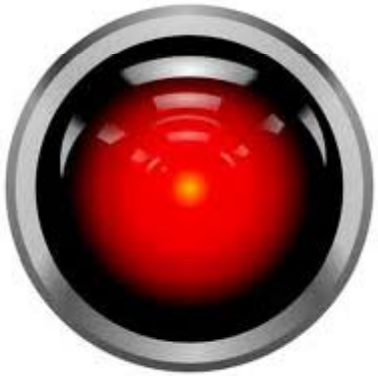
How Big Data and Artificial Intelligence Change the Game for Insurance Professionals



IBM WATSON

presented by

Jamie Bisker

Senior Analyst, P&C Insurance

Aite Group

Innovation Provocateur

November 2014

Agenda

Opening remarks: Big data, artificial intelligence, and insurance

Big Data: 4 Vs

Cognitive Computing : From Top Down to Bottom Up

Use of Learning Systems for Risk Discovery (U^2-RDD)

Opening Remarks

Opening remarks: Big data, artificial intelligence, and insurance

Big Data & AI already in use in the insurance industry

Are Actuaries the Original Data Scientists?

Expert underwriting, claims decisioning, business rules

OCR, language translation, image recognition…

Agenda

Opening remarks: Big data, artificial intelligence, and insurance

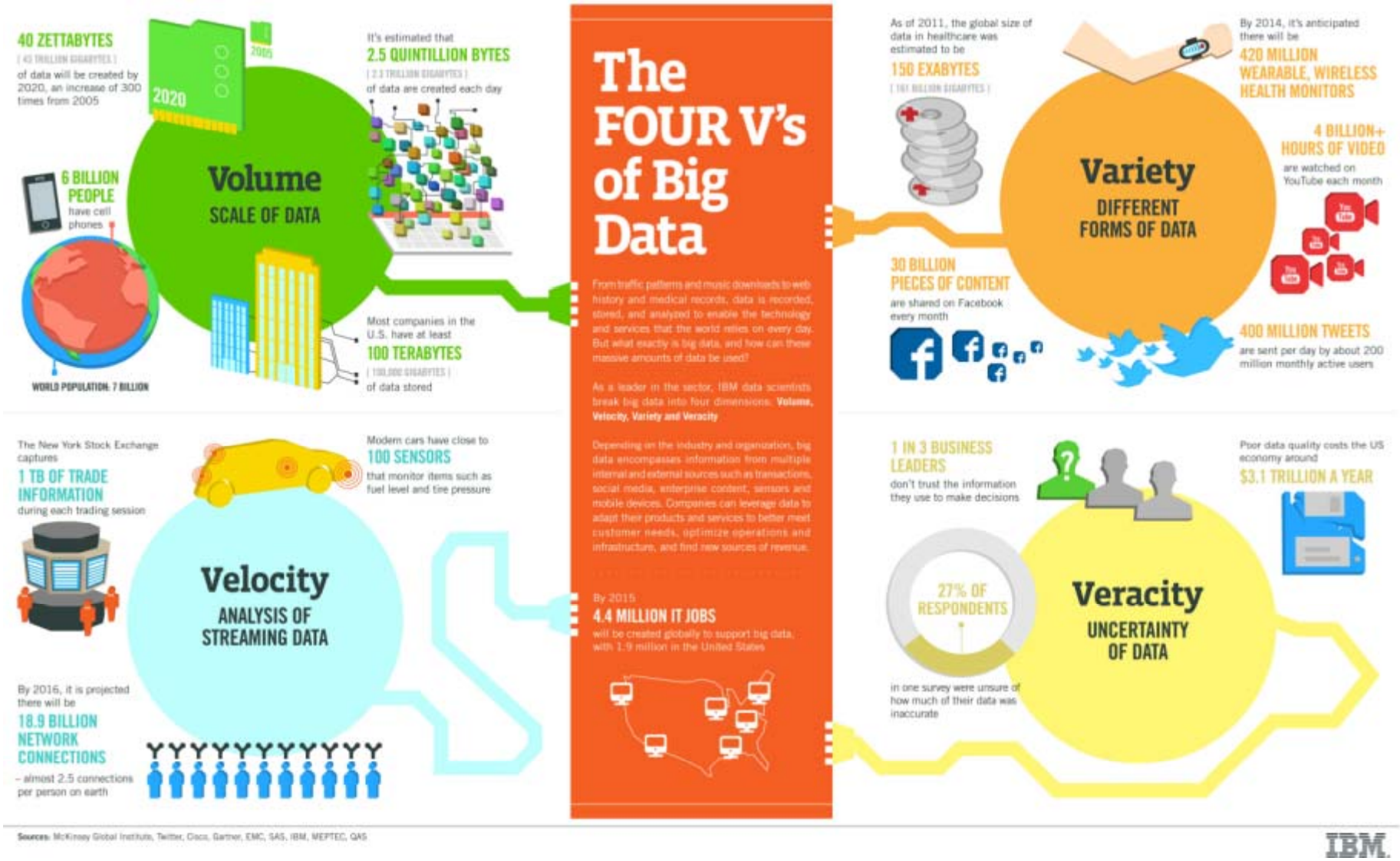
Big Data: 4 Vs

Cognitive Computing : From Top Down to Bottom Up

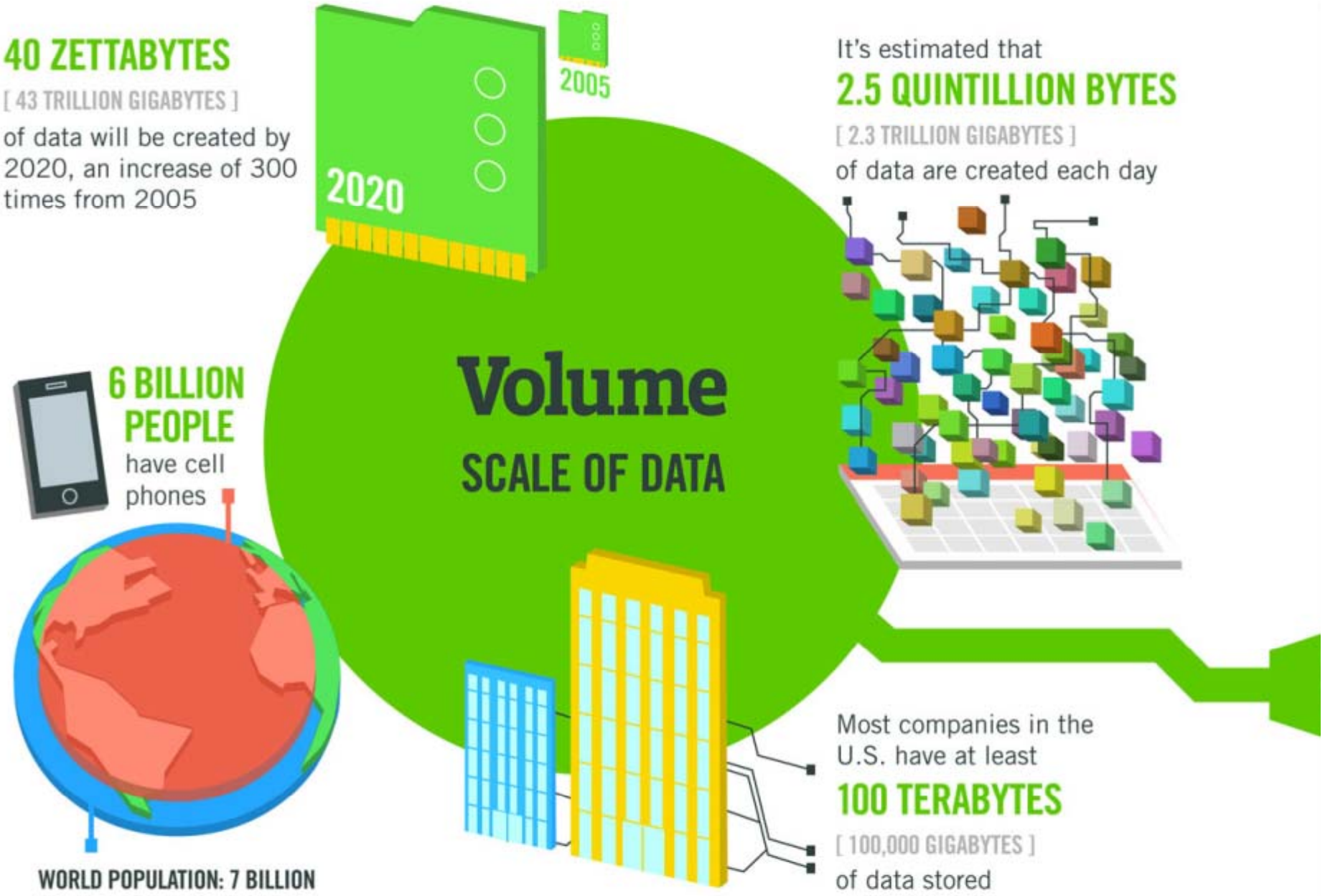
Use of Learning Systems for Risk Discovery (U2-RDD)


