packaged)

Based on System x3630: Ultra-dense, storage-rich server for *BigInsights*

Configuration	Starter	Half Rack	1st Thru 4th Full Rack	Additional Full Racks
Usable Storage	Up to 108TB	Up to 324TB	Up to 684TB per rack	Up to 720TB per rack
User space	Up to 36TB	Up to 108TB	Up to 228TB per rack	Up to 240TB per rack

axel.koester@de.ibm.com

Real time traffic jam prediction in Munich (InfoSphere Streams)



95% reliable prediction of traffic jams near Munich from scanning induction sensors on A8, A9 and A99

85% reliability for 2 hours future jam prognosis

"Learned" patterns

http://www.bild.de/auto/auto-news/cebit/ibm-computer-watson-stau-vorhersagen-prognose-23004292.bild.html
http://www.computerwoche.de/hardware/data-center-server/2506319/

axel.koester@de.ibm.com

Lab Research
Cognitive Computing

axel.koester@de.ibm.com

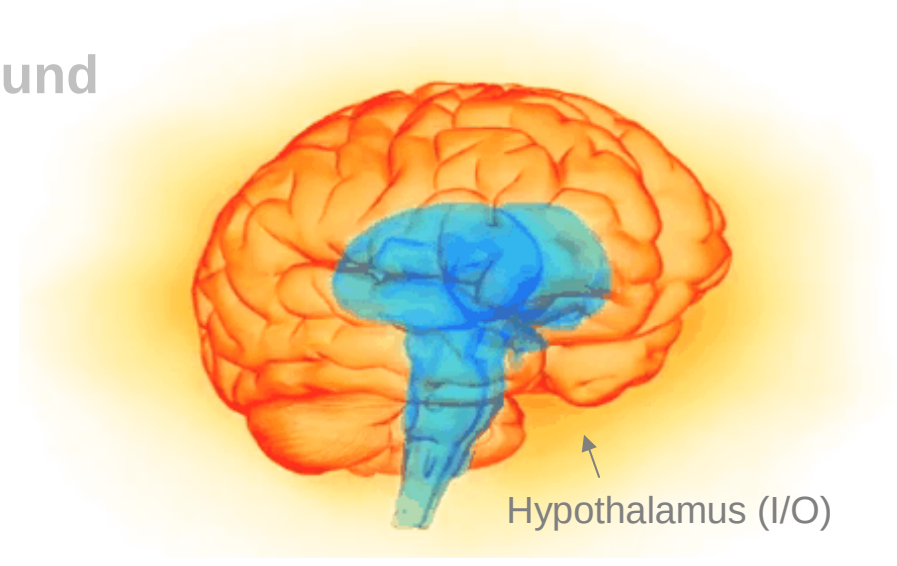
Human brains are incredibly good at "low power"

Recognize a face in a crowd – efficiently

Distinguish own **from** outside **sound**

Combine unrelated facts

Filter & distill information



20 Watt

How?

Brains are not 100% accurate.

axel.koester@de.ibm.com

http://www.ibm.com/smarterplanet/us/en/business_analytics/article/cognitive_computing.html

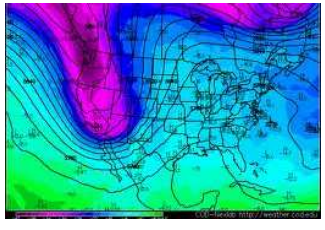
"Lower power" from abandoning the 100% accurate IT

