

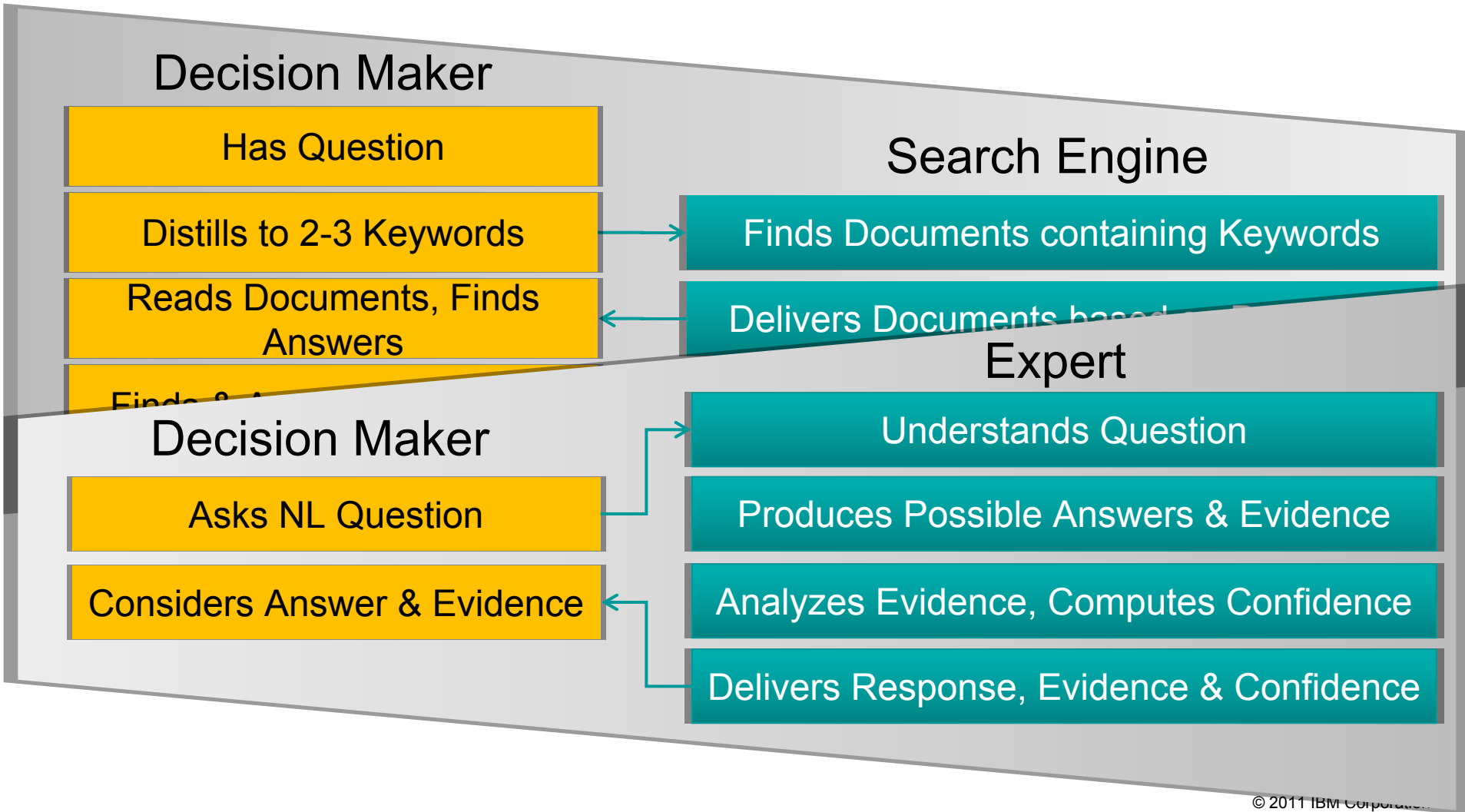
Building *Watson*

A Brief Overview of DeepQA and What Comes Next

Scott Spangler
IBM Research



Informed Decision Making: Search vs. Expert Q&A





Automatic Open-Domain Question Answering

A Long-Standing Challenge in Artificial Intelligence to emulate human expertise

- Given
 - Rich **Natural Language Questions**
 - Over a **Broad Domain of Knowledge**

- Deliver
 - **Precise Answers:** Determine what is being asked & give precise response
 - **Accurate Confidences:** Determine likelihood answer is correct
 - **Consumable Justifications:** Explain why the answer is right
 - **Fast Response Time:** Precision & Confidence in <3 seconds



A Grand Challenge Opportunity

- Capture the imagination
 - The Next *Deep Blue*
- Engage the scientific community
 - Envision new ways for computers to impact society & science
 - Drive important and measurable scientific advances
- Be Relevant to IBM Customers
 - Enable better, faster decision making over unstructured and structured content
 - Business Intelligence, Knowledge Discovery and Management, Government, Compliance, Publishing, Legal, Healthcare, Business Integrity, Customer Relationship Management, Web Self-Service, Product Support, etc.



Real Language is Real Hard

■ Chess

- A finite, mathematically well-defined search space
- Limited number of moves and states
- Grounded in **explicit, unambiguous** mathematical rules



■ Human Language

- Ambiguous, contextual and implicit
- Grounded only in **human cognition**
- Seemingly infinite number of ways to express the same meaning



What Computers Find Easier (and Hard)

$$(\ln(12,546,798 * \pi) ^ 2) / 34,567.46 = 0.00885$$

Select *Payment* where *Owner*="David Jones" and *Type(Product)*="Laptop",

Owner	Serial Number
David Jones	45322190-AK

Invoice #	Vendor	Payment
INV10895	MyBuy	\$104.56

Serial Number	Type	Invoice #
45322190-AK	LapTop	INV10895

David Jones
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 David Jones

=

Dave Jones
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 David Jones ≠

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.