BIG
DATA

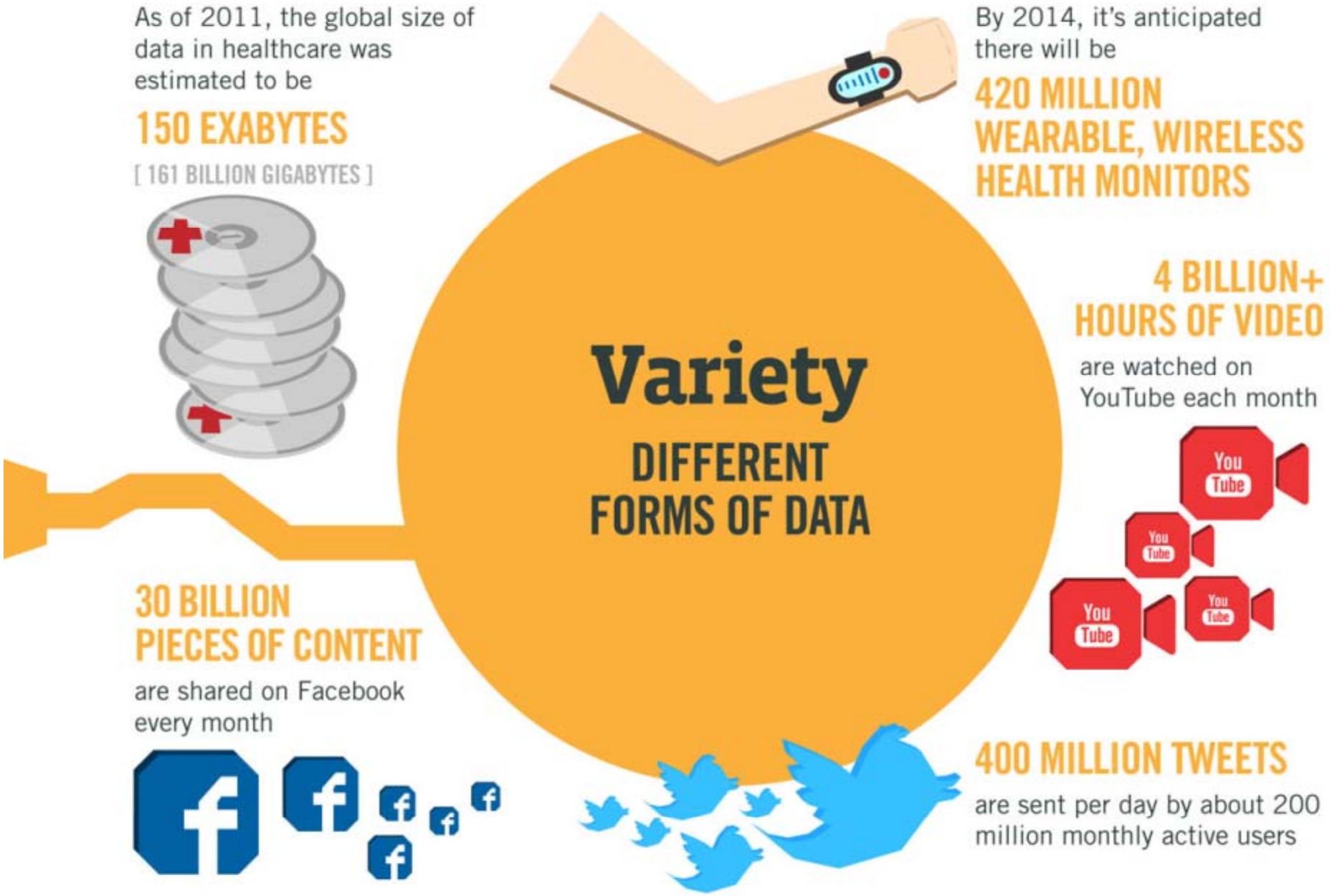
The 4 V's of Big Data:



The 4 V's of Big Data:



The 4 V's of Big Data:

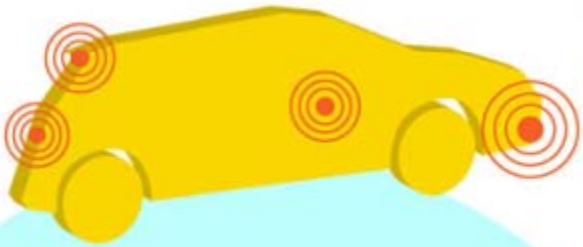


The 4 V's of Big Data:

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

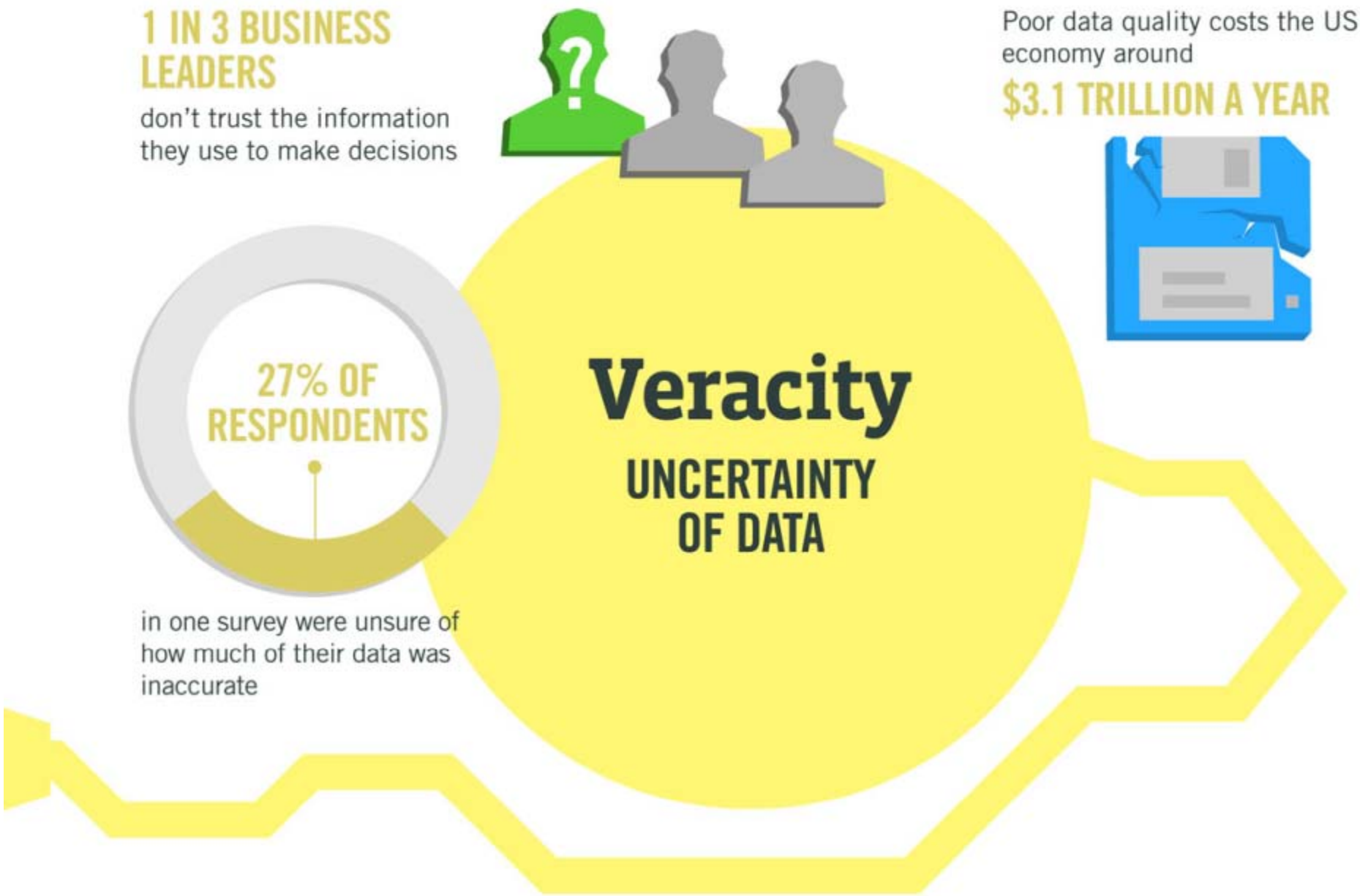
ANALYSIS OF STREAMING DATA

By 2016, it is projected there will be

18.9 BILLION NETWORK CONNECTIONS

– almost 2.5 connections per person on earth

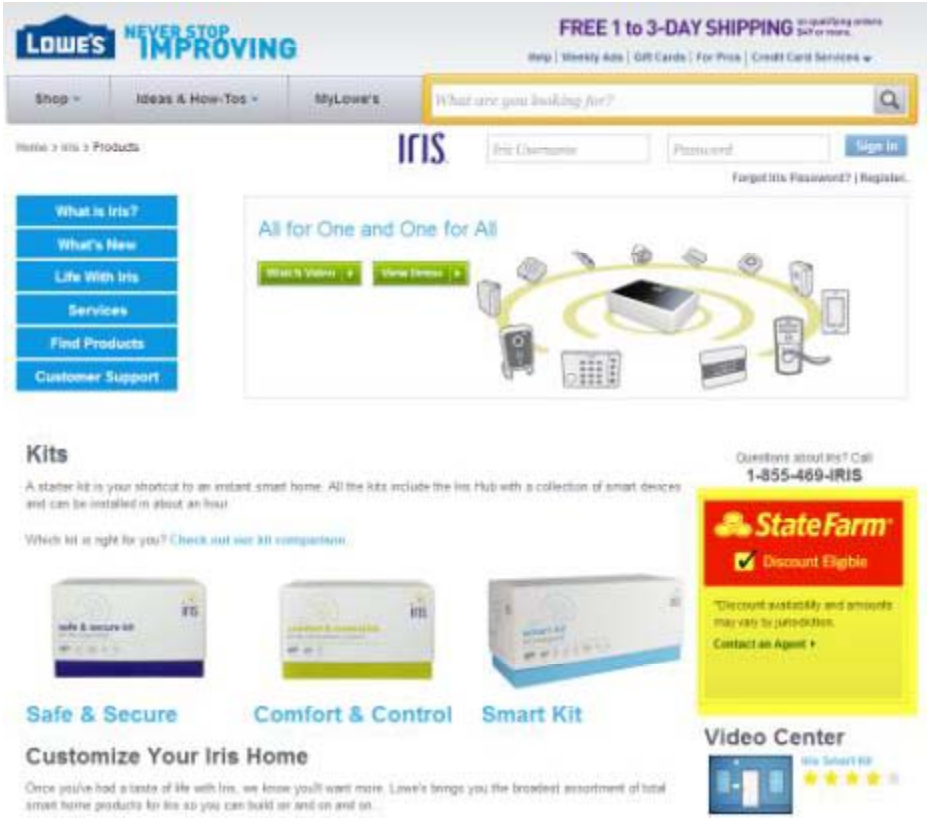
The 4 V's of Big Data:



Big Data and Insurance: Home monitoring

Canary: Home sensor

- HD Camera
- Microphone
- Night Vision
- Motion Detection
- Siren
- Temperature
- Air Quality
- Humidity



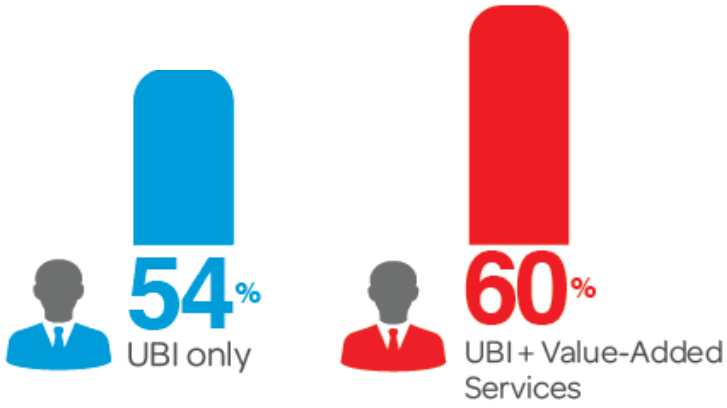
Big Data and Insurance: Usage-based insurance (UBI)

Offering popular value-added services in addition to the discount increases demand

Value-added services include:

UBI vs. UBI with value-added services

Enrollment interest @ 10% Discount



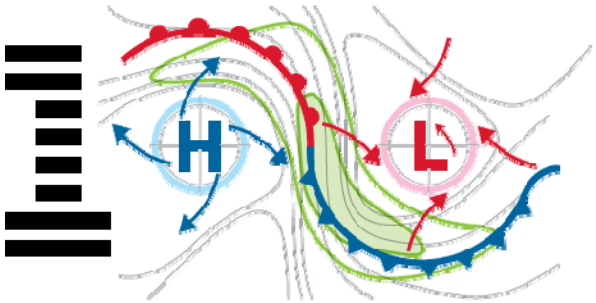
Big Data and Insurance: Beyond risk mitigation

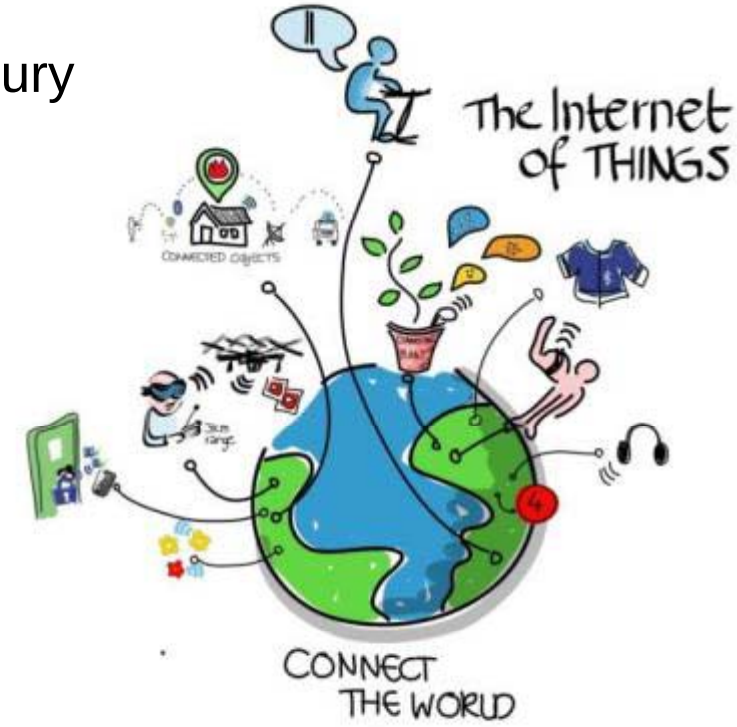
Mining social media networks for precision marketing, building predictive models of consumer behavior

IoT: Sensors for health data, video, audio, financial, appliances, etc. => Risk Wellness, Risk Discovery

High-frequency trading algorithms for Treasury and Risk

Weather Prediction not just forecasting: IBM's Deep Thunder for CATs, Energy, Wildfires, Fleet Mgmt…





Agenda

Opening remarks: Big data, artificial intelligence, and insurance

Big Data: 4 Vs

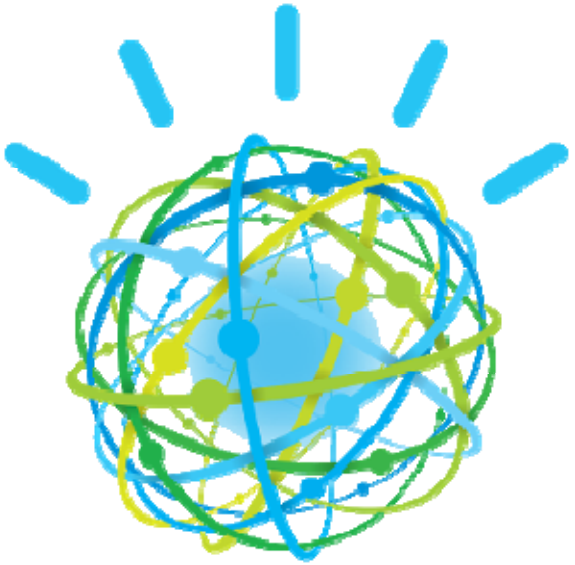
Cognitive Computing : From Top Down to Bottom Up

Use of Learning Systems for Risk Discovery (U2-RDD)

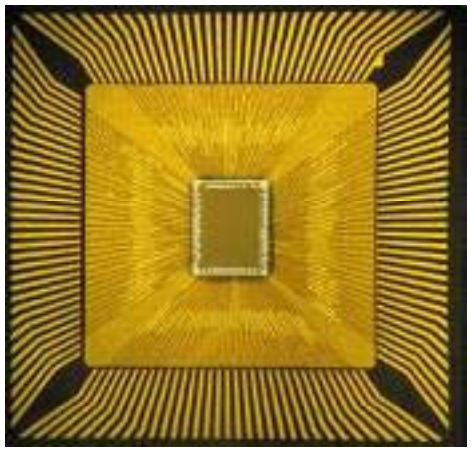
Cognitive Computing:

From Top Down

To Bottom Up



IBM WATSON



True North – Neurosynaptic Chip

Why does IBM pursue 'Grand Challenges' such as DeepBlue, BlueGene, and now Watson?

Grand Challenges forces IBM to stretch, collaborate and grow - they are an essential element of who we are

- **Chess: IBM's Deep Blue**
  - A finite, mathematically well-defined search space
  - Large but limited number of moves and states
  - Everything explicit, unambiguous mathematical rules

JEOPARDY! The IBM Challenge

- **Human Language: DeepQA and Watson**
  - Ambiguous, contextual and implicit
  - Grounded only in **human cognition**
  - Seemingly innumerable number of ways to express the same meaning

The Grand Challenge that created Watson:

"Can we design a system ***rivaling a human's ability*** to answer questions posed in natural language, interpreting meaning and context to retrieve, analyze and understand vast amounts of information in real-time?"

What is Watson?

Watson is a question answering information system from IBM

It was developed over several years as part of IBM's ***DeepQA*** research, and combines several technologies & areas of research

Why is it called Watson?

It is named after IBM's founder, Thomas J. Watson, Sr.

Initially, Watson was a computer system designed especially to play "Jeopardy!"

It represents a significant advancement in the field of natural language processing - question answering

What Watson Was Not ...

… a 'supercomputer' (like Blue Gene/P →)

It does use parallel processing to achieve high levels of performance required to play Jeopardy! and a large amount of active system memory (15 terabytes)

…*at the time*, a generic dialogue system intended for back-and-forth conversations
However, the current physician's assistant pilot that IBM Research's DeepQA Team is working on does include a dialogue mechanism using speech recognition

…a reasoning system that works *externally* with its own answers…

Many people feel that full artificial intelligence software could be used when in fact there are several easier tools available – text mining for example

Natural language understanding: Why is it so hard for computers to understand humans? A few clues…

- Noses that run and feet that smell?

- Ship by truck and send cargo by ship?
- How can a slim chance and a fat chance be the same, while a wise man and a wise guy are opposites?
- How can a house burn up as it burns down?
- Why do we fill in a form by filling it out?

- How does an alarm go off by going on?

What Computers Find Hard

Computer programs are natively **explicit, fast** and **exacting** in their calculation over numbers and symbols….But **Natural Language** is implicit, highly contextual, ambiguous and often imprecise.

Person	Birth Place
A. Einstein	ULM

Structured

Unstructured

- Where was X born?

