



What is Watson?

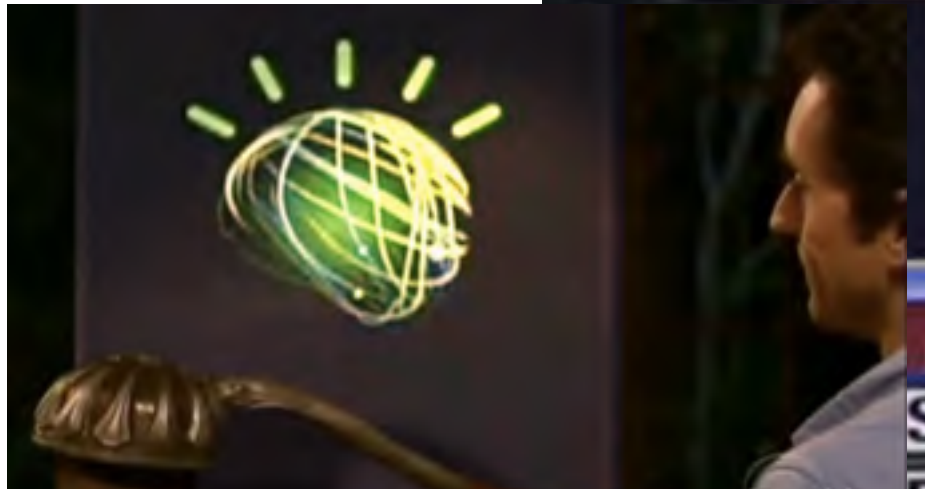


David Gondek, Ph.D.
Knowledge Capture and Learning
Watson Healthcare Adaptation
IBM Research

Watson



The Future of Watson?



5

Which of these kitchen items most closely resembles a hemisphere?

- A: Salad fork
- B: Cookie sheet
- C: Spatula
- D: Cereal bowl

SlateTV.com



Informed Decision Making: Search vs. Expert Q&A

Decision Maker

Has Question

Distills to 2.7 Keywords

Reads Documents, Finds Answers

Finds 8.4 Documents

Search Engine

Finds Documents containing Keywords

Delivers Documents based on Relevance

Expert

Understands Question

Produces Possible Answers & Evidence

Analyzes Evidence, Computes Confidence

Delivers Response, Evidence & Confidence

Decision Maker

Asks NL Question

Considers Answer & Evidence



The Jeopardy! Challenge: *A compelling and notable way to drive and measure the technology of automatic Question Answering along 5 Key Dimensions*

**Broad/Open
Domain**

**Complex
Language**

**High
Precision**

**Accurate
Confidence**

**High
Speed**

\$200

Keanu Reeves had a Nokia phone, but it took a land line to slip in & out of this, the title of a 1999 sci-fi flick

\$1000

1948: Johns Hopkins scientists find that this antihistamine alleviates motion sickness

\$600

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

\$2000

The first person mentioned by name in 'The Man in the Iron Mask' is this hero of a previous book by the same author.

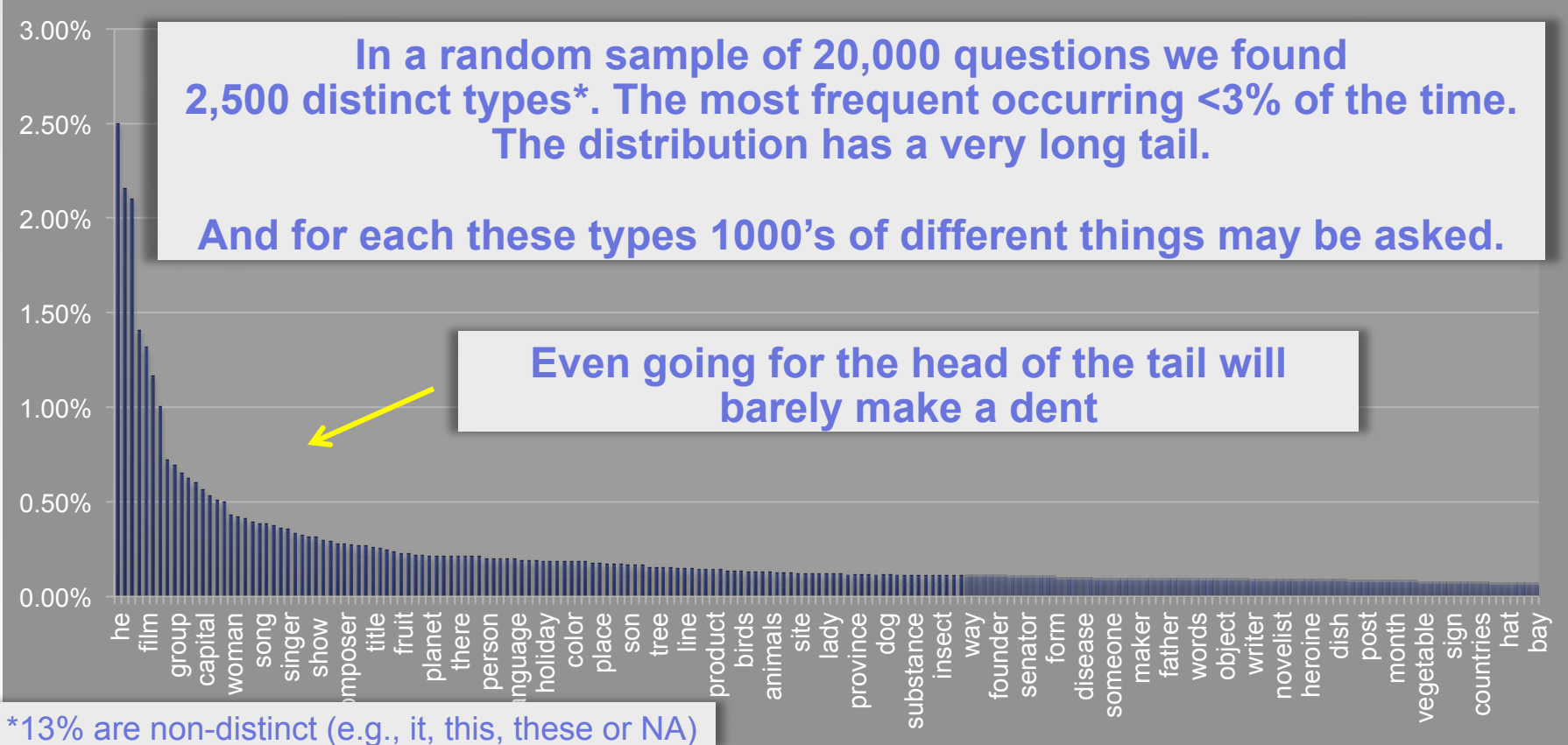
Broad Domain

We do NOT attempt to anticipate all questions and build specialized databases.

In a random sample of 20,000 questions we found 2,500 distinct types*. The most frequent occurring <3% of the time. The distribution has a very long tail.

And for each these types 1000's of different things may be asked.

Even going for the head of the tail will barely make a dent



*13% are non-distinct (e.g., it, this, these or NA)

Our Focus is on reusable NLP technology for analyzing volumes of *as-is* text. Structured sources (DBs and KBs) are used to help interpret the text.

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

From Chess to Chatting

■ Chess

- A finite, mathematically well-defined search space
- Limited number of moves and states
- All the symbols are completely grounded in the mathematical rules of the game

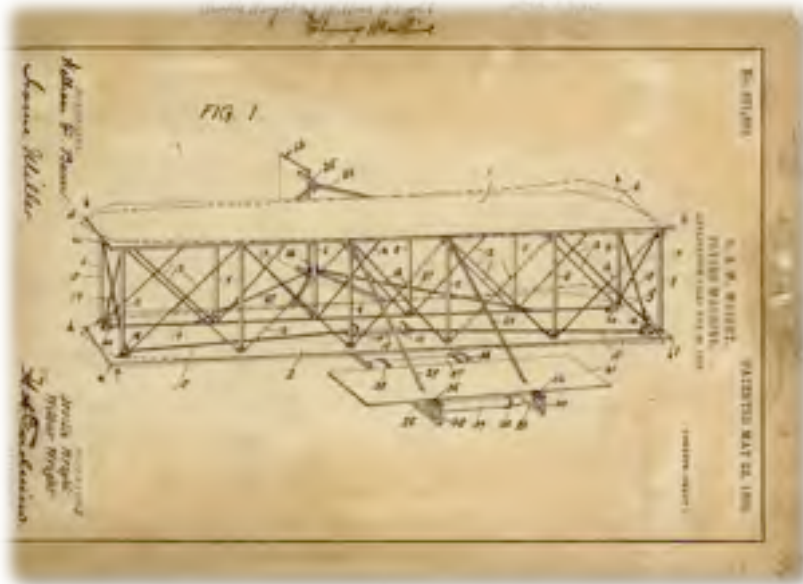


■ Human Language

- Words by themselves have no meaning
- Only grounded in **human cognition**
- Words navigate, align and communicate an infinite space of intended meaning
- Computers can **not** ground words to human experiences to derive meaning



Distance between question and justifying passage



The Wright brothers' first flight was this long.