The Jeopardy! Challenge

A palpable, compelling and notable way to drive the technology of Question Answering along Key Dimensions

Broad/Open Domain

Complex Language

High Precision

Accurate Confidence

High Speed

\$200

The juice of this bog fruit is sometimes used to treat urinary tract infections

\$400

A round type of this structure, with the silo inside it, was popular in the 1900s; it's rare now

\$1000

Of the 4 countries in the world that the U.S. does not have diplomatic relations with, the one that's farthest north

\$600

In cell division, mitosis splits the nucleus & cytokinesis splits this liquid *cushioning* the nucleus

\$800

Grace Murray Hopper is credited with applying this 3-letter term to a mysterious computer problem

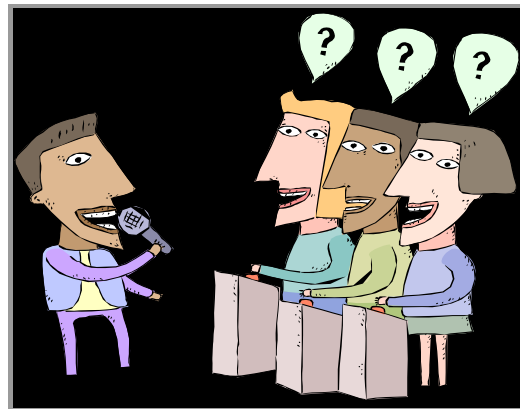
Basic Game Play

Technology	Clue	Before and After
\$200	TECHNOLOGY	\$200
\$400	ALL POLICEMEN CAN THANK STEPHANIE KWOLEK FOR HER INVENTION OF THIS POLYMER FIBER, 5 TIMES TOUGHER THAN STEEL	\$400
\$600		\$600
\$800		\$800
\$1000		\$1000

Annotations:

- 6 Categories (points to the 'Before and After' column)
- 5 Levels of Difficulty (points to the rows of monetary values)

- ❑ 1 of 3 Players Selects a Clue
- ❑ Host reads Clue out loud
- ❑ All Players compete to answer
- ❑ 1st to buzz-in gets to answer



- ❑ **IF correct**
 - earns \$ value
 - selects Next Clue
- ❑ **IF wrong**
 - **loses** \$ value
 - other players buzz again (rebounds)

- ❑ Two Rounds Per Game + Final Question
- ❑ ONE Daily Double in First Round, TWO in 2nd Round

Some Basic Jeopardy! Clues

- This **fish** was thought to be extinct millions of years ago until one was found off South Africa in 1938
 - Category: ENDS IN "TH"
 - Answer: **coelacanth**

- When hit by electrons, a phosphor gives off electromagnetic energy in this **form**
 - Category: General Science
 - Answer: **light (or photons)**

- Secy. Chase just submitted **this** to me for the third time--guess what, pal. This time I'm accepting **it**
 - Category: Lincoln Blogs
 - Answer: **his resignation**

The *type* of thing being asked for is often indicated but can go from specific to very vague

Automatic Learning for “Reading”

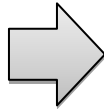
Sentence
Parsing

Generalization &
Statistical Aggregation

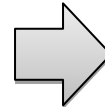
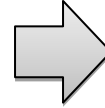
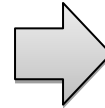
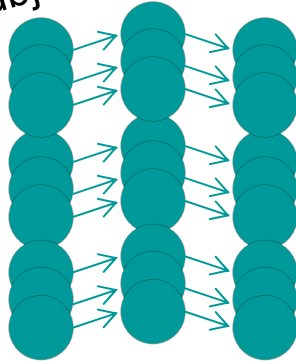
Volumes of Text

Syntactic Frames

Semantic Frames



subject verb object



Inventors patent inventions (.8)
 Officials Submit Resignations (.7)
 People earn degrees at schools (0.9)
 Fluid is a liquid (.6)
 Liquid is a fluid (.5)
 Vessels Sink (0.7)
 People sink 8-balls (0.5) (in pool/0.8)



Evaluating Possibilities and Their Evidence

In cell division, mitosis splits the nucleus & cytokinesis splits this **liquid** cushioning the nucleus.

- *Organelle*
- *Vacuole*
- *Cytoplasm*
- *Plasma*
- *Mitochondria*
- *Blood ...*

- Many candidate answers (CAs) are generated from many different searches
- Each possibility is evaluated according to **different dimensions of evidence**.
- **Just One** piece of evidence is if the CA is of the right type. In this case a "liquid".

Is("Cytoplasm", "liquid") = 0.2↑

Is("organelle", "liquid") = 0.1

Is("vacuole", "liquid") = 0.2

Is("plasma", "liquid") = 0.7

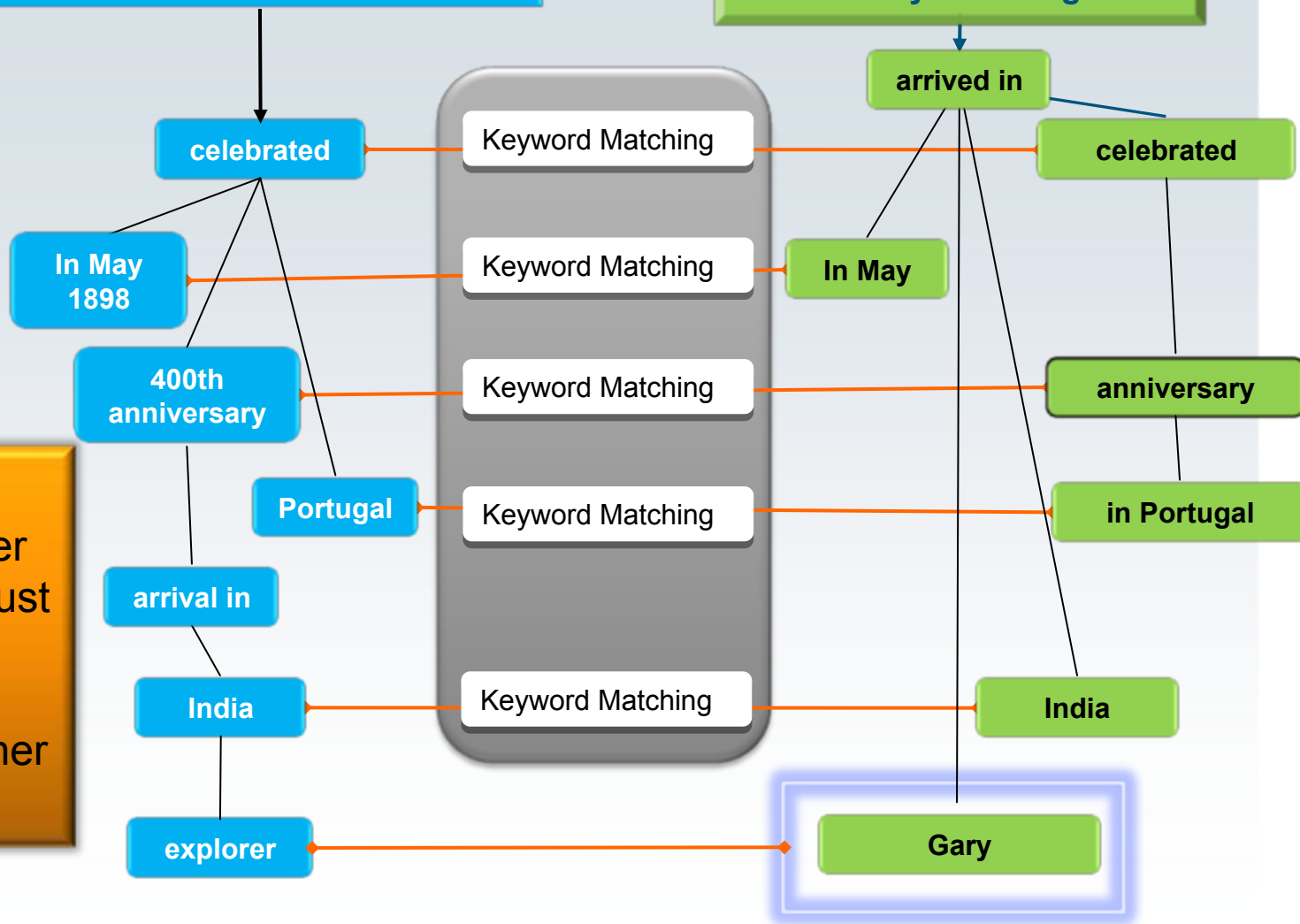
"Cytoplasm is a **fluid** surrounding the nucleus..."