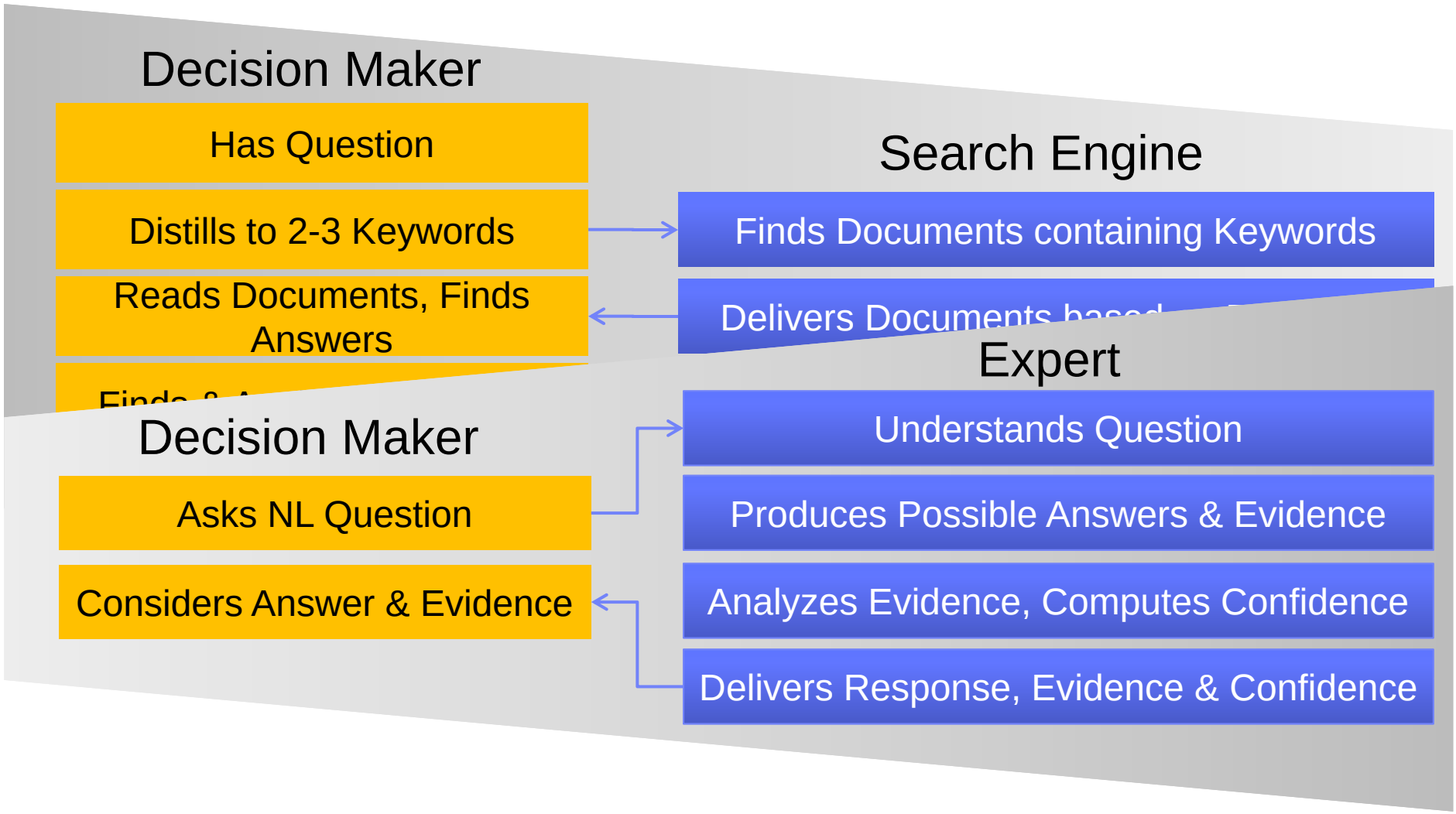
  One day, from among his city views of Ulm, Otto chose a water color to send to Albert Einstein as a remembrance of Einstein´s birthplace.

Person	Organization
J. Welch	GE

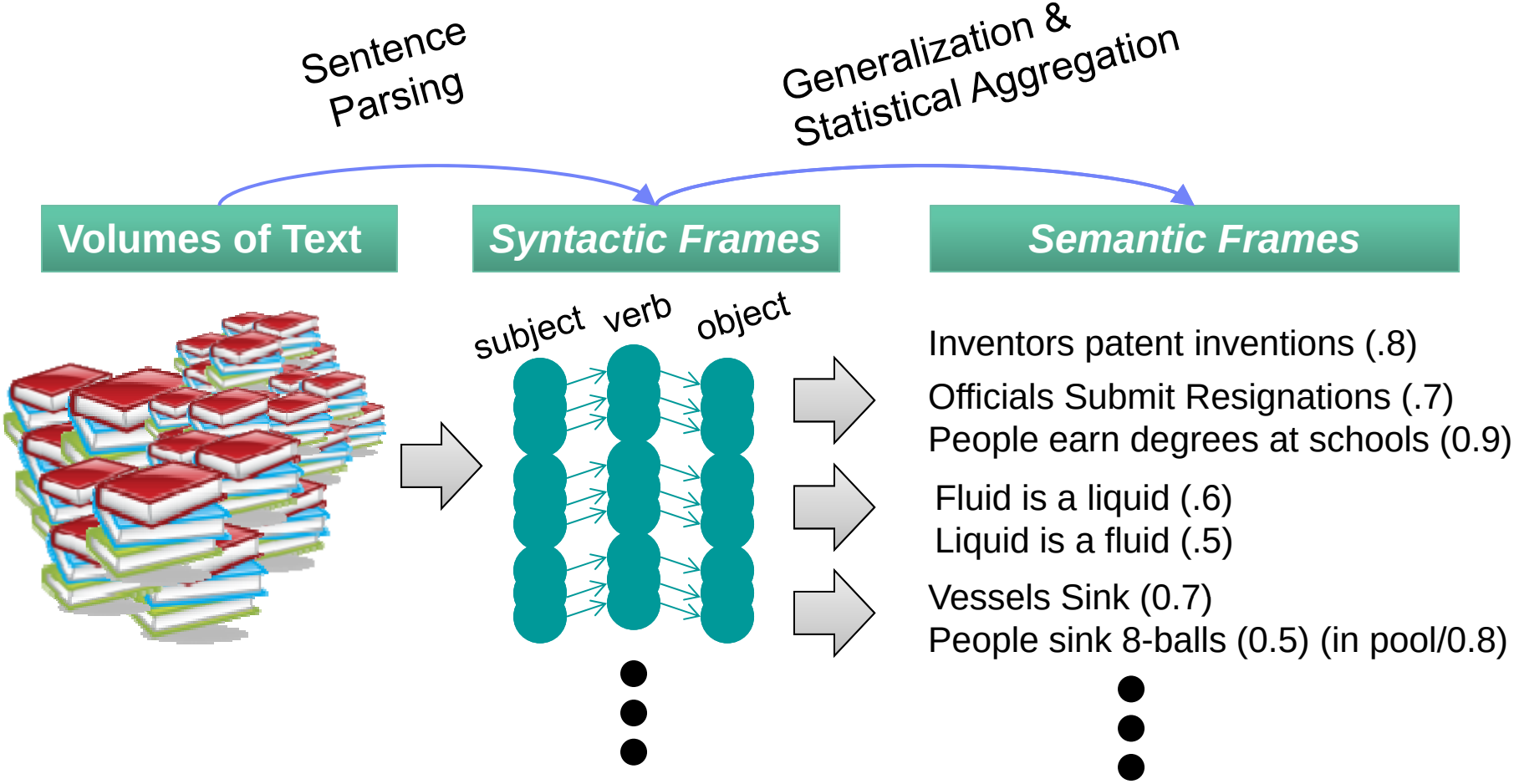
- X ran this?

  If leadership is an art then surely Jack Welch has proved himself a master painter during his tenure at GE.

Informed Decision Making: Search vs. Expert Q&A



Loading Watson - automatic learning from “reading” – the system was not connected to any external network during competition