The Wright brother's first flight was 120 feet long

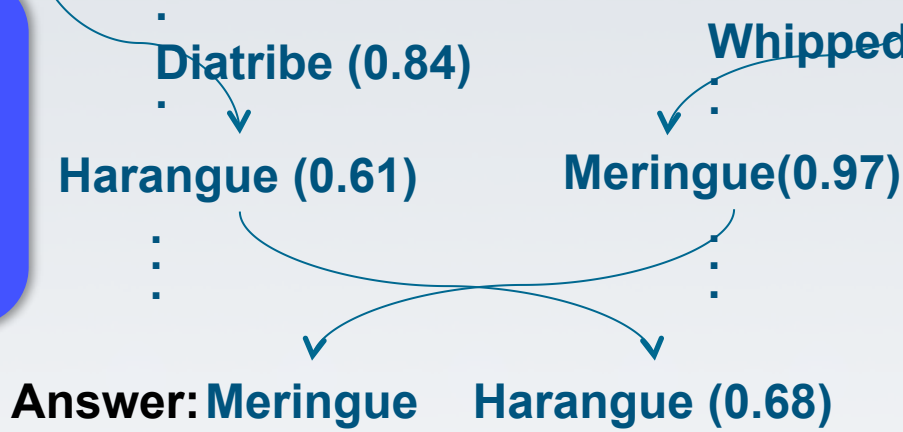
With Orville at the controls and Wilbur running along side to steady the wing, the plane rose 12 feet into the air and went about 120 feet on its 12-second trip. This marked the beginning of air travel for mankind.

Decomposition and Synthesis

A long, tiresome speech delivered by a frothy pie topping

EDIBLE RHYME
TIME: A long tiresome speech delivered by a frothy pie topping

BOXING TERMS:
Rhyming term for a hit below the belt



Correct: Low Blow

Watson: Wang Bang (0.43)

Some Questions require Decomposition and Synthesis

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

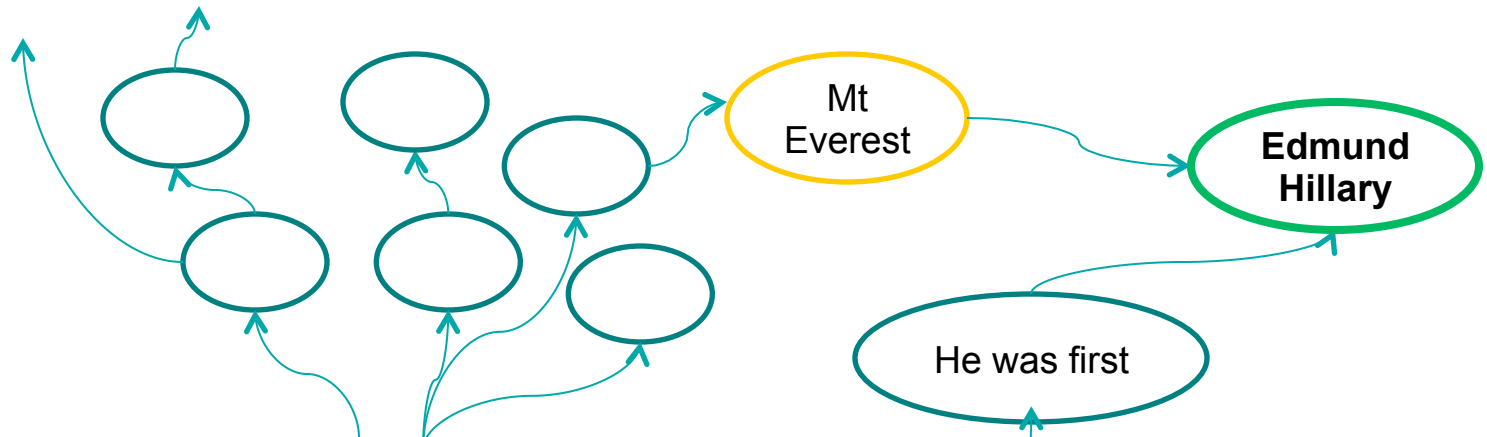
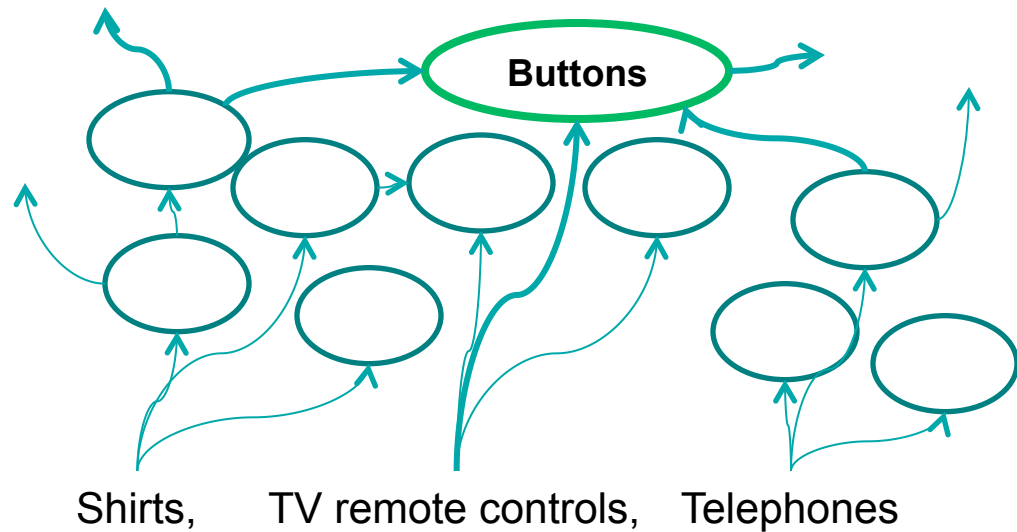
this man was U.S. president.

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

?

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

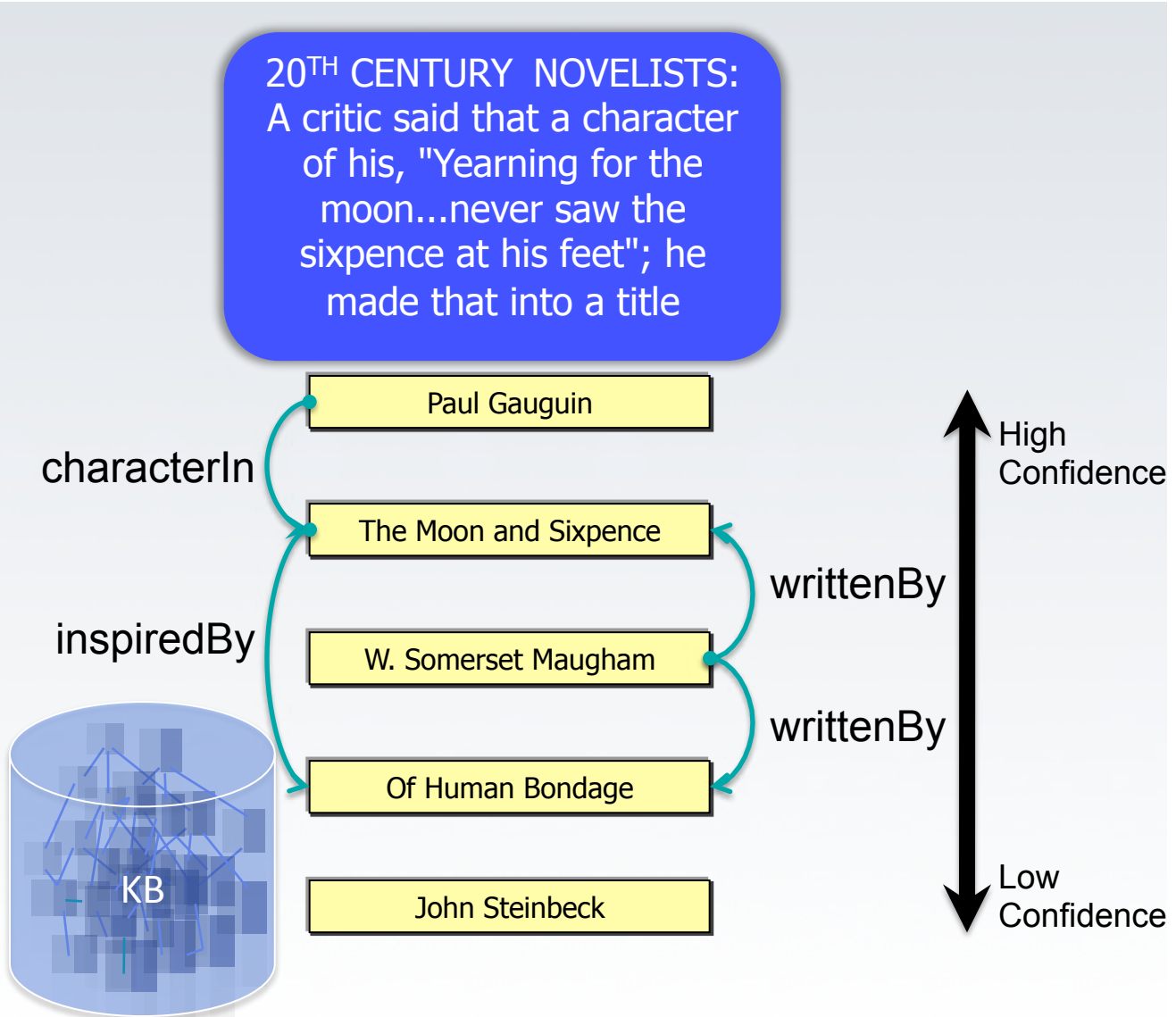
The Missing Link



Evidence Diffusion

Watson's Candidate Answers are not independent.