Wordnet → Is_a(Fluid, Liquid) → ?

Learned → Is_a(Fluid, Liquid) → yes.

In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.

In May, Gary arrived in India after he celebrated his anniversary in Portugal.

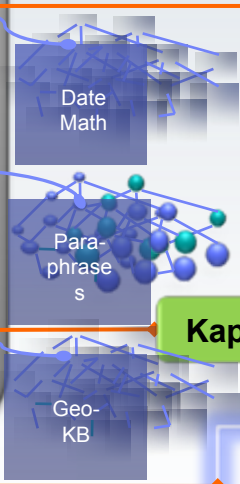
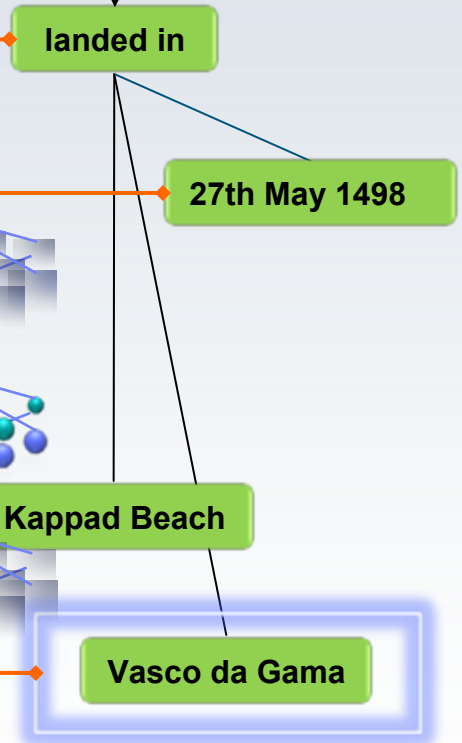
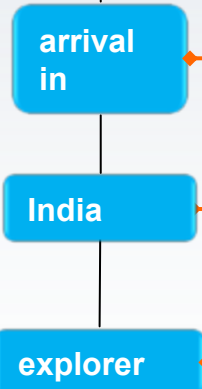
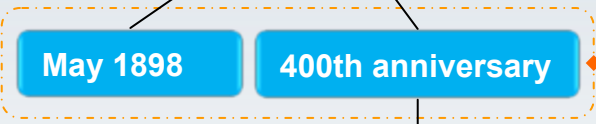
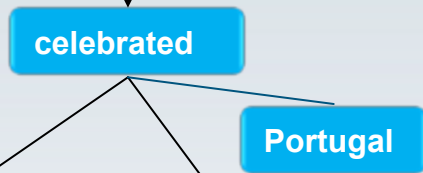
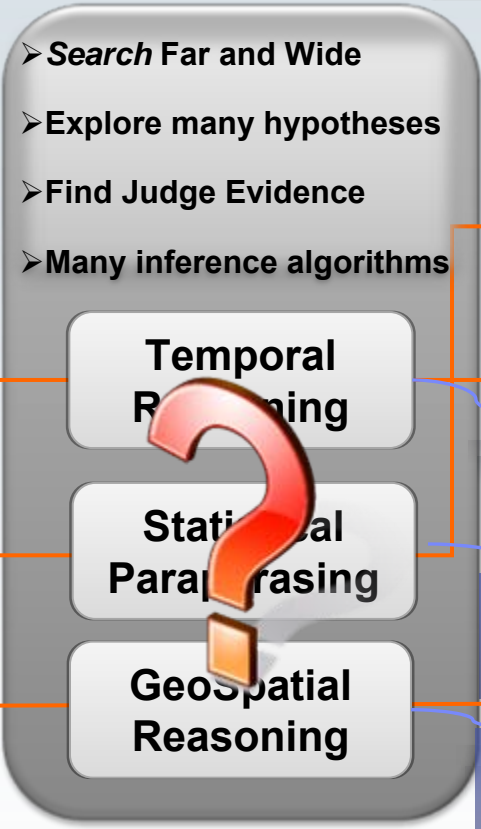


Evidence suggests "Gary" is the answer BUT the system must learn that keyword matching may be weak relative to other types of evidence

Different Types of Evidence: Deeper Evidence

In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.