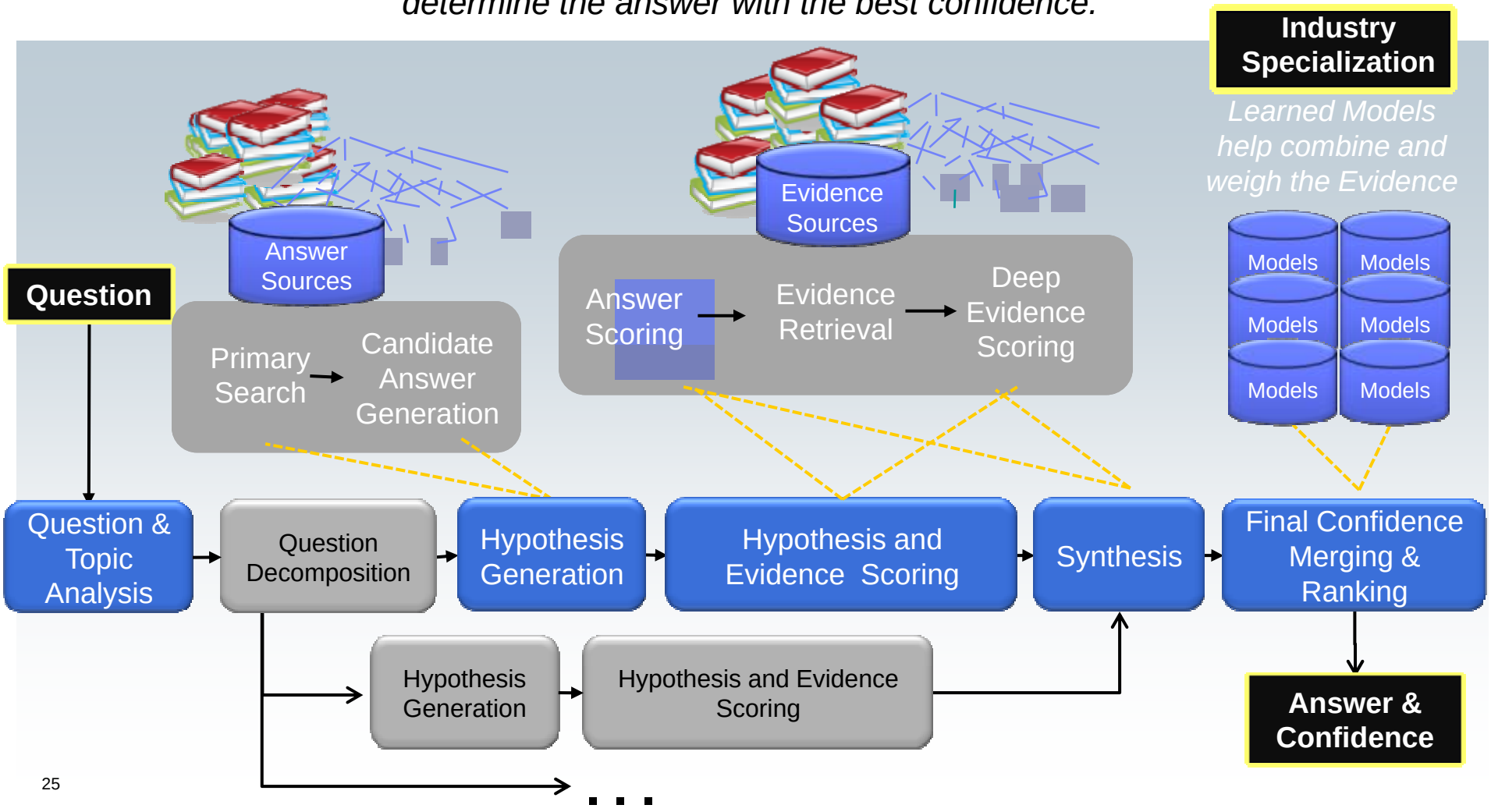
Once loaded with domain knowledge, the system begins its work…
DeepQA Architecture Phases:

- **Question Analysis:** Variety of NLP algorithms analyze the question to attempt to figure out what it is asking
- **Primary Search:** Retrieve content related to the question (including both unstructured text documents/passages and structured knowledge-base entries)
- **Candidate Answer Generation:** From the retrieved content, extract the words or phrases that could be possible answers
- **Evidence Retrieval:** For each Candidate Answer, retrieve more content that relates that answer to the question
- **Evidence Scoring:** Many algorithms attempt to determine the degree to which the retrieved evidence supports the candidate answers.
- **Merging and Ranking:** Consider all the scored evidence to produce a final ranked list of answers with confidences

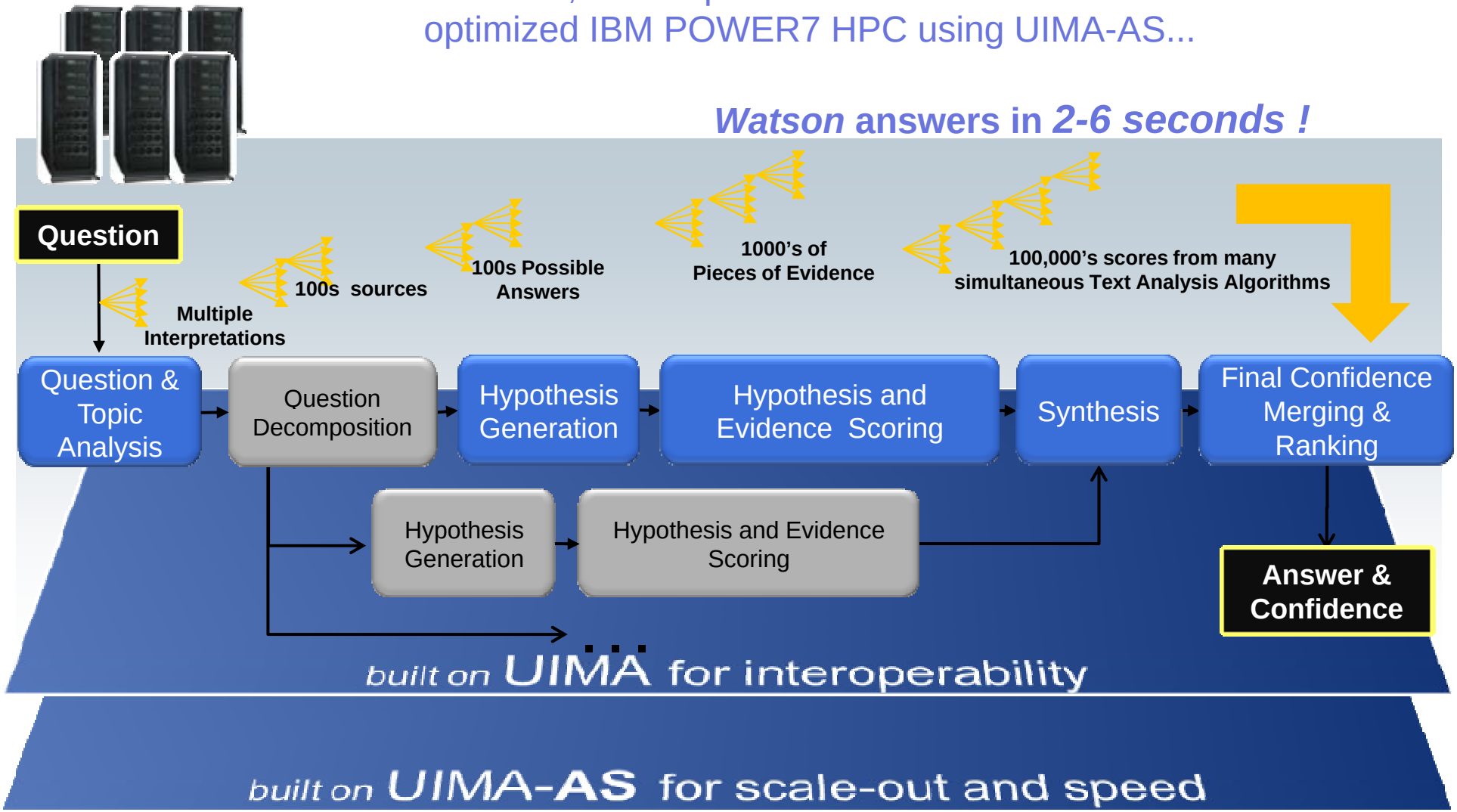
DeepQA: The Technology Behind Watson

Massively Parallel Probabilistic Evidence-Based Architecture

DeepQA *generates and scores many hypotheses using an extensible collection of* ***Natural Language Processing***, ***Machine Learning*** *and* ***Reasoning Algorithms.*** *These gather and weigh evidence over both unstructured and structured content to determine the answer with the best confidence.*



One Jeopardy! question can take
2 hours on a single 2.6Ghz Core (i.e. – a laptop).
However, when optimized & scaled out on a workload optimized IBM POWER7 HPC using UIMA-AS...



Smarter Insurance and Watson…Tomorrow

Today, Watson could be thought of as a very thorough, but uneducated Research Librarian:

Speculations on Tomorrow?

- Working on tasks that require **reasoning** about questions, the answers uncovered, and implications**:**
  - Holding a dialogue about the question asked
  - Does this action fit within our strategy?
- A Watson-type system could use a **business ethics** knowledge base for corporate, regional, national, or situational events:
  - reviewed project or strategic decisions
  - support an executive or a manager in review of unit/individual actions
- Extend its capabilities via use of **existing intelligent system tools**: such as IBM's ILOG, SPSS, or Entity Analytics could empower much broader reasoning and higher quality responses.

Watson has been busy… there's now a Watson Group

Watson Today: 24 times faster, smarter and 90 percent smaller; IBM has shrunk Watson from the size of a master bedroom to three stacked pizza boxes. Cloud-based, mobile enabled.

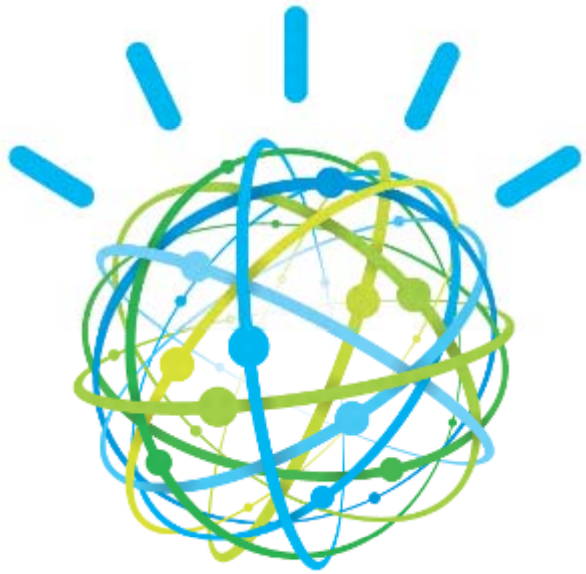
- **Watson Discovery Advisor**: Save researchers the time needed to pore though millions of articles, journals and studies. After quickly reading through, determining context and synthesizing vast amounts of data, it will help users pinpoint connections within the data that can strengthen and accelerate their work

- **Watson Analytics**: Service that allows users to explore Big Data insights through visual representations without the need for advanced analytics training.