By sharing evidence based on the relationships between candidate answers we raise the score of the right answer



Category Inference

- What the Jeopardy! Clue is asking for is NOT always obvious
- Watson can try to infer the type of thing being asked for from the previous answers.
- In this example after seeing 2 correct answers Watson starts to dynamically learn a confidence that the question is asking for something that it can classify as a “month”.

CELEBRATIONS OF THE MONTH

Clue	Type	Watson's Answer	Correct Answer
D-DAY ANNIVERSARY & MAGNA CARTA DAY	day	Runnymede	June
NATIONAL PHILANTHROPY DAY & ALL SOULS' DAY	day	Day of the Dead	November
NATIONAL TEACHER DAY & KENTUCKY DERBY DAY	day / month	Churchill Downs	May
ADMINISTRATIVE PROFESSIONALS DAY & NATIONAL CPAS GOOF-OFF DAY	month	April	April
NATIONAL MAGIC DAY & NEVADA ADMISSION DAY	month	October	October

Evidence Dimensions



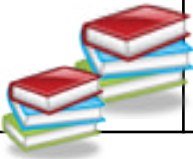


\$200
Keanu Reeves had a Nokia phone, but it took a land line to slip in & out of this, the title of a 1999 sci-fi flick

Evidence Dimensions

Keanu Reeves	had a Nokia Phone	took a land line to slip in & out of this	1999	Sci-fi flick
--------------	-------------------	---	------	--------------

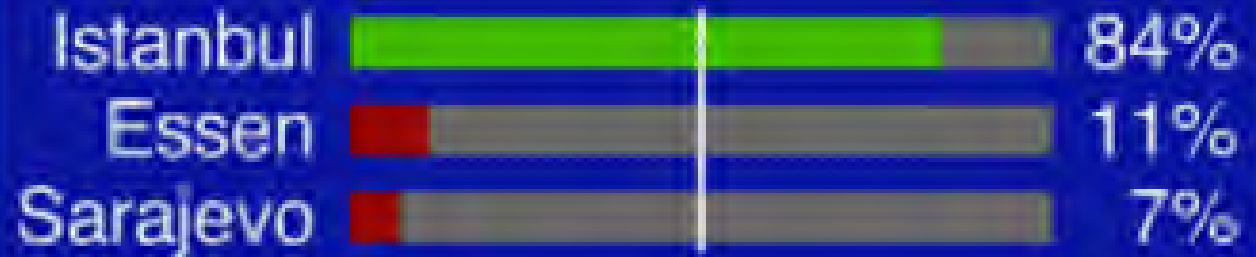
Justification

Evidence Dimensions

Evidence Sources KB: ActedIn(Keanu Reeves, X) 	Text Passage: has(Keanu Reeves, Nokia) 	Text Passage: Align 	KB: OccurredIn(X, 1999) 	KB: Isa(X,sci-fi flick) 
--	--	--	---	---

**EACH YEAR THE EU
SELECTS CAPITALS OF
CULTURE; ONE OF THE 2010
CITIES WAS THIS TURKISH
"MEETING PLACE OF
CULTURES"**

WHAT IS ISTANBUL? ✓

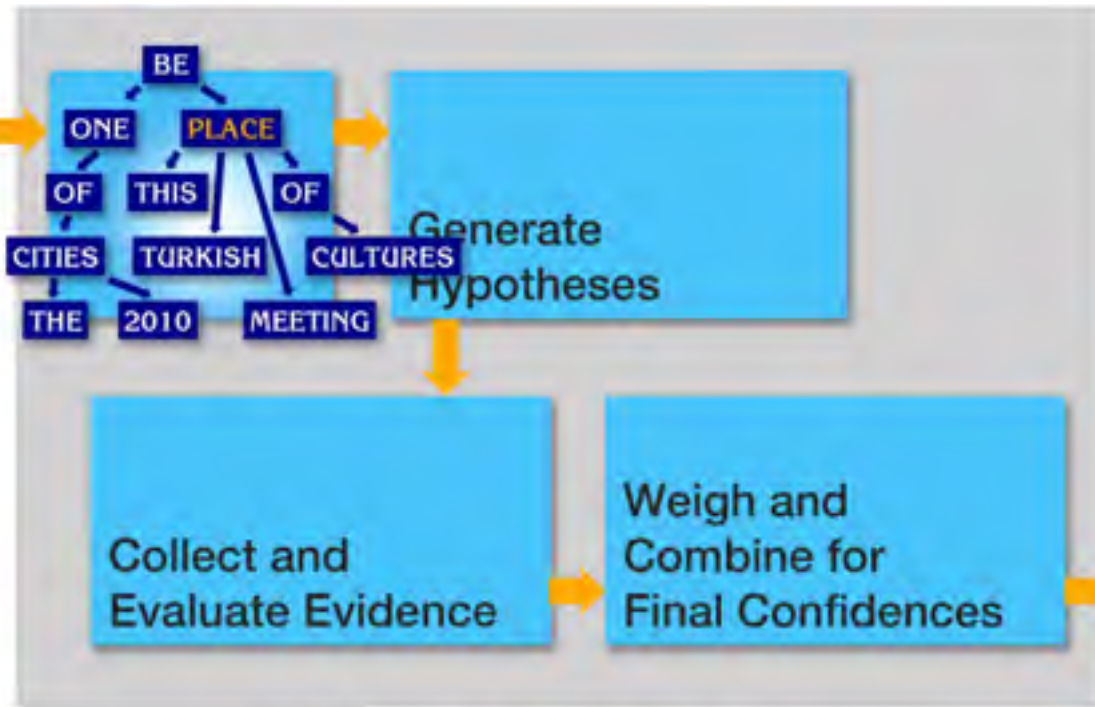




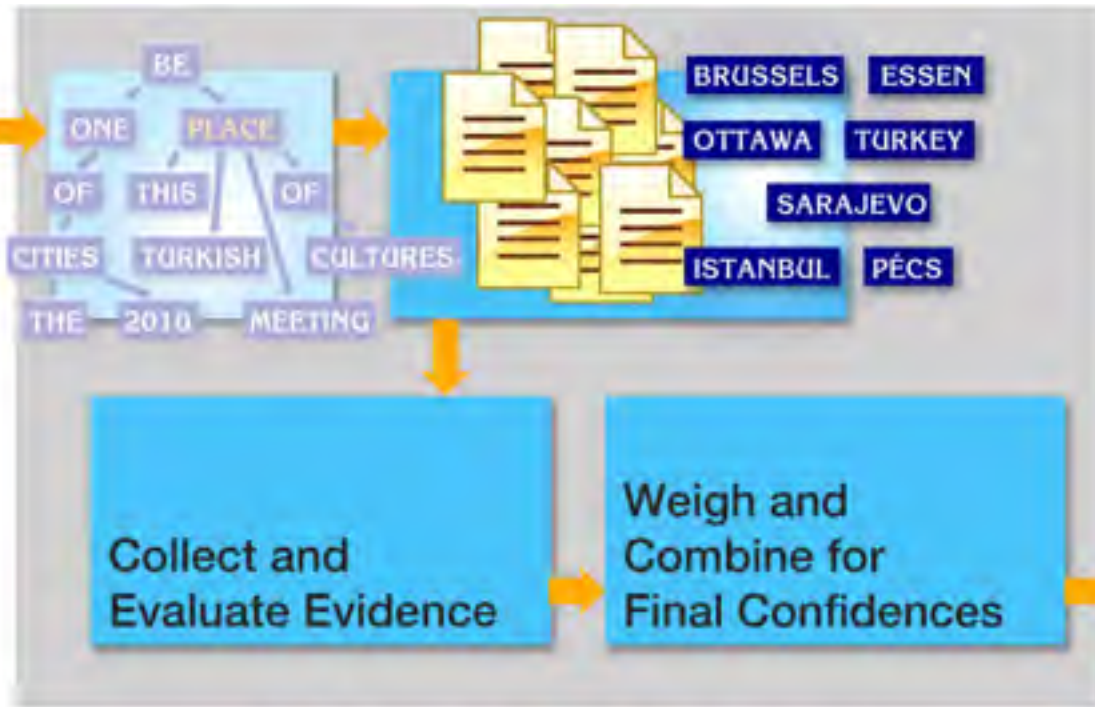
EACH YEAR THE EU
SELECTS
CAPITALS OF CUL-
TURE; ONE OF THE
2010 CITIES WAS
THIS TURKISH
"MEETING PLACE
OF CULTURES"