On the 27th of May 1498, Vasco da Gama landed in Kappad Beach



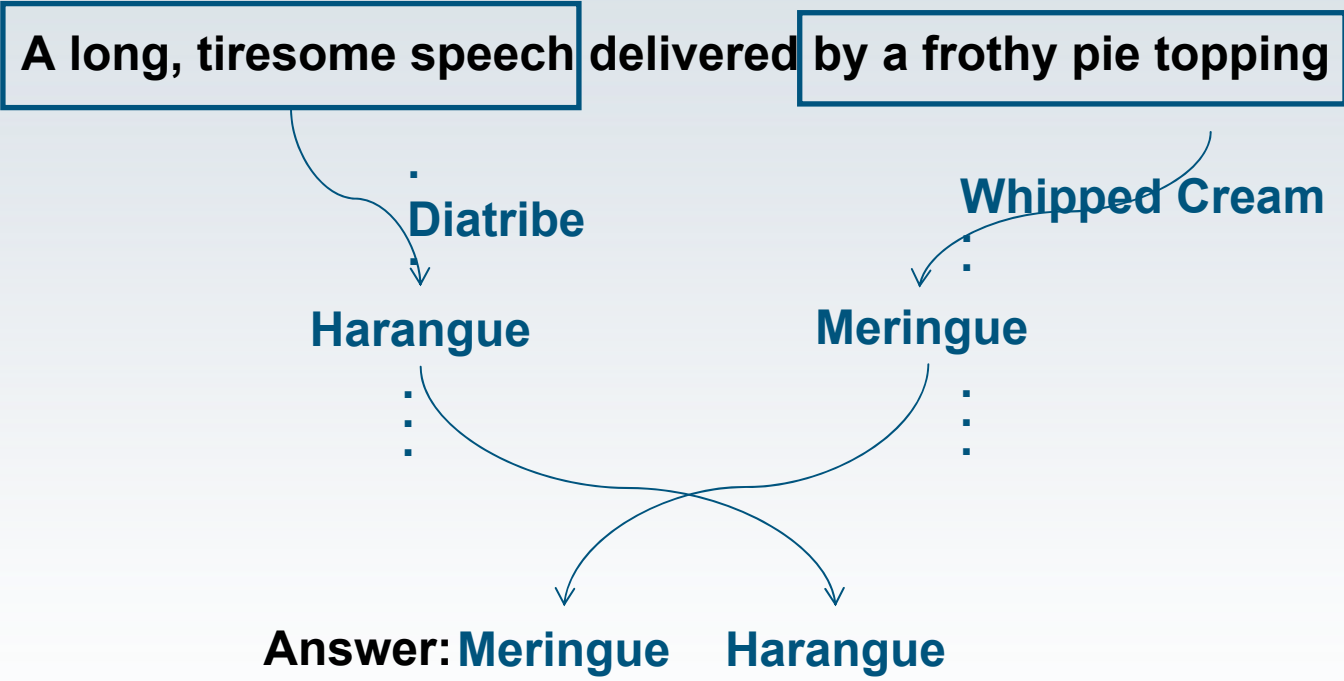
Stronger evidence can be much harder to find and score.

The evidence is still not 100% certain.

Not *Just* for Fun

Some Questions require
Decomposition and Synthesis

Category: Edible Rhyme Time



Divide and Conquer (Typical in Final Jeopardy!)

Must identify and solve sub-questions from different sources to answer the top level question

Lyndon B Johnson

In 1968

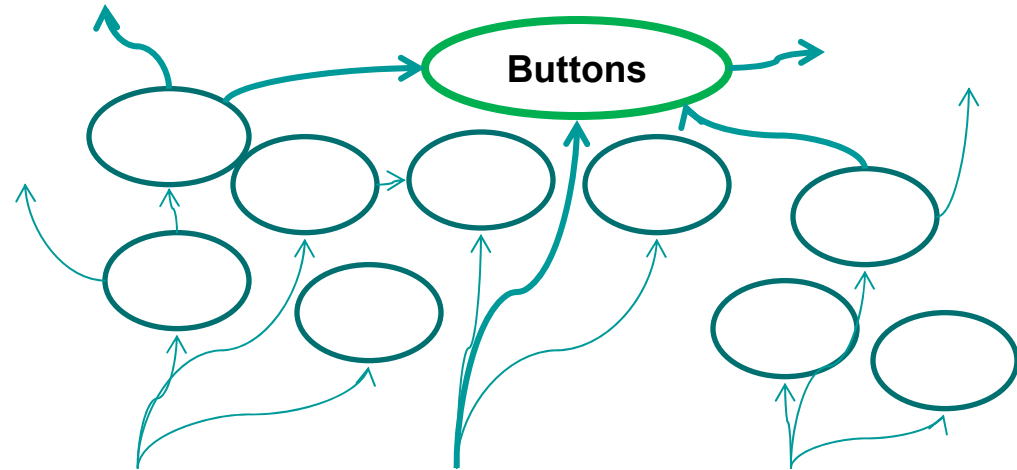
this man was U.S. president.

When "60 Minutes" premiered this man was U.S. president.

?

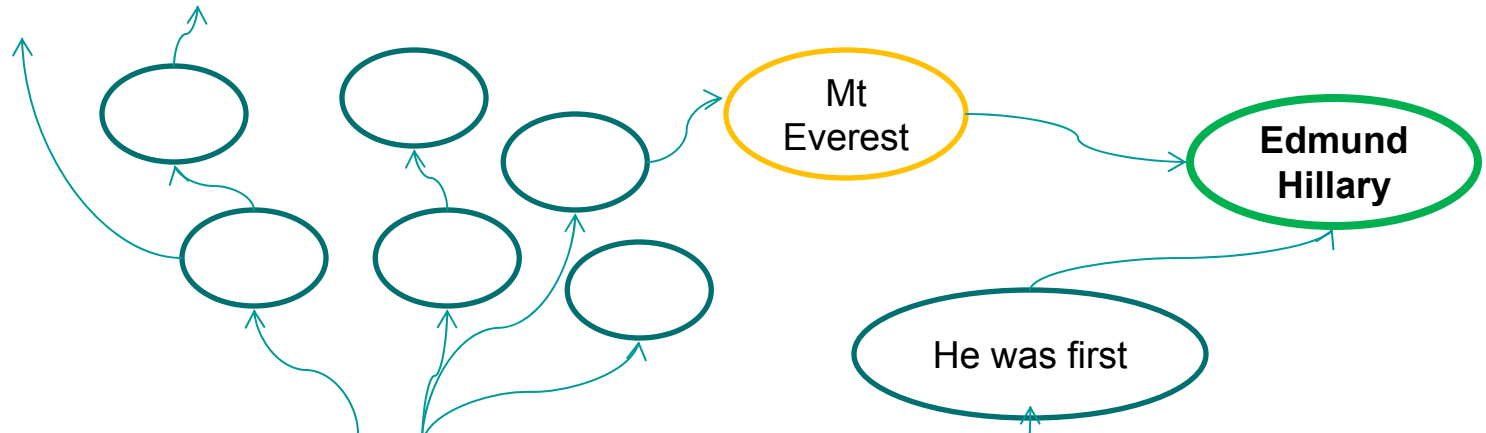
The DeepQA architecture attempts different *decompositions* and recursively applies the QA algorithms