- **Watson Explorer**: provides a unified view of all of a user's information. The service provides data discovery, navigation and search capabilities that are secure, unified and span a broad range of applications, data sources and data formats – both inside and outside an enterprise.

- **Watson Foundations**: tools and capabilities to tap into all relevant data – regardless of source or type.

Questions ?



Agenda

Opening remarks: Big data, artificial intelligence, and insurance

Big Data: 4 Vs

Cognitive Computing : From Top Down To Bottom Up

Use of Learning Systems for Risk Discovery (U^2-RDD)

Overview

There are known knowns.
These are ***things we know that we know*.**

There are known unknowns.
That is to say, ***there are things that we know we don't know*.**

But there are also unknown unknowns.
There are things ***we don't know we don't know*.**

- Donald Rumsfeld

Read more at
http://www.brainyquote.com/quotes/quotes/d/donaldrums148142.html#mUeoxsr2YZzRBwDt.99

Discovering and connecting to UUs determine impacts

Current knowledge mining capabilities are sufficient to expose information within bodies of structured and unstructured data based on search criteria via:

- Review of emergent information sources (news feeds, web sites, periodicals, academic papers for example)
- Reviewing any corpus in a more historical context
- Web crawling programs or robots

The domains of a given Risk Discovery are established
Ex. - risks to commercial establishments, individuals, families, etc.;

Does X threaten families? Does Y limit the exposure of T?
Does contract language include…

A series of **standing DQA queries works against these corpora**: this becomes a portion of a use case.

Discovering and connecting UUs to determine impacts; reasoning about risks

- DQA's reasoning is limited to the context of the corpora it reviews;
- Additional reasoning must be provided – or perhaps the mechanisms can be tuned to this function;
- Example: Does Fred risk death by driving to Paris? Where is Fred now? Lyon. Yes, Fred faces the risk of death when operating a motor vehicle (based on what has been read – a report on highway deaths, specific information for the routes from Paris to Lyon, etc.)
- DQA reads text that contains information on the risks in question – there would have to be *explicit statements* about the risk of driving and/or the risk of driving between cities, Lyon and Paris, etc.;
- GOAL: a mechanism that can establish that a risk exists or is likely to exist, and then a DQA-based mechanism could look for linkages;
- Can DQA be tuned to such that the lexicon, ontologies, etc., are sufficiently primed with risk (or any other topic for that matter) context so that causal relationships can be discovered without explicit (meaning brittle) coding to do so??

Autonomous Risk Discovery Research:

ARDUOUS:
Autonomous Risk Discovery of Unknown Or Unlikely Scenarios

A new identification mechanism that tags risk event streams (RES) as they are automatically discovered (ex. What is trending on Twitter, news feeds, or financial markets?)

A supervisor reviews the blackboard for connectable events (those that can be connected to existing contracts for example) and scores them;

UUs need to be themselves coded on a spectrum:

- UU_1 = discoverable unknowns
- UU_3 = deeply hidden unknowns
- UU_5 = unknowable unknowns

RESs also need to be coded:

- RES_1 = easily linkable risk or event stream
- RES_3 = sparsely linkable
- RES_5 = unique risk or event streams

Autonomous Risk Discovery Research:

ARDUOUS:
Autonomous Risk Discovery of Unknown Or Unlikely Scenarios

- With Sherlock Holmes, uncanny knowledge was explained away as being a result of a lack of attention to details
- Event streams or scenarios need to compared to risk hypothesis for match potential and scoring; Can a hypothesis be created such that one element is a known thing or event and the another aspect is artificial – made up ?
- Such scenarios can be positive (new markets, new opportunities) or negative (new dangers or risks)
- What does event X mean to us (our company)? What risk hypothesis will negatively impact us? Can opportunity cost be calculated on Y?
- The concept of what is unknown depends on the observer; there are things not known by anyone, but others things that are known to someone (U_2 ?) not known because a question was not asked (discoverable). I wonder if there is a large quantity of fissionable material in the Earth's crust that will be compressed at a future date by tectonic events that will cause a natural atomic bomb to explode and cause an earthquake, tsunami, sinkhole, etc. ??

ARDUOUS - Autonomous Risk Discovery of Unknown Or Unlikely Scenarios

- Event streams or scenarios need to compared to risk hypothesis for match potential and scoring; Can a hypothesis be created such that one element is a known thing or event and the another aspect is artificial – made up ?

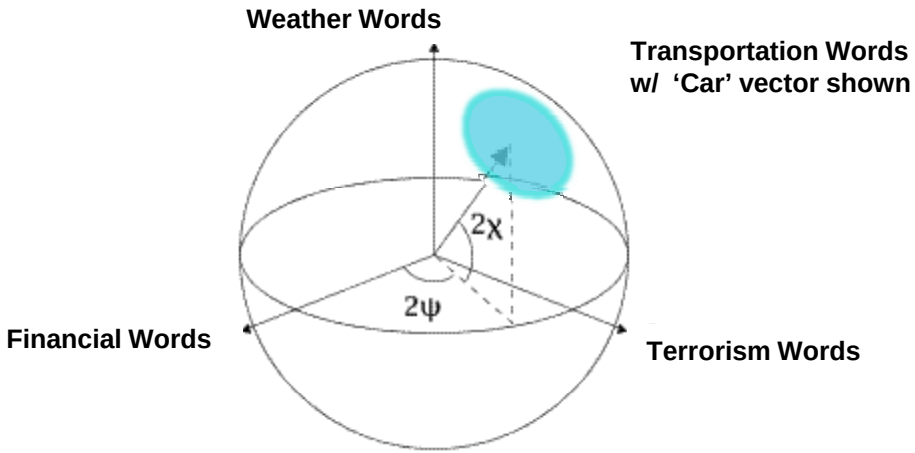
- Such scenarios can be positive (new markets, new opportunities) or negative (new dangers or risks)

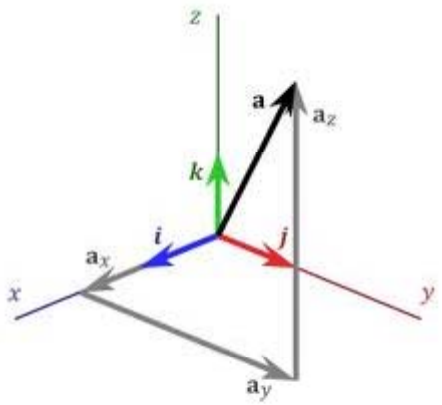
- What does event X mean to us (our company)? What risk hypothesis will negatively impact us? Can opportunity cost be calculated on Y?

- The concept of what is unknown depends on the observer; there are things not known by anyone, but others things that are known to someone (U2 ?) not known because a question was not asked (discoverable). I wonder if there is a large quantity of fissionable material in the Earth's crust that will be compressed at a future date by tectonic events that will cause a natural atomic bomb to explode and cause an earthquake, tsunami, sinkhole, etc. ??

ARDUOUS - Autonomous Risk Discovery of Unknown Or Unlikely Scenarios

- Use ”Risk Vectors” in a way that is similar to how HNC used “Context Vectors”
- Explore an indentified RV for connecting context; compare connecting context and any scoring so that they bubble to the top
- Scoop top CCs and begin a context based search of static and active corpora
- Form hypothesis from this and report to a ‘workbench’ mechanism for additional queries and further automated discovery searches.