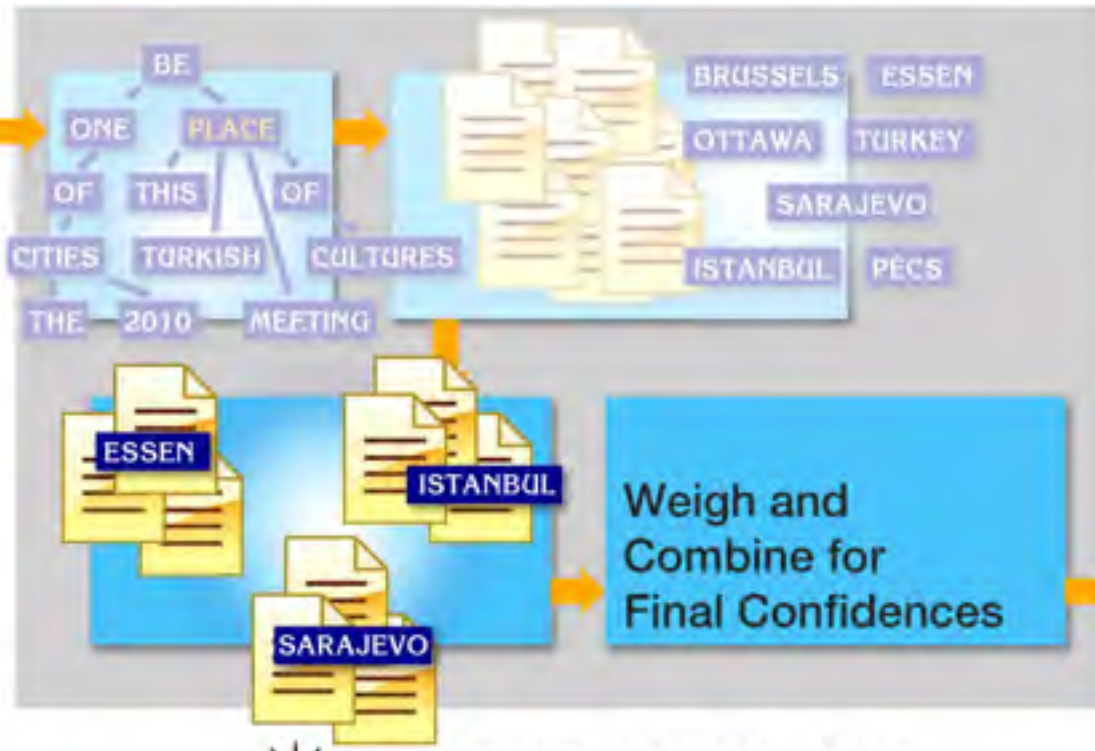
EACH YEAR THE EU
SELECTS
CAPITALS OF CULTURE; ONE OF THE
2010 CITIES WAS
THIS TURKISH
"MEETING PLACE
OF CULTURES"



EACH YEAR THE EU
SELECTS
CAPITALS OF CULTURE;
ONE OF THE
2010 CITIES WAS
THIS TURKISH
"MEETING PLACE
OF CULTURES"



EACH YEAR THE EU
SELECTS
CAPITALS OF CULTURE; ONE OF THE
2010 CITIES WAS
THIS TURKISH
"MEETING PLACE
OF CULTURES"

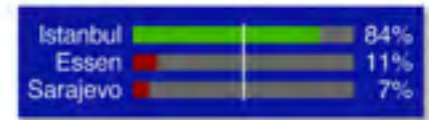


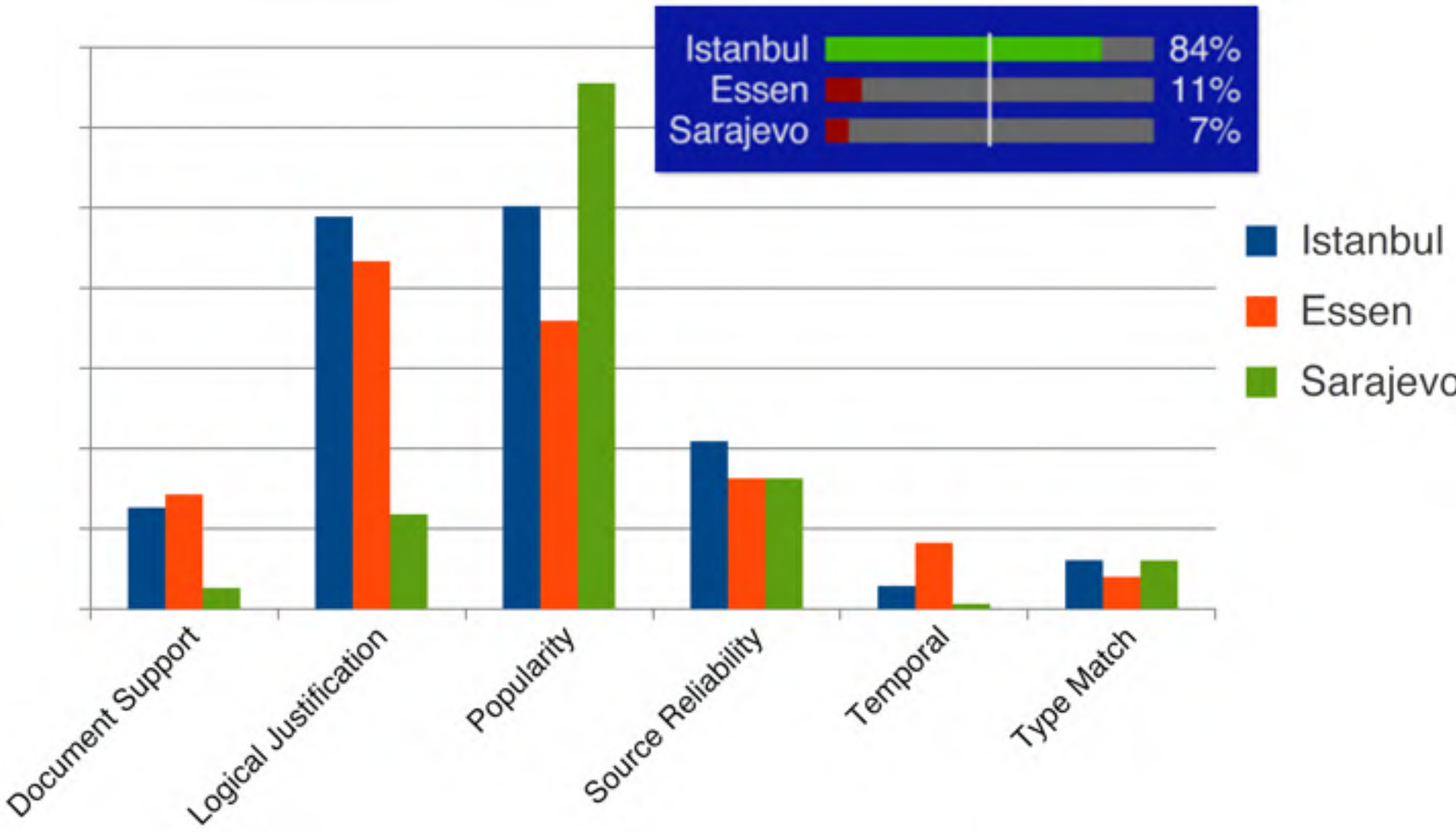
EACH YEAR THE EU
SELECTS
CAPITALS OF CUL-
TURE; ONE OF THE
2010 CITIES WAS
THIS TURKISH
"MEETING PLACE
OF CULTURES"



ANSWER

EACH YEAR THE EU
SELECTS
CAPITALS OF CULTURE;
ONE OF THE
2010 CITIES WAS
THIS TURKISH
"MEETING PLACE
OF CULTURES"