Unreliable input data

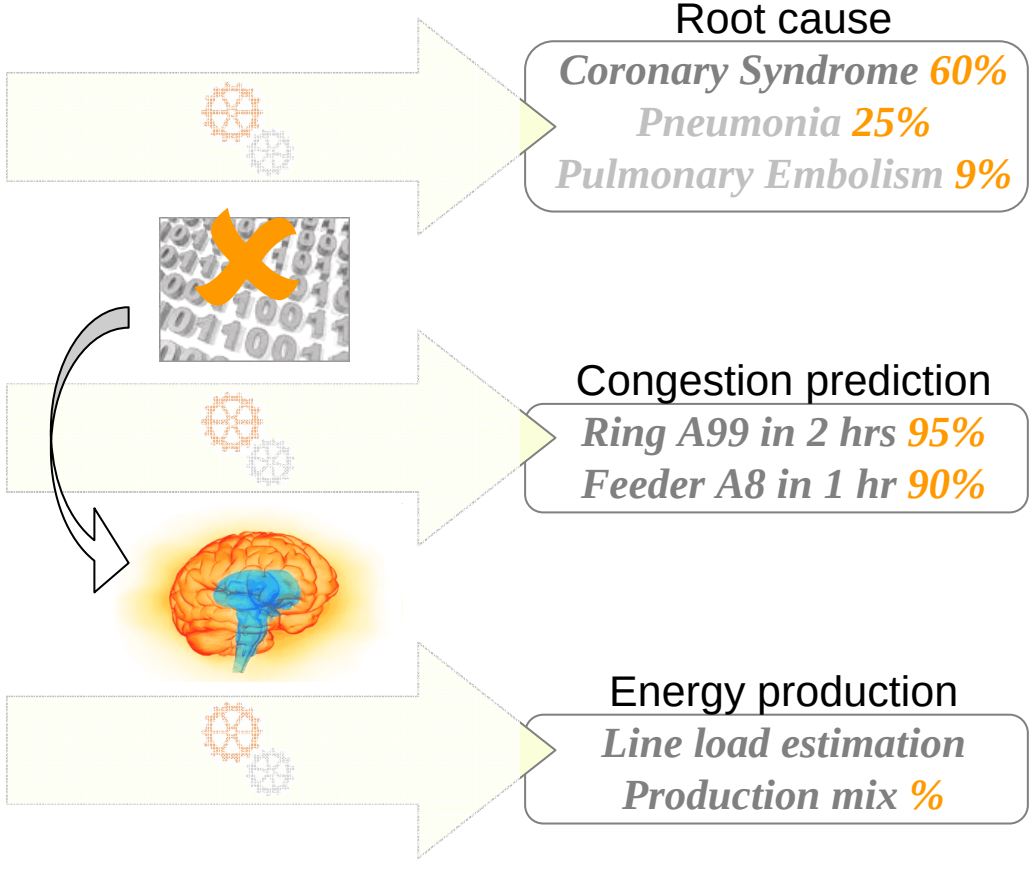
Patient symptoms

Road traffic sensors

Wind & sun forecast



:



Root cause

Coronary Syndrome 60%
Pneumonia 25%
Pulmonary Embolism 9%

Congestion prediction

Ring A99 in 2 hrs 95%
Feeder A8 in 1 hr 90%

Energy production

Line load estimation
Production mix %

:

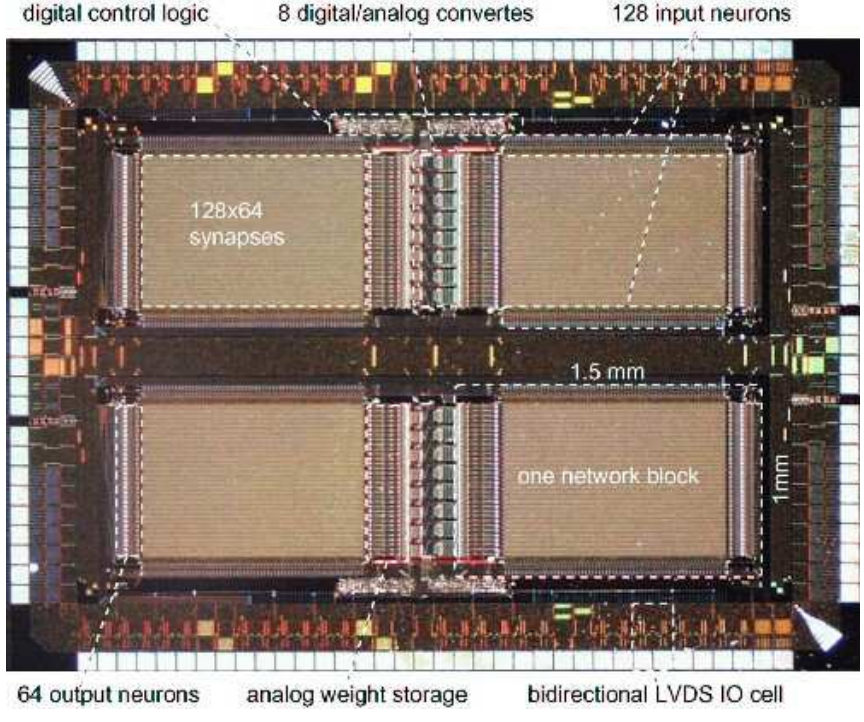
Approximation

Approximations based on unreliable data should not require bit-accurate processing !

axel.koester@de.ibm.com

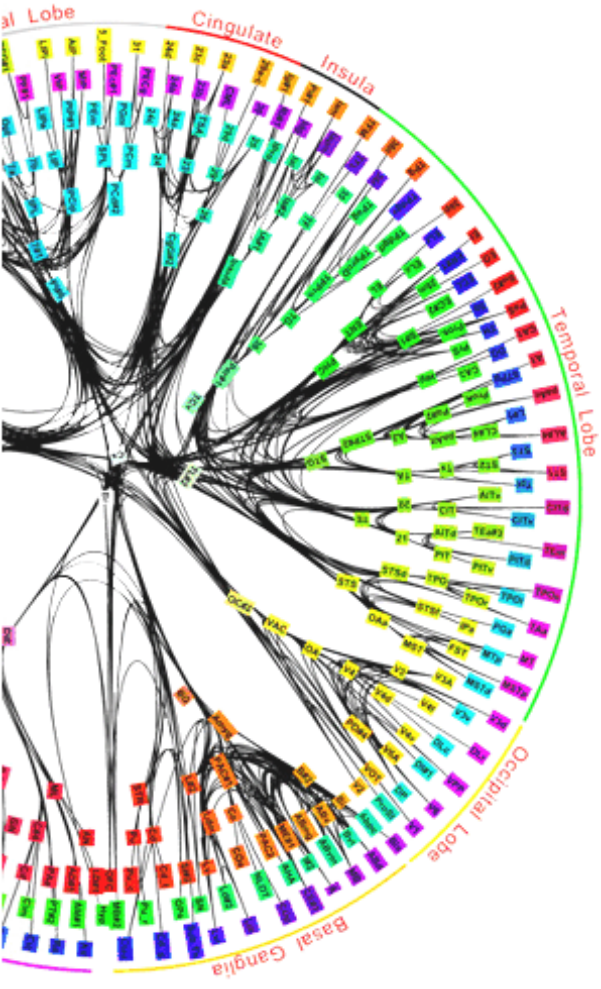
Future Microchips = analog technology?

- Digital = low power: **0 (no current)** or **1 (no resistance)**
- no more true at ultrahigh speeds, constantly *between* 0 and 1



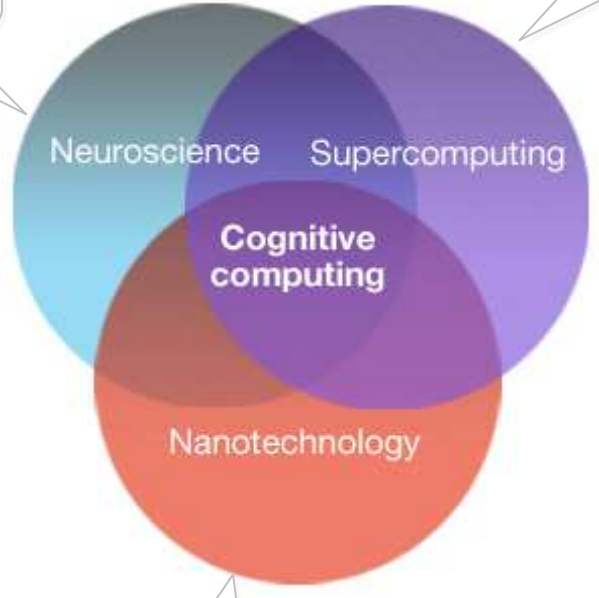
Source: University Heidelberg AnaloG Evolvable Neural Network