Missing Links



Category: Common Bonds

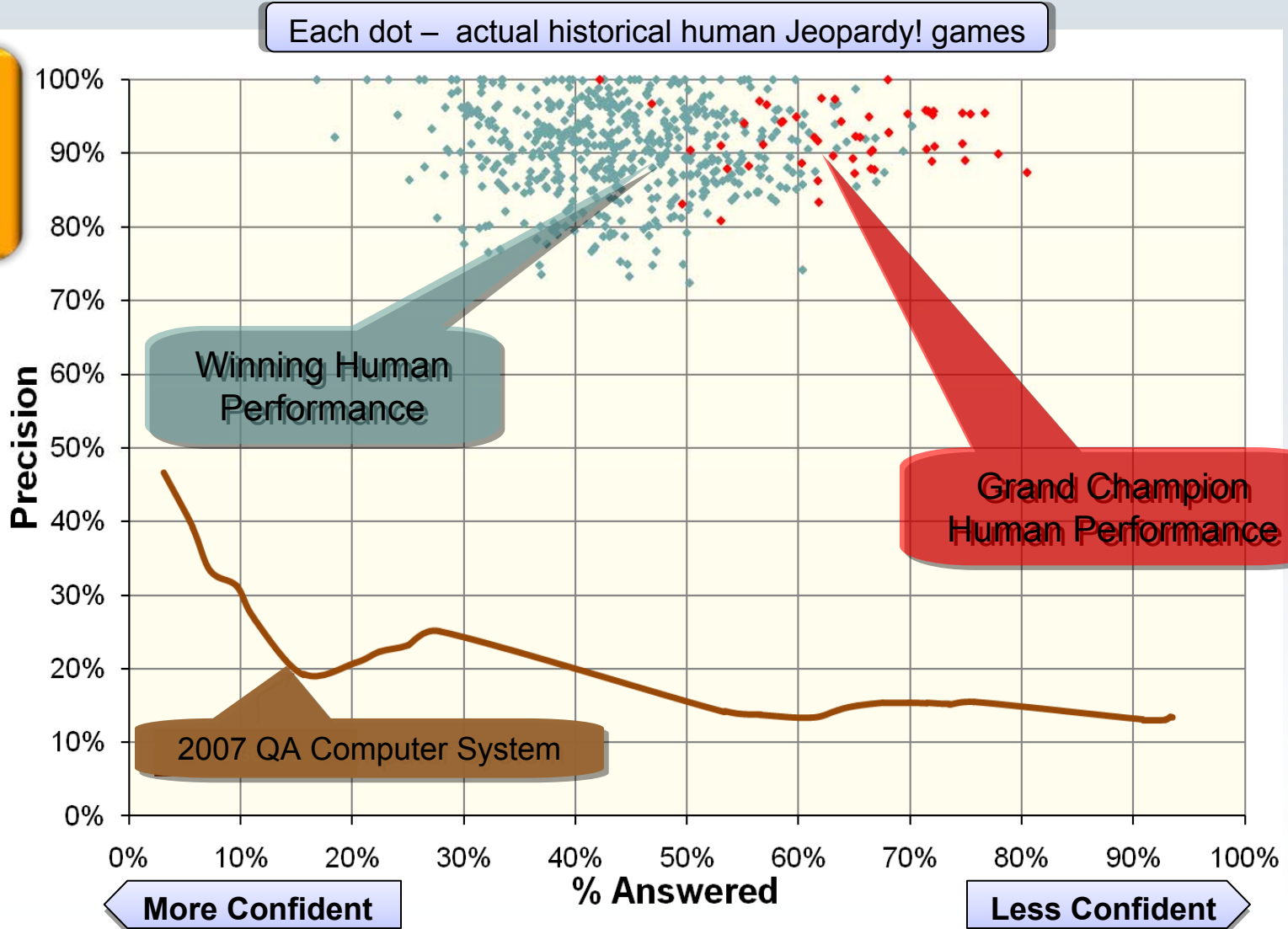
Shirts, TV remote controls, Telephones



On hearing of the discovery of George Mallory's body, he told reporters he still thinks he was first.