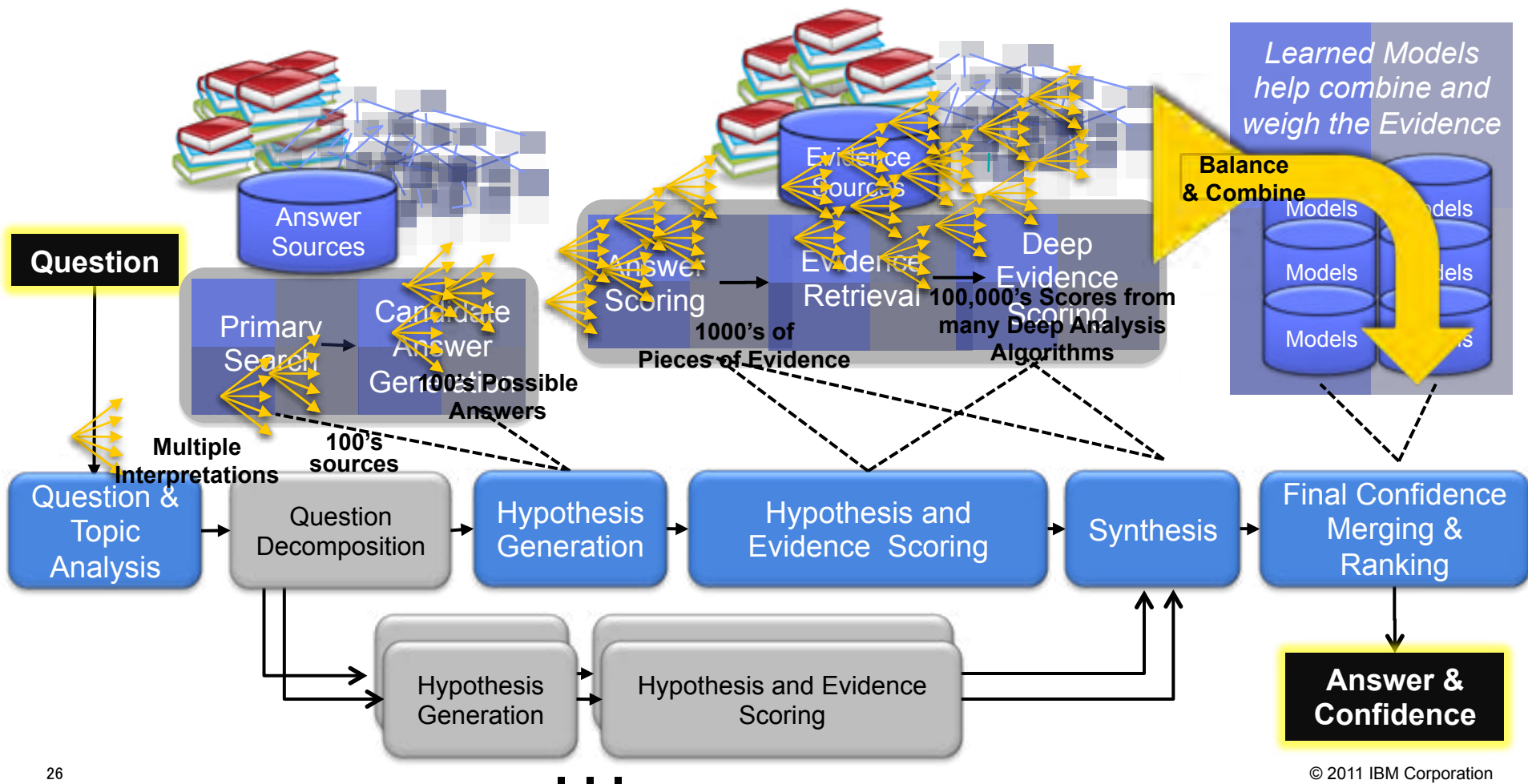




Massively Parallel Probabilistic Evidence-Based Architecture

Generates and scores many hypotheses using a combination of 1000's **Natural Language Processing, Information Retrieval, Machine Learning and Reasoning Algorithms.**

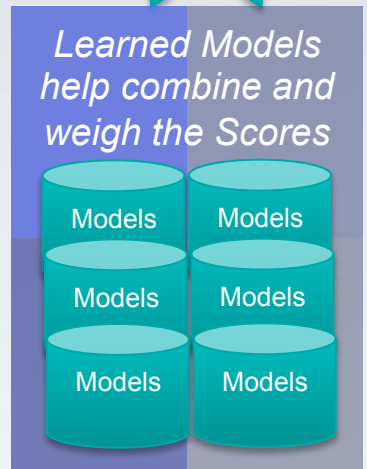
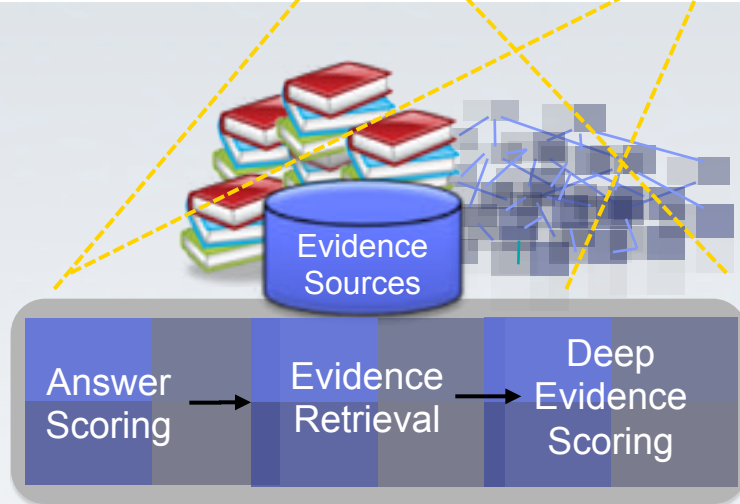
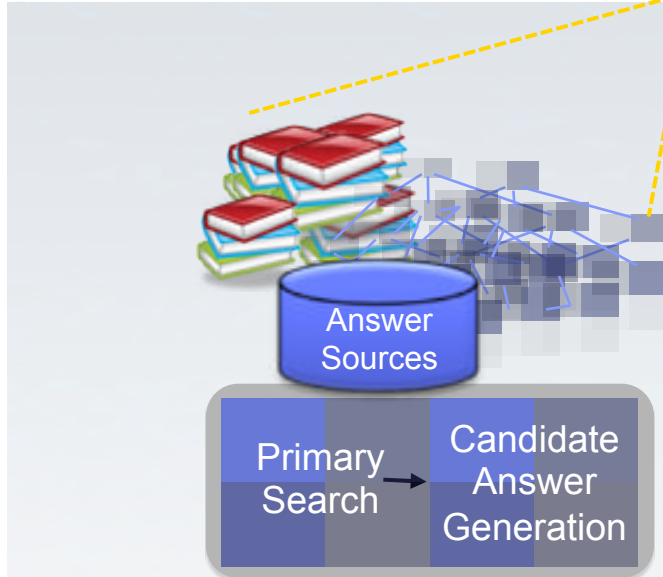
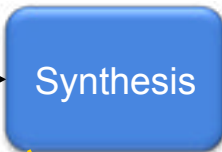
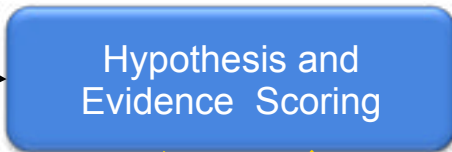
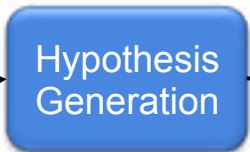
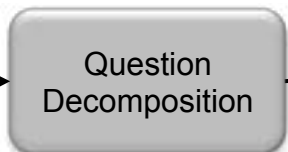
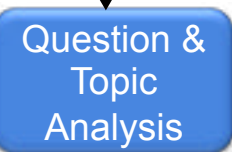
These gather, evaluate, weigh and balance different types of **evidence** to deliver the answer with the best support it can find.



Train-time

Machine Learning: System runs all the algorithms on 100,000's of questions & answers and trains on them to produce a set of *trained models*.