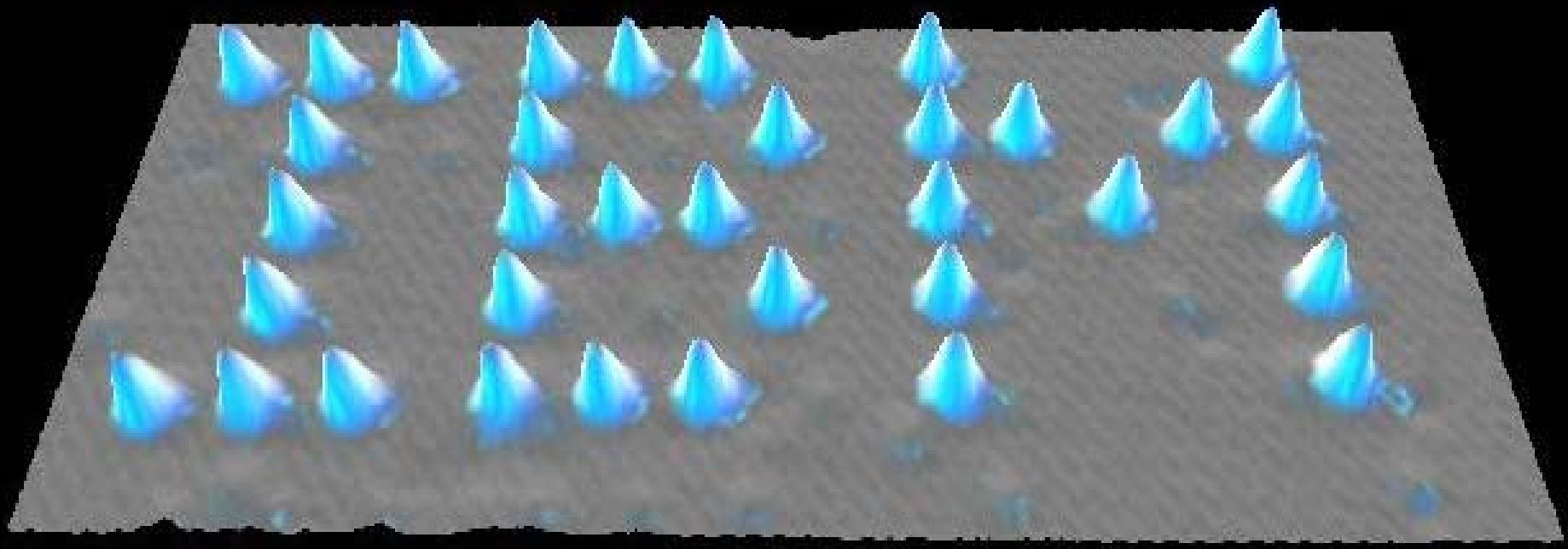




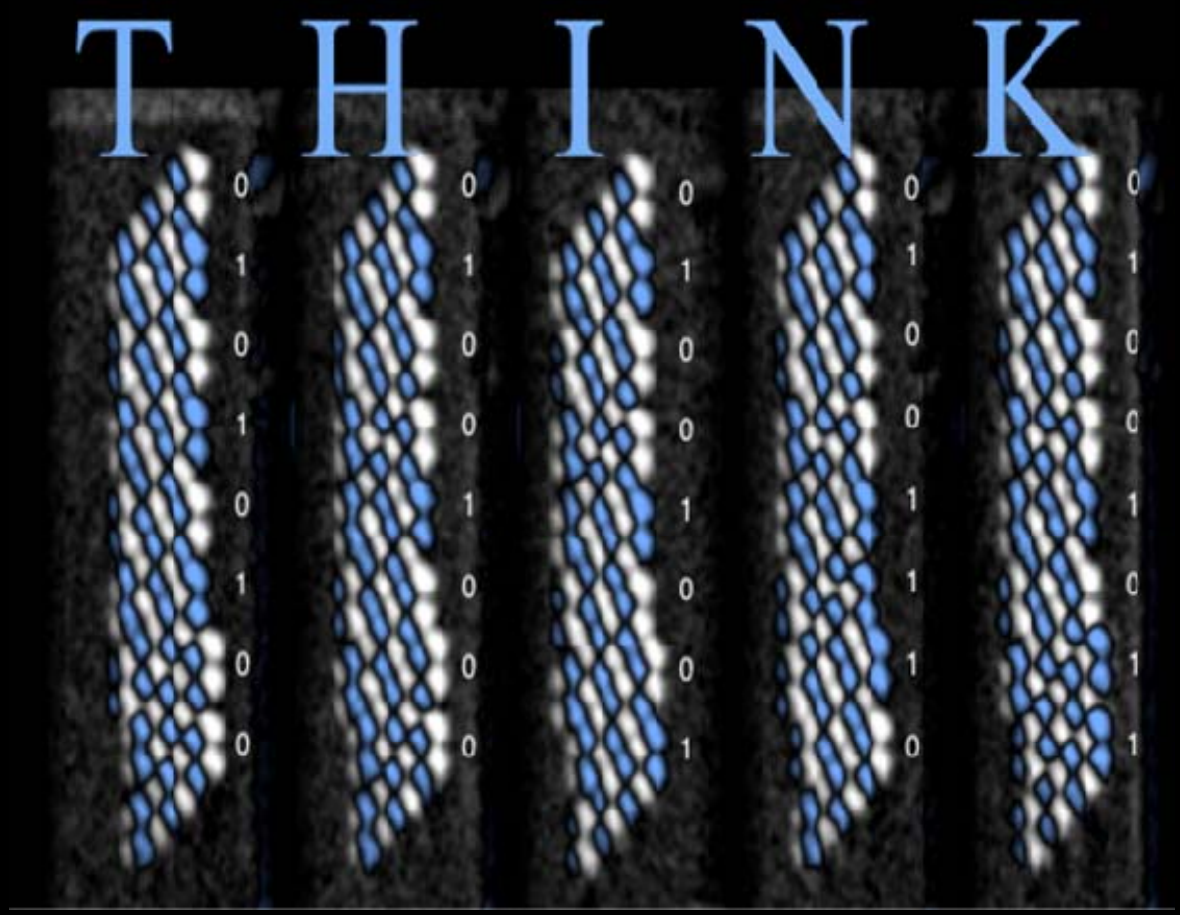
ARDUOUS - Autonomous Risk Discovery of Unknown Or Unlikely Scenarios

- http://senseclusters.sourceforge.net/
- http://www.google.com/patents/US5619709

From the Bottom Up:
Nanotechnology meets Brain



Cognitive Systems: *Atomic storage*



Atomic limits of magnetic storage

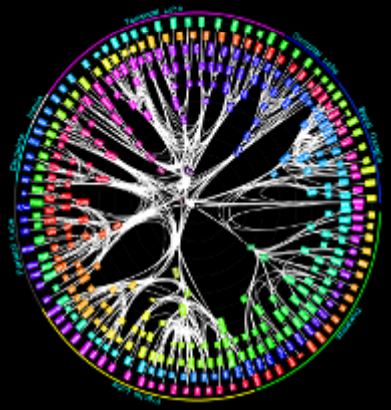
96 iron atoms store one byte of data

Note: There are ~11 sextillion iron atoms in 1 gram of iron;

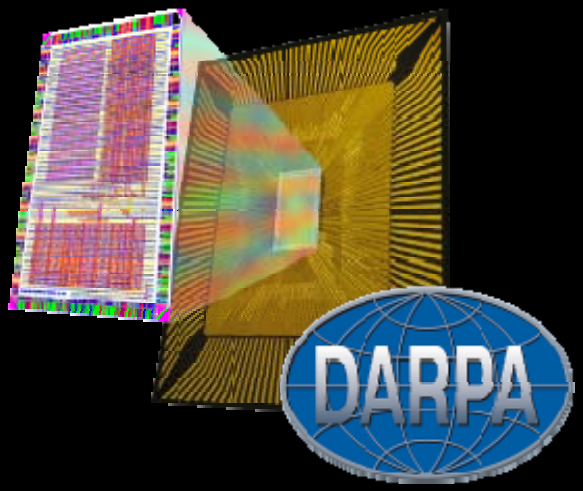
13,067,848,000,000,000,000 Bytes (13 Quintillion Bytes) in 1 BB

Cognitive Systems: *SyNAPSE*

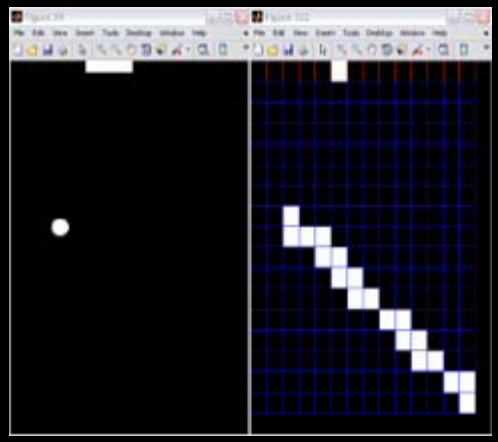
- "Neuron" and "synapse" -like computing model – the bottom-up, biologically inspired model of computing
- Systems learn through analytics/experience
- Advantages: Ultra energy-efficient, flexible, learning



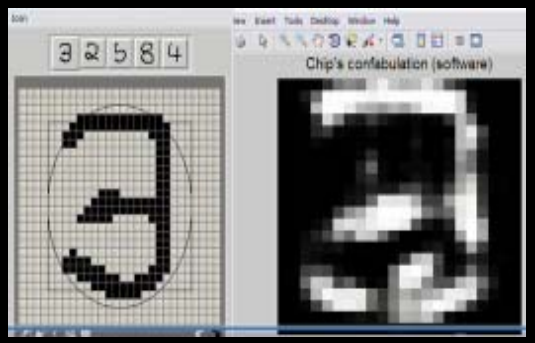
Wiring diagram – monkey brain



'True North'