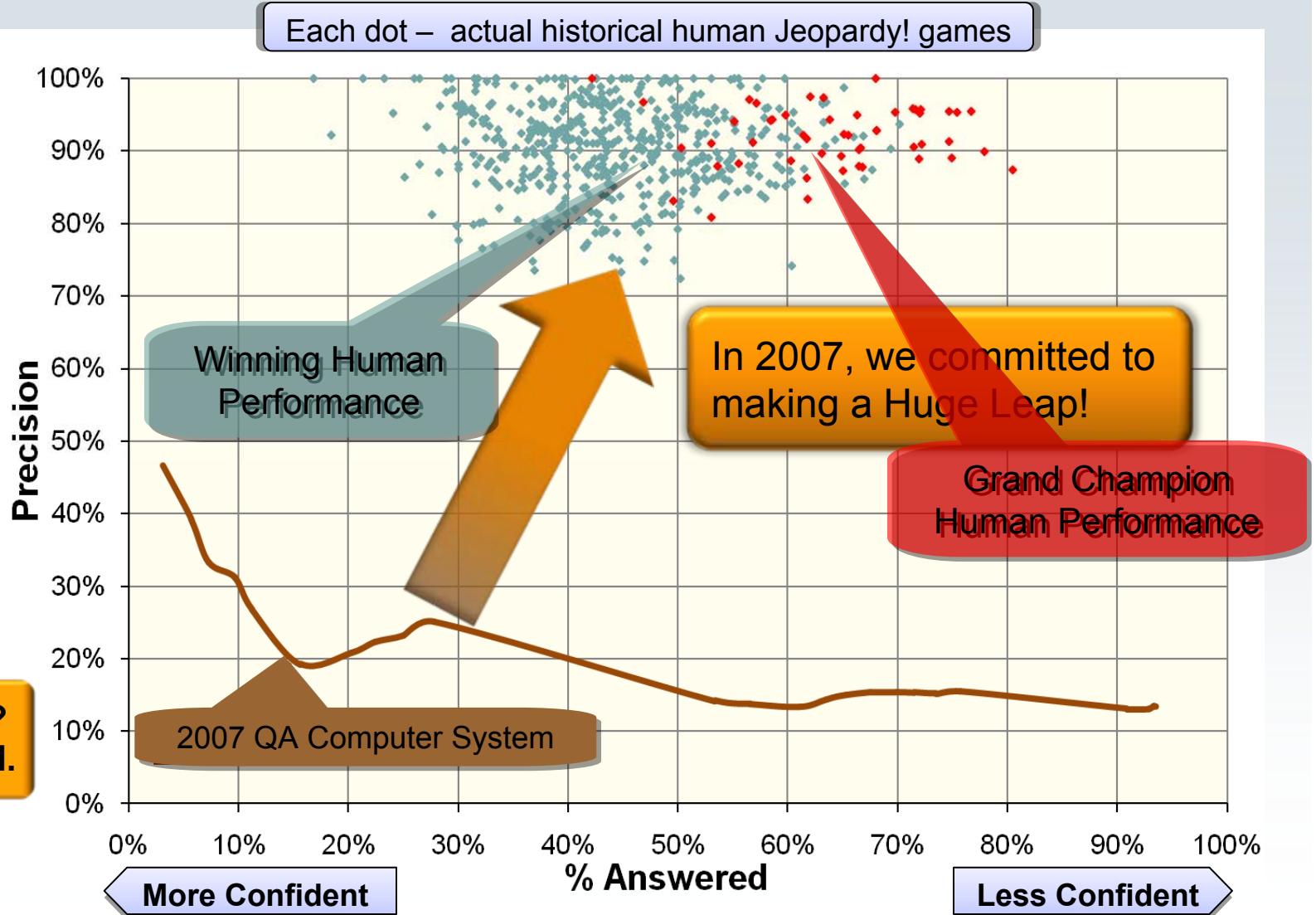
What It Takes to compete against Top Human Jeopardy! Players

Our Analysis Reveals the Winner's Cloud



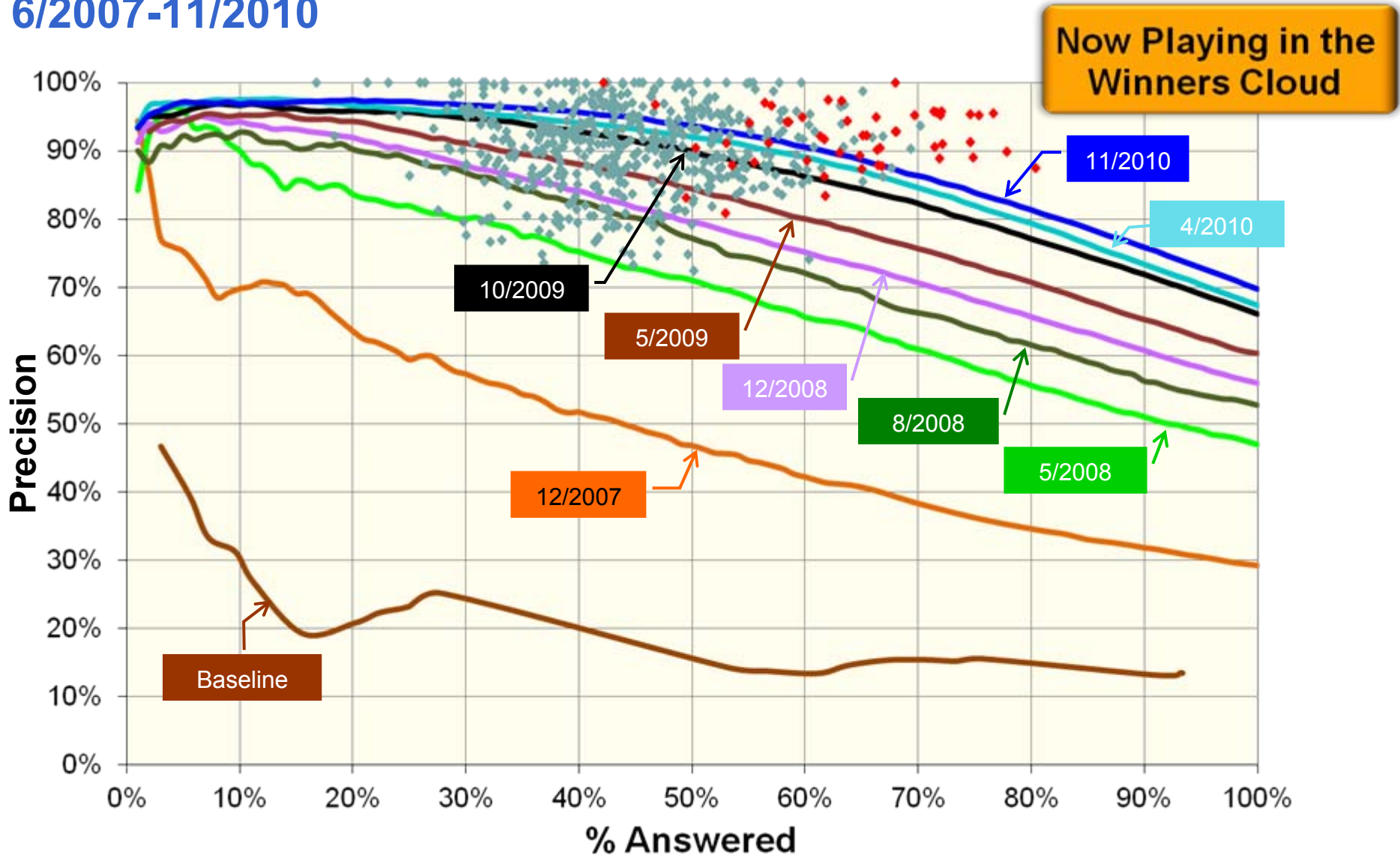
What It Takes to compete against Top Human Jeopardy! Players