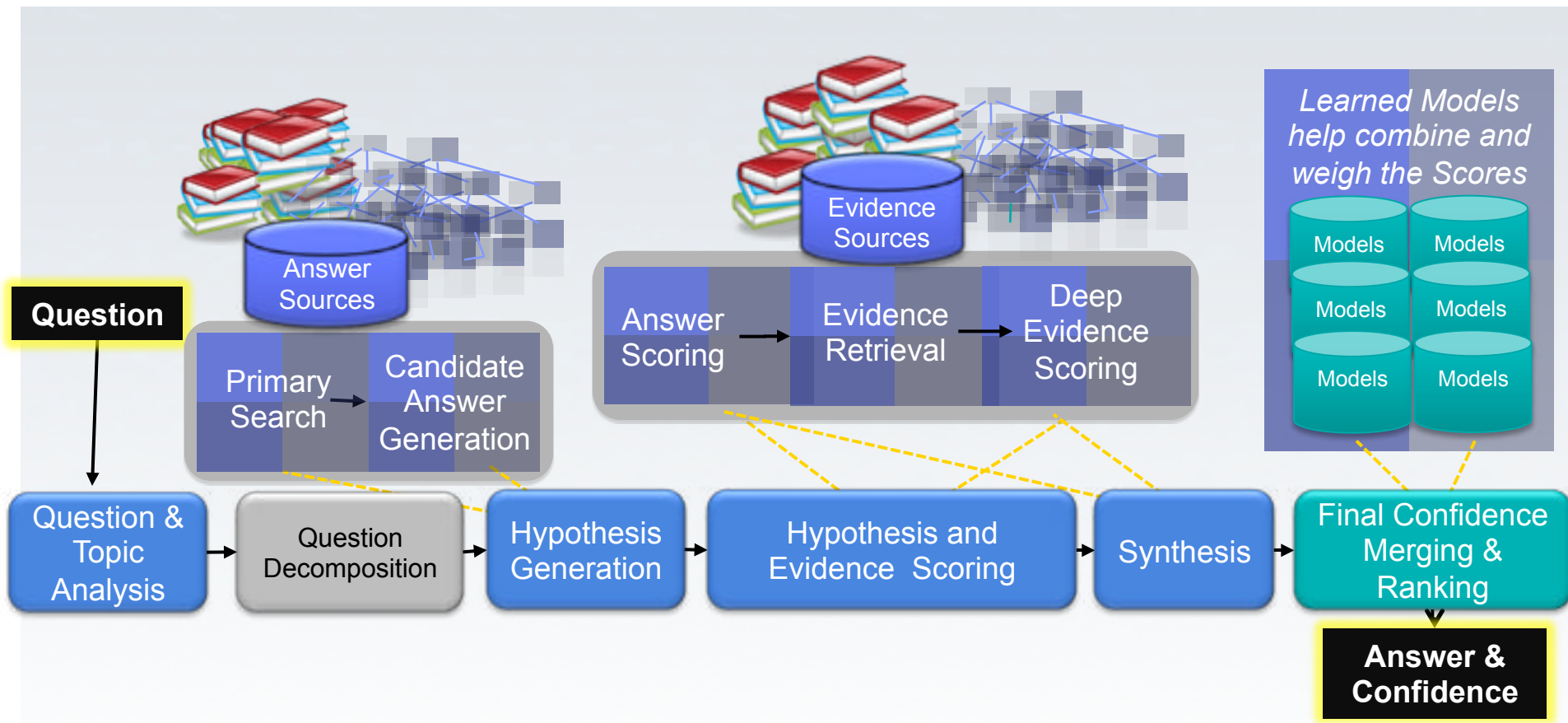
- Mixture of Experts to handle question types,
- Stacking for Domain Adaptation, Multi-answers, etc.



In 4 years DeepQA has trained for over 8500 recorded experiments

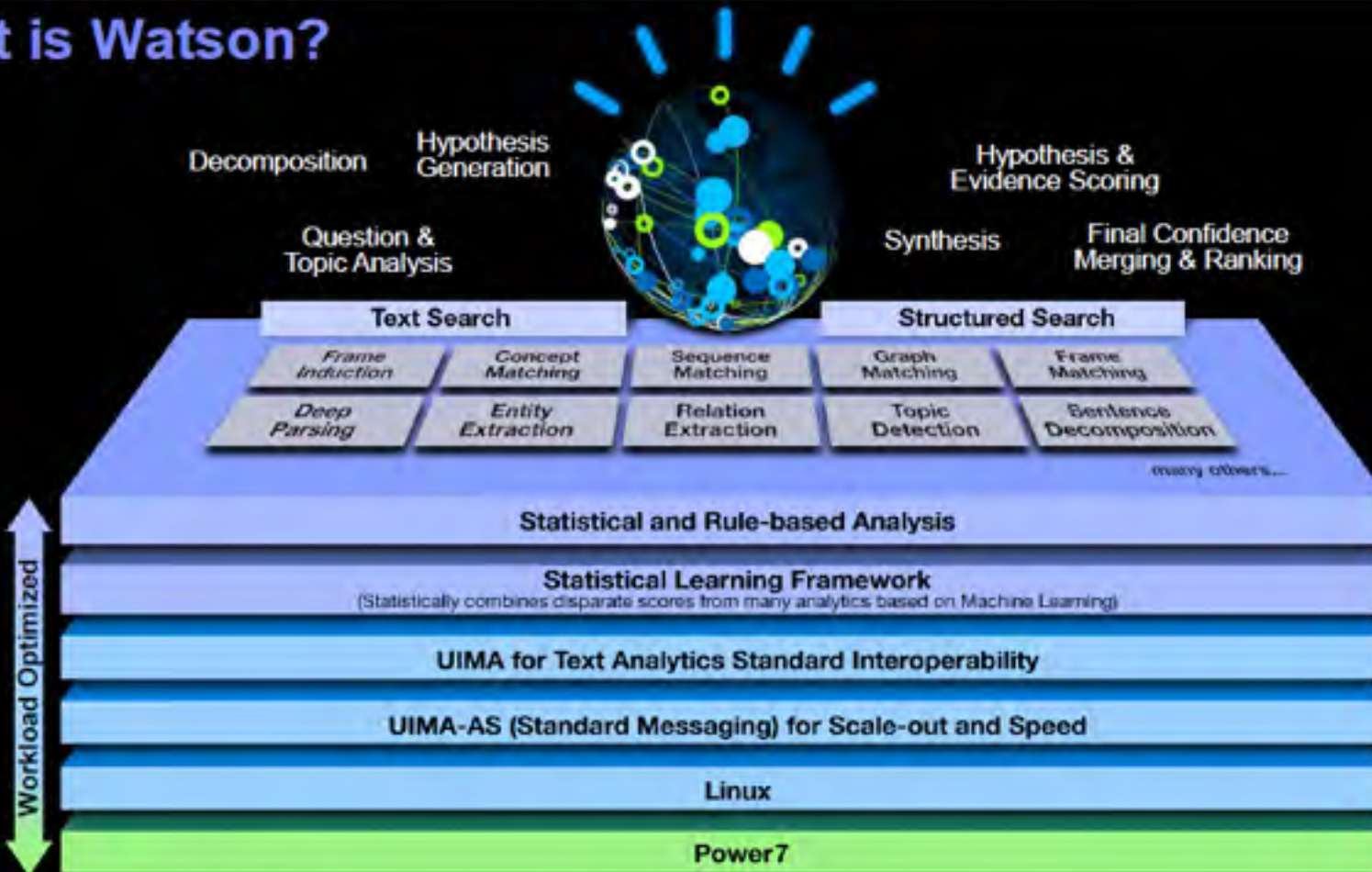
System runs all the algorithms on a *single question* and applies models to the features they produce to select an answer with a confidence

Apply-time



Run-Time Stack: Natural Language Processing on top of a complex stack

What is Watson?



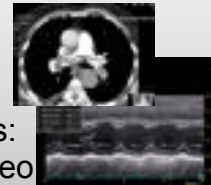
What is Watson?



UIMA Interoperability

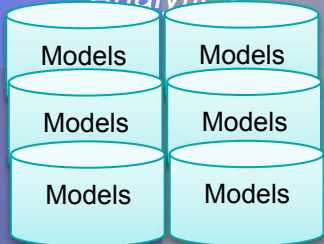
Loose Integration of Heterogeneous Analytics For Rapid Prototyping

Allows for Multimodal analytics: Image, Speech, Video



Statistical Integration of Analytics

Trained Models combine and weigh Analytics

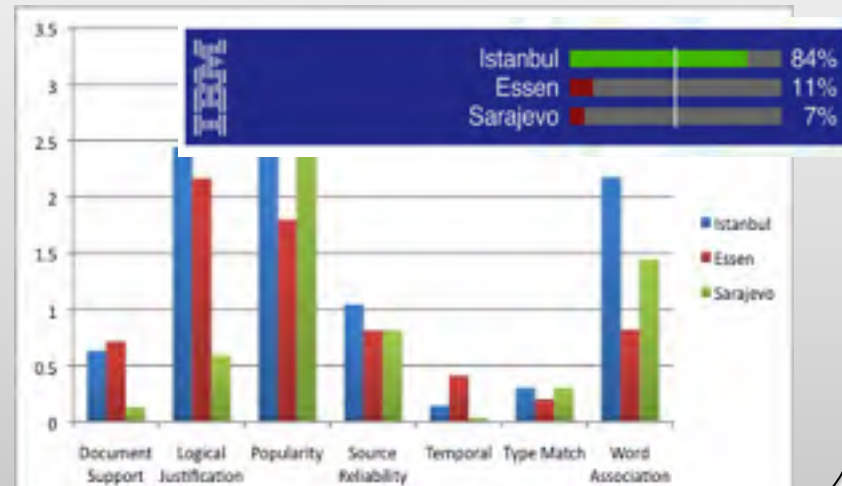


Register scores from analytics

Estimate confidence in hypothesis

- Answer ranking and confidence estimation
- Different question classes
- Small training sets
- Hypothesis decomposition and synthesis

Confidence and Evidence Dimensions



Evidence Dimensions

\$200
Keanu Reeves had a Nokia phone, but it took a land line to slip in & out of this, the title of a 1999 sci-fi flick



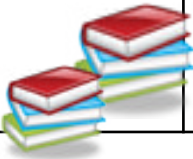


Evidence Dimensions

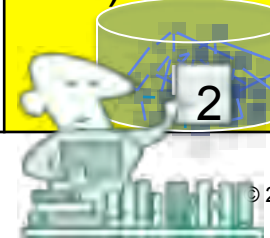
Keanu Reeves	had a Nokia Phone	took a land line to slip in & out of this	1999	Sci-fi flick
--------------	-------------------	---	------	--------------



Justification

Evidence Dimensions

KB: ActedIn(Keanu Reeves, X) 	Text Passage: had(Keanu Reeves, Nokia) 	Text Passage: Align 	KB: OccurredIn(X, 1999) 	KB: Isa(X, sci-fi flick) 
--	---	---	---	--

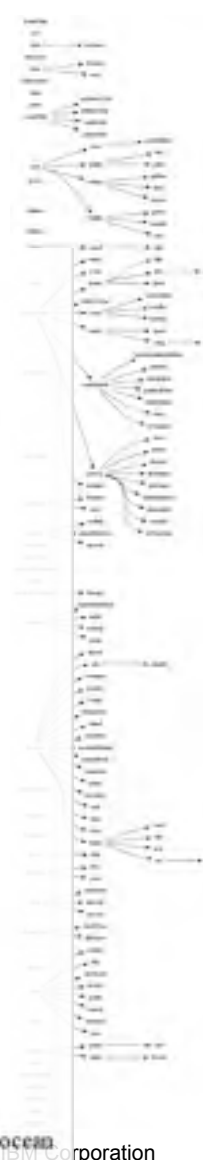
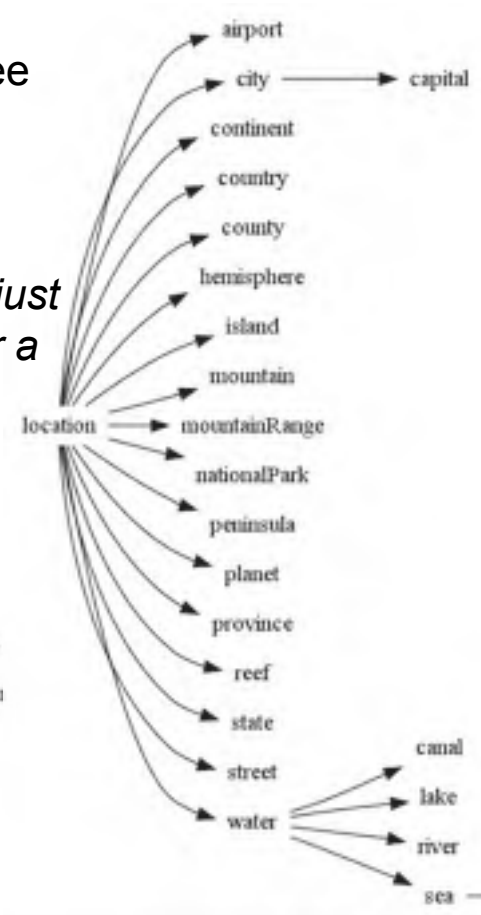
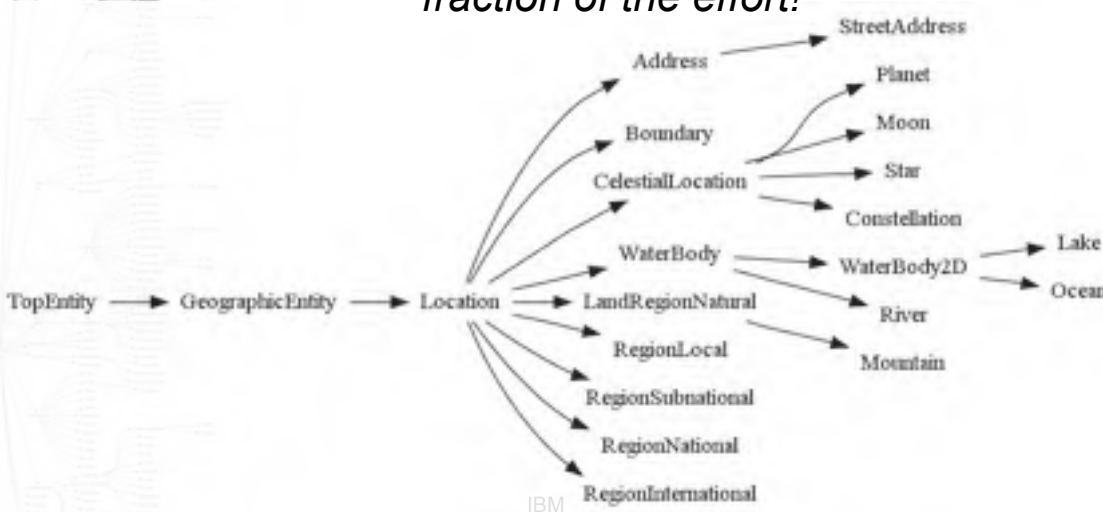
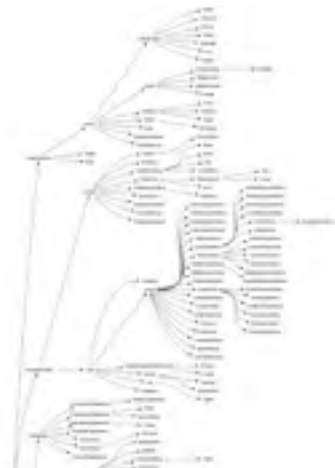


Taxonomies and Named Entity Detectors

HUTT

Ephyra

- **Adding a component which identifies types of entities**
 - Reconciling ontologies is a difficult, labor-intensive process
 - Ontologies overlap and disagree
 - **DeepQA allows for loose coupling: instead of reconciling ontologies, integrate as a new analytic**
 - **In this case, loose coupling worked just as well as ontology reconciliation for a fraction of the effort!**



The Jeopardy! Challenge: *A compelling and notable way to **drive** and **measure** the technology of automatic Question Answering along 5 Key Dimensions*

**Broad/Open
Domain**

**Complex
Language**

**High
Precision**

**Accurate
Confidence**

**High
Speed**

\$200

Keanu Reeves had a Nokia phone, but it took a land line to slip in & out of this, the title of a 1999 sci-fi **flick**

\$1000

1948: Johns Hopkins scientists find that this **antihistamine** alleviates motion sickness

\$600

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

\$2000

The first person mentioned by name in 'The Man in the Iron Mask' is this **hero** of a previous book by the same author.

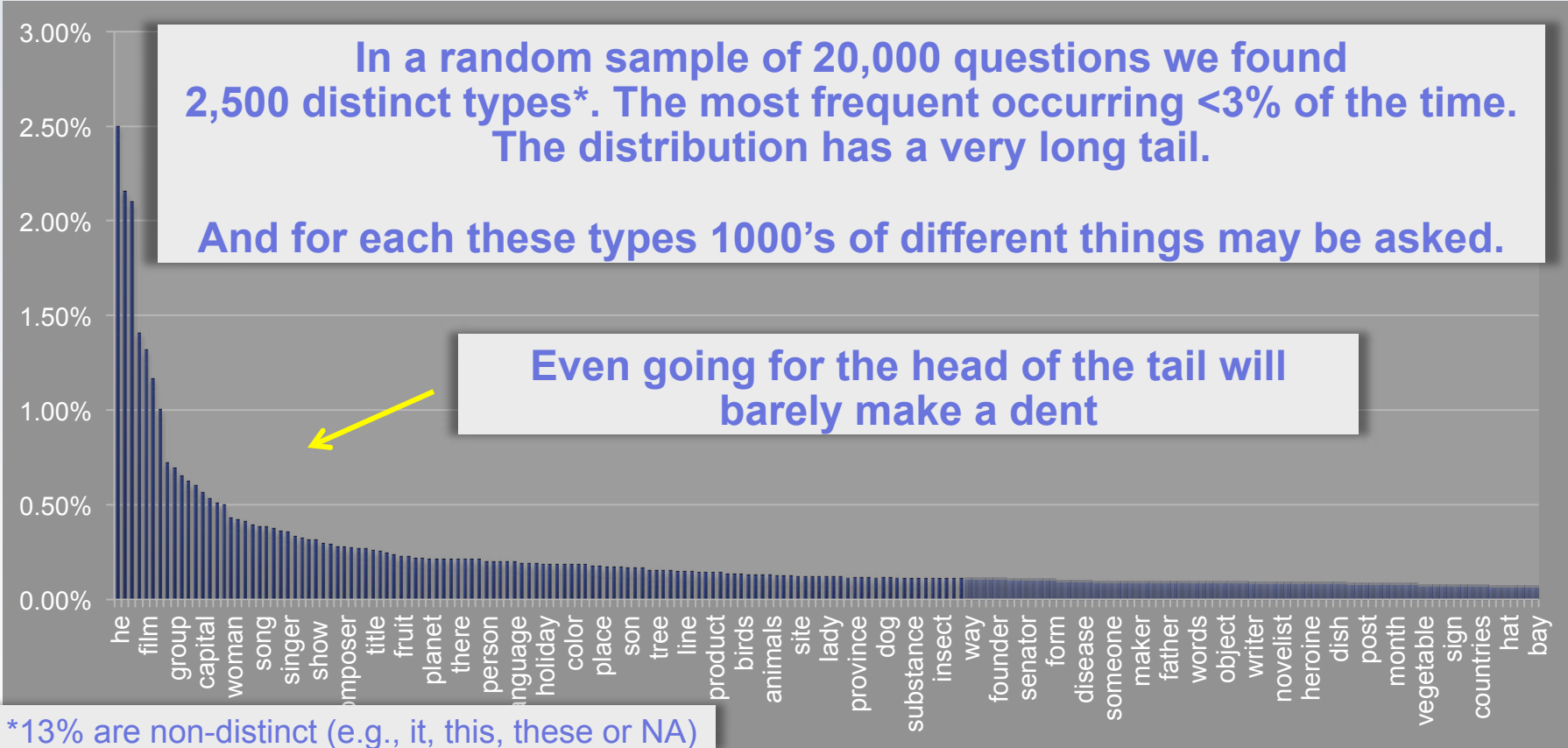
Broad Domain

We do NOT attempt to anticipate all questions and build specialized databases.

In a random sample of 20,000 questions we found 2,500 distinct types*. The most frequent occurring <3% of the time. The distribution has a very long tail.

And for each these types 1000's of different things may be asked.

Even going for the head of the tail will barely make a dent

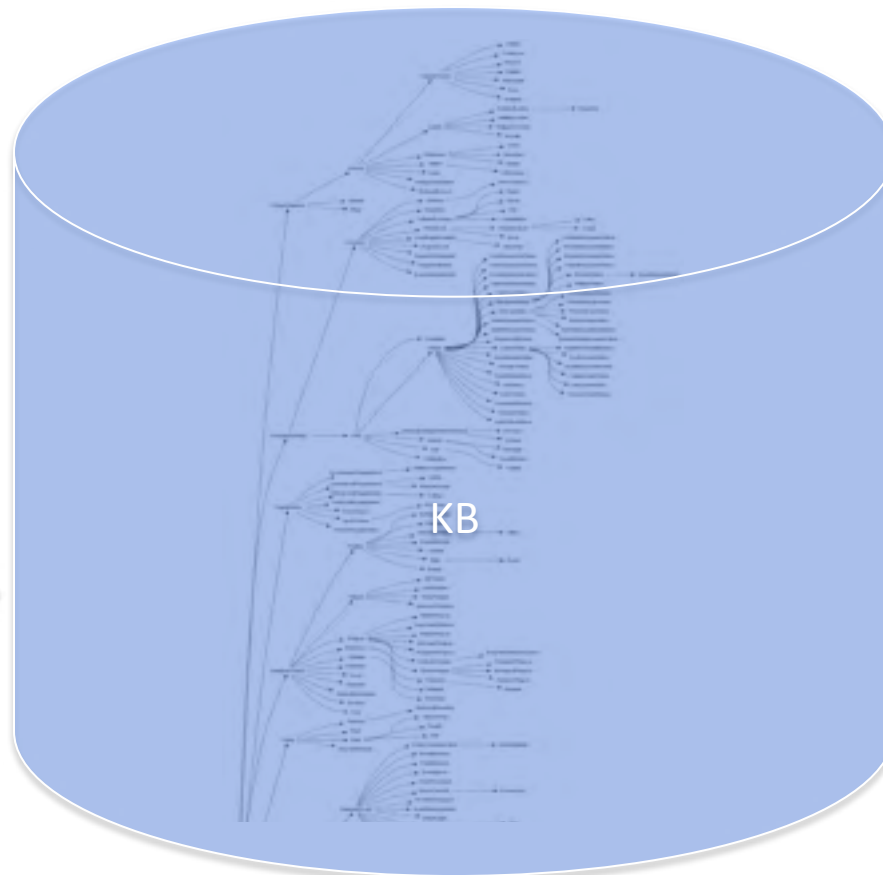
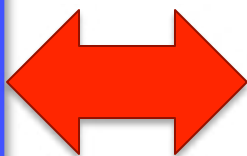


Our Focus is on reusable NLP technology for analyzing volumes of *as-is* text. Structured sources (DBs and KBs) are used to help interpret the text.

To Knowledge From Extractions (DeepQA KAFE framework)

Challenge in Open Domain
Question Answering:
*Question text doesn't match
the ontology*

The Lenin shipyards in
this Polish port city,
where Solidarity began
in the 1980s, were later
renamed

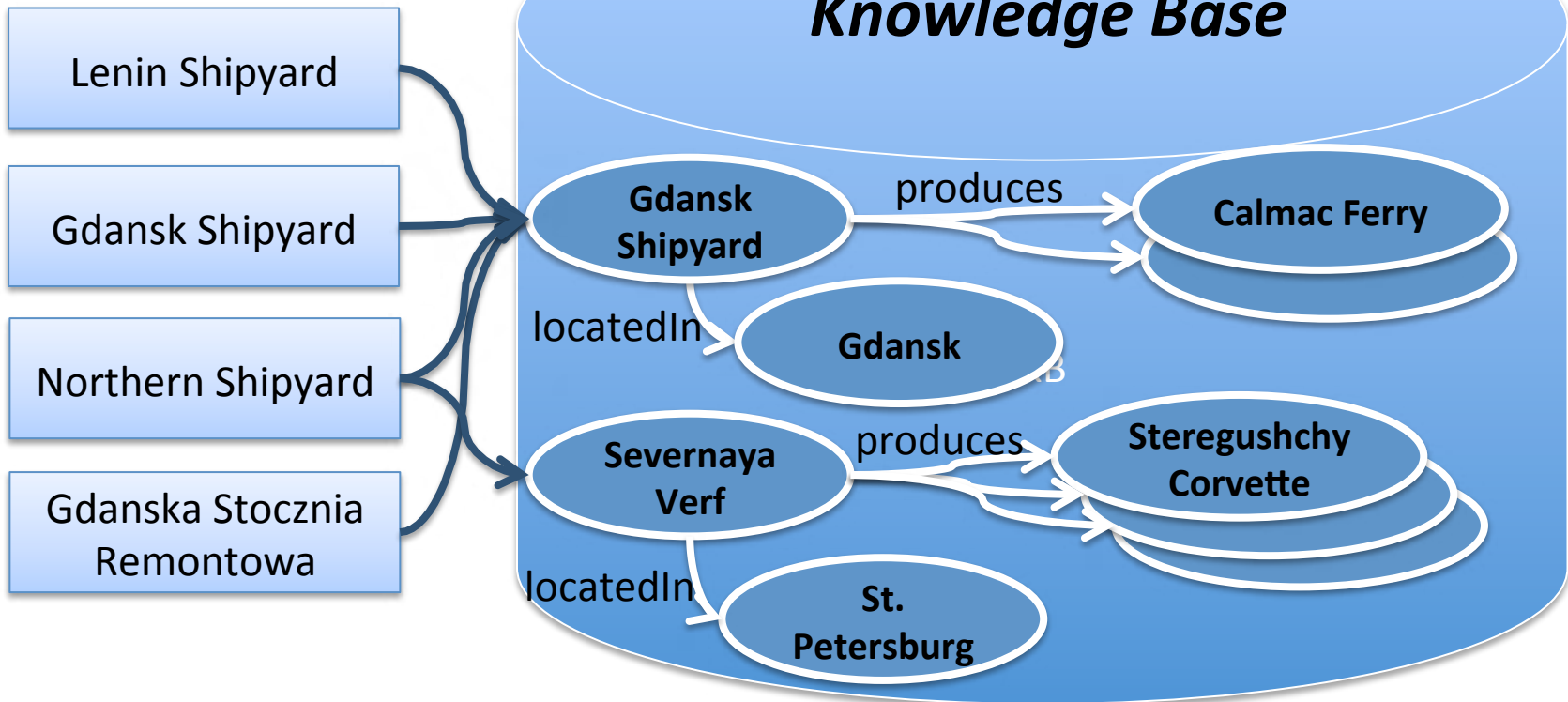


Mapping Language to Knowledge: KAFE

The Lenin shipyards in this Polish port city, where Solidarity began in the 1980s, were later renamed

KaFe Stack
Instance to Type Mapping (IceT)
Verb/Predicate to Relation Mapping (Vrappe)
Type to Type Mapping (Latte)
Instance to Instance Mapping (CHAI)

Text