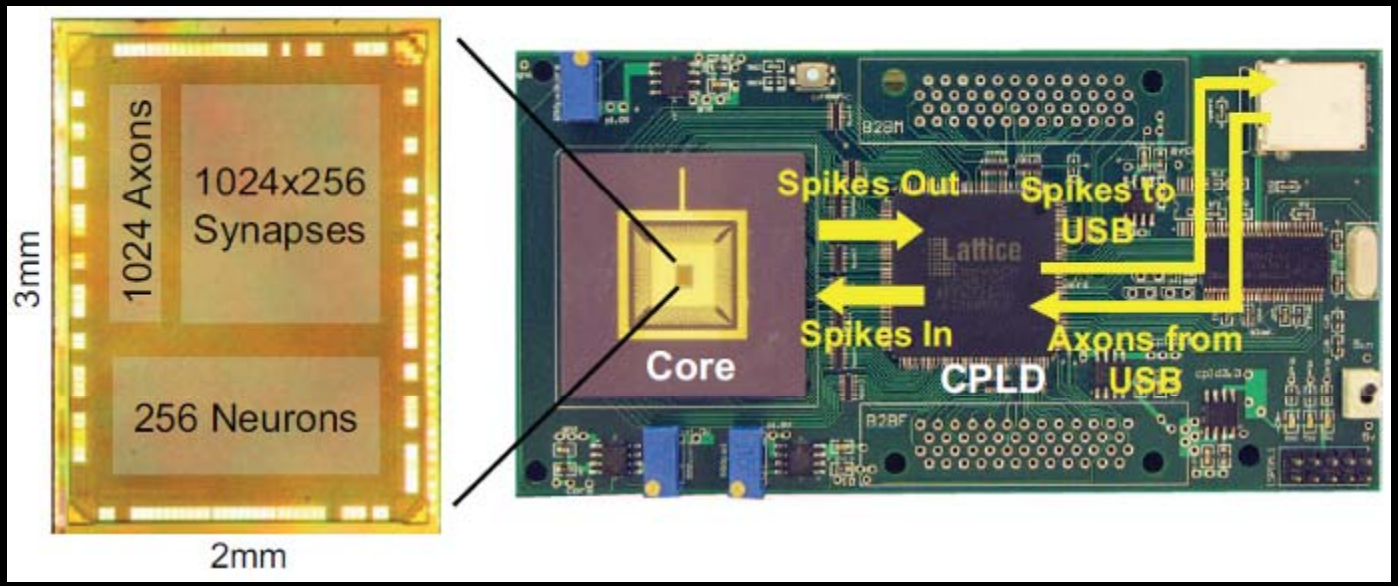


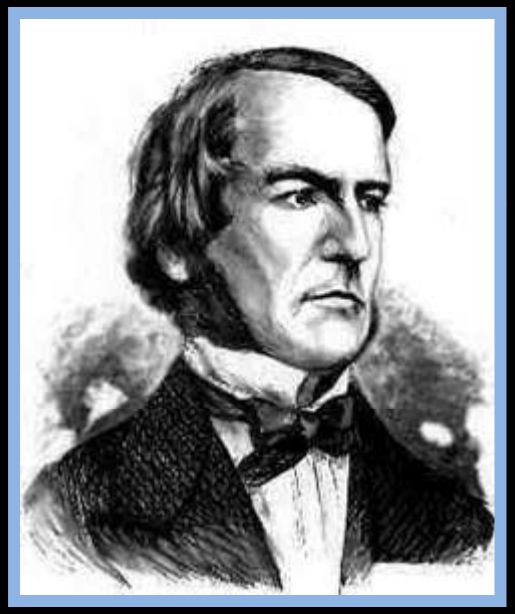
Learning Pong

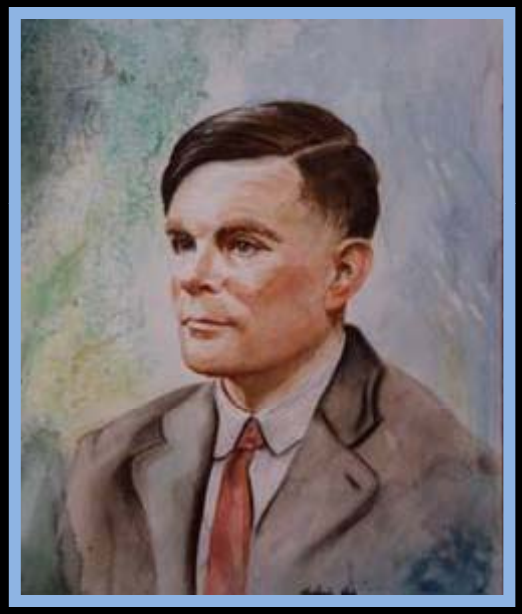


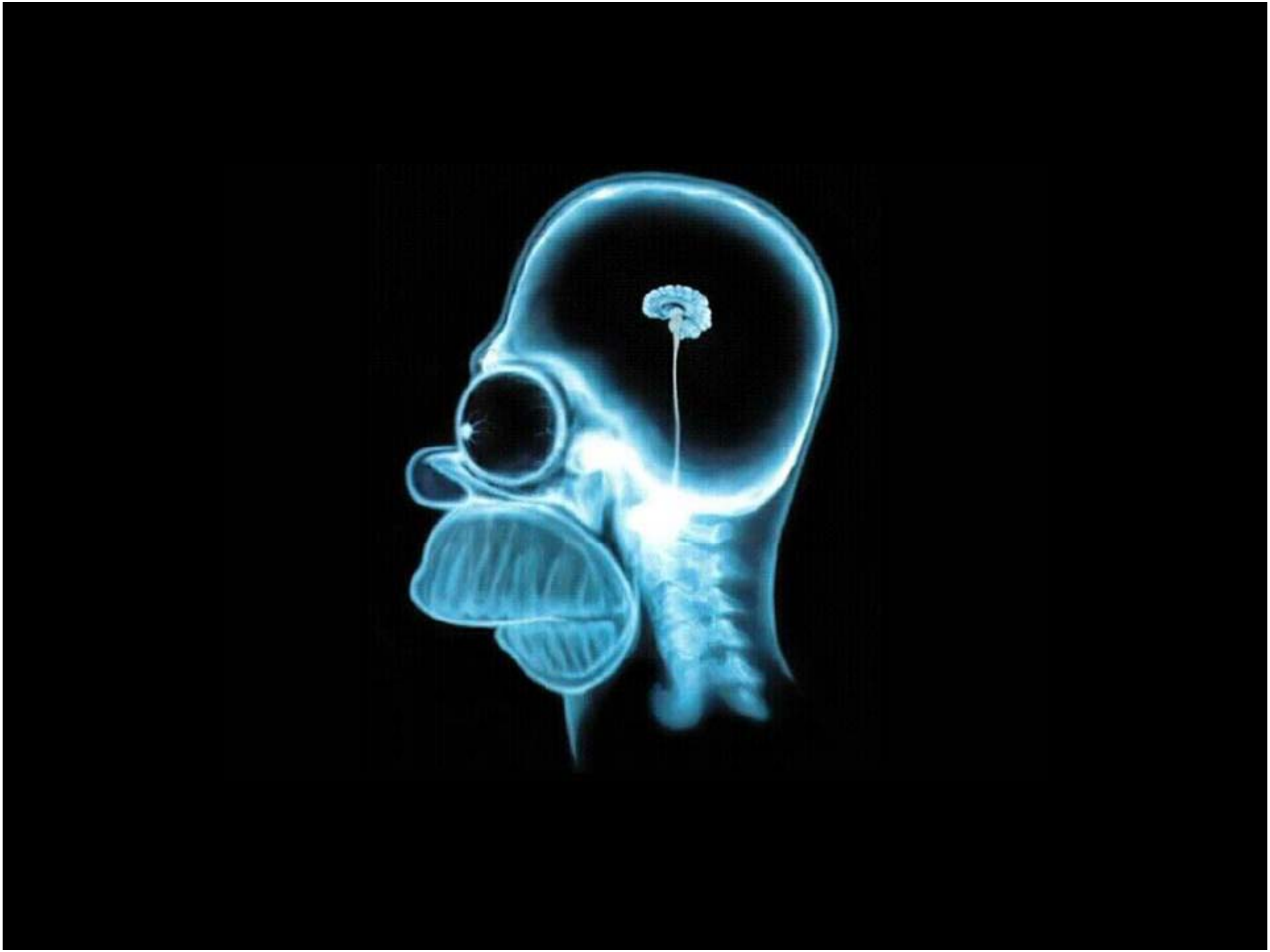
Character ID

Neurosynaptic Core

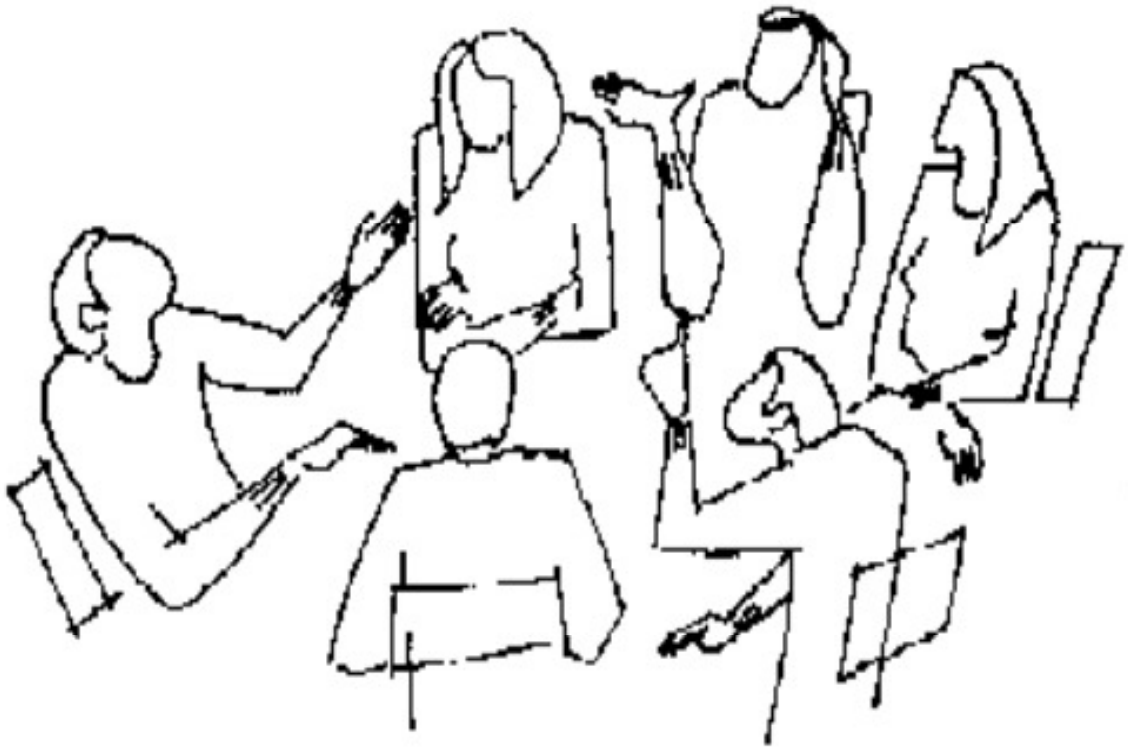








Discussion



Wrap-Up, Feedback & Next Steps