Our Analysis Reveals the Winner's Cloud



Deep QA: Incremental Progress in Precision and Confidence

6/2007-11/2010



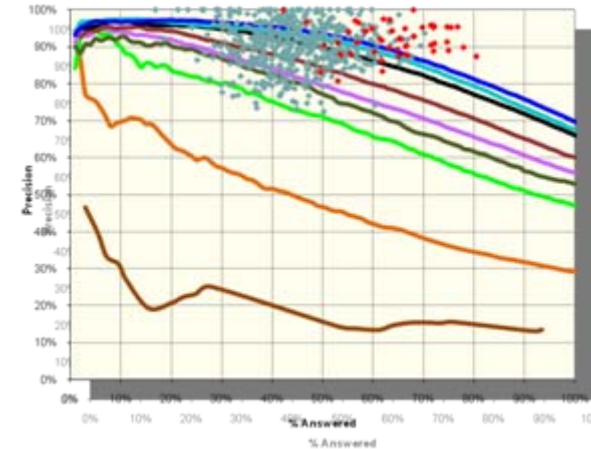
Watson: Precision, Confidence & Speed

- **Deep Analytics** – We achieved champion-levels of *Precision* and *Confidence* over a huge variety of expression



- **Speed** – By optimizing Watson’s computation for Jeopardy! on **2,880 POWER7** processing cores we went **from 2 hours per question on a single CPU to an average of just 3 seconds** – fast enough to compete with the best.

- **Results** – in 55 real-time sparring against former **Tournament of Champion Players last year**, Watson put on a very competitive performance, winning 71%. In the final Exhibition Match against Ken Jennings and Brad Rutter, Watson won!



Potential Business Applications



Healthcare / Life Sciences: Diagnostic Assistance, Evidenced-Based, Collaborative Medicine

Tech Support: Help-desk, Contact Centers



Enterprise Knowledge Management and Business Intelligence

Government: Improved Information Sharing and Security



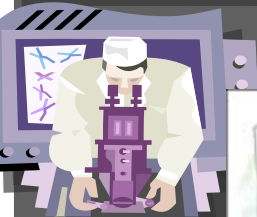
Considers and synthesizes a broad range of evidence improving quality, reducing cost



Symptoms



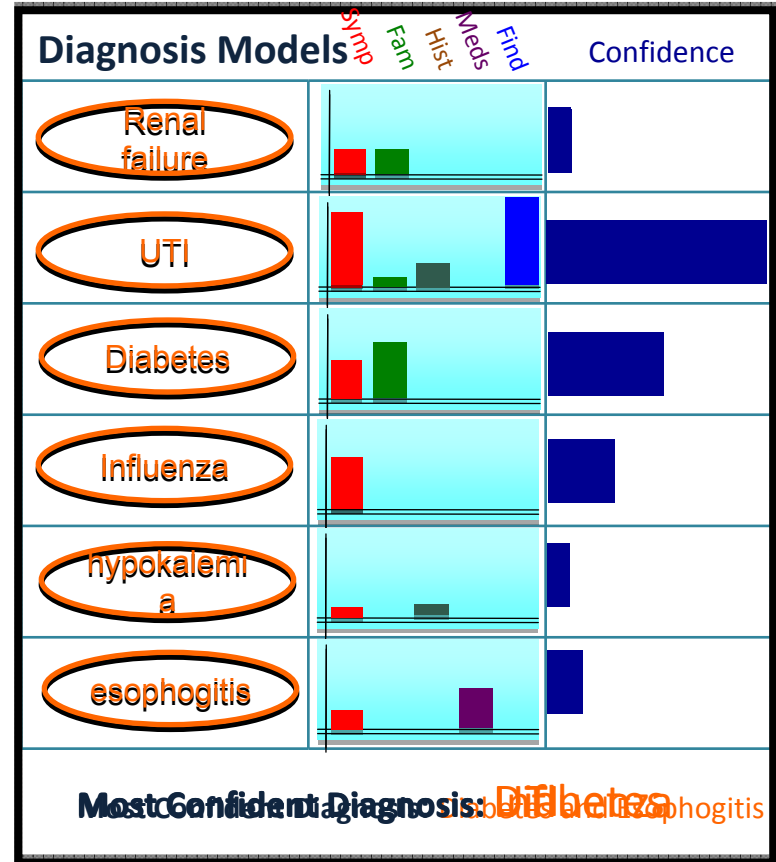
Family History
Patient History
Medications
Tests/Findings



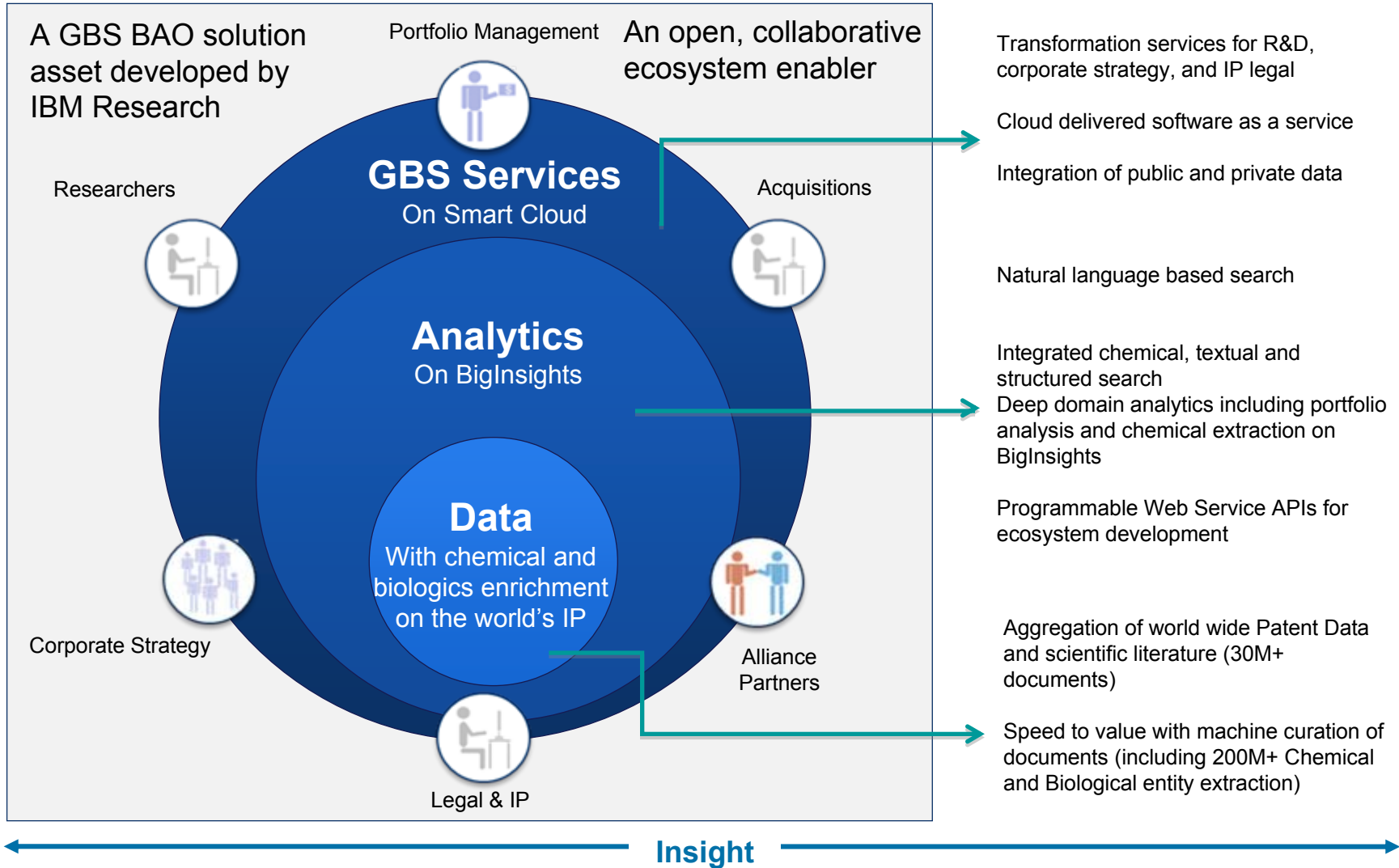
Notes/Hypotheses



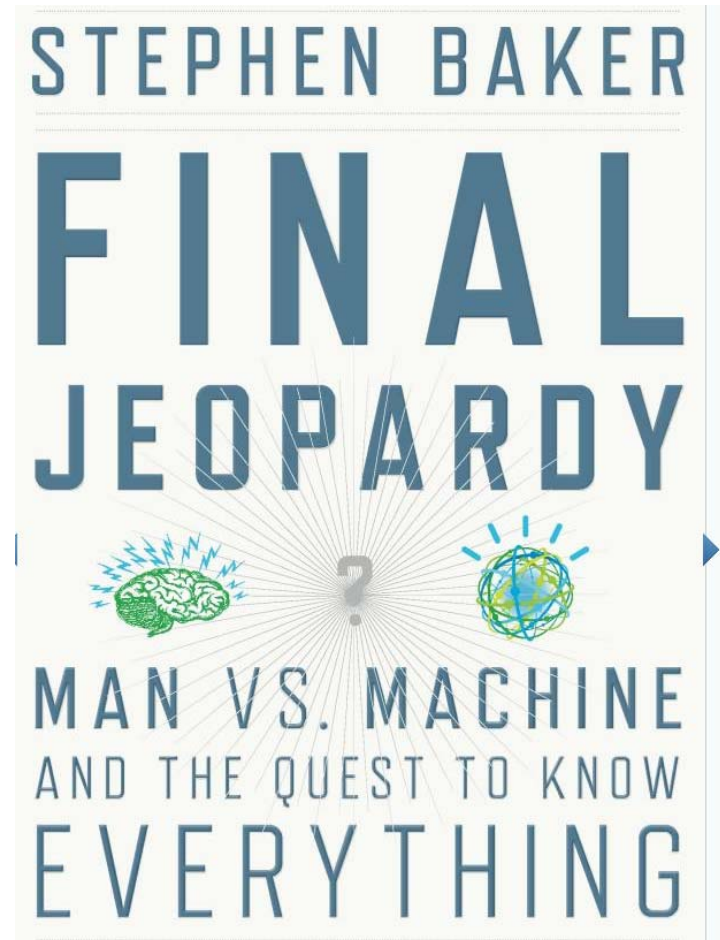
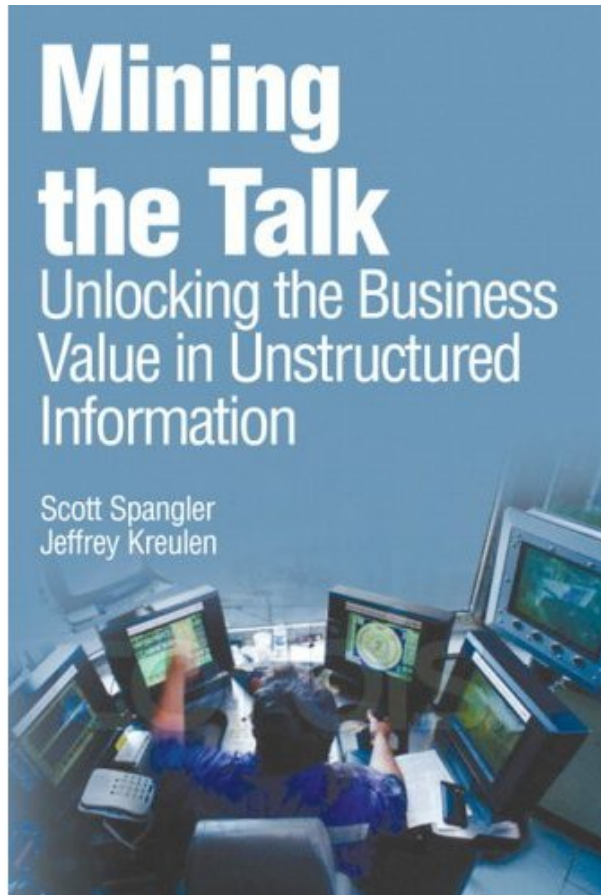
Huge Volumes of Texts, Journals, References, DBs etc.



IBM SIIP - Deriving strategic insights and driving business transformations by combining deep text analytics with the world's IP on IBM Smart Cloud



Further Reading



IBM Almaden Research Center