SyNAPSE and BlueBrain: Simulating the brain



understand *learning* down to synapses

explore large-scale brain simulations



design a chip that "learns" at molecular level

http://www.ibm.com/smarterplanet/us/en/business_analytics/article/cognitive_computing.html

"Cognitive Computing Research" at IBM

SyNAPSE – Systems of Neuromorphic Adaptive Plastic Scalable Electronics

"…simulating the entire human cortex, about 25 billion neurons, at full speed? To do that, you'll need to find 1000 times more computing power (than today's supercomputers)."

A SyNAPSE chip could do that at a fraction of the required energy.

"I'll have it ready for you within the next decade."

Dharmendra Modha, IBM
SyNAPSE program leader

The initial phase of the SyNAPSE program developed nanometer scale electronic synaptic components capable of adapting the connection strength between two artificial neurons in a manner analogous to that seen in biological systems (Hebbian learning), and simulated the utility of these synaptic components in core microcircuits that support the overall system architecture. *(Wikipedia)*

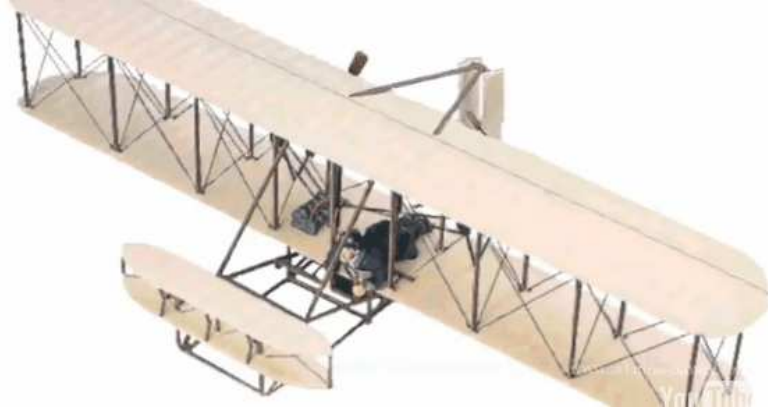
Research VP

Bruce Hillsberg
blh@us.ibm.com
Director, Storage Systems

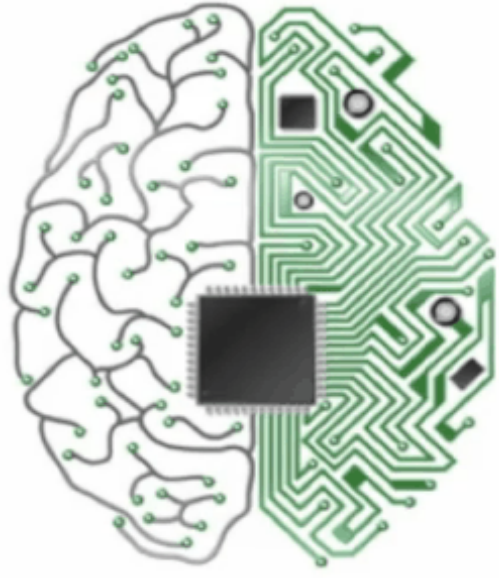
Dharmendra Modha
dmodha@almaden.ibm.com
Manager, Cognitive Computing

http://en.wikipedia.org/wiki/Hebbian_theory

Truth is in the middle



copied nature «



» optimized engineering

axel.koester@de.ibm.com

axel.koester@de.ibm.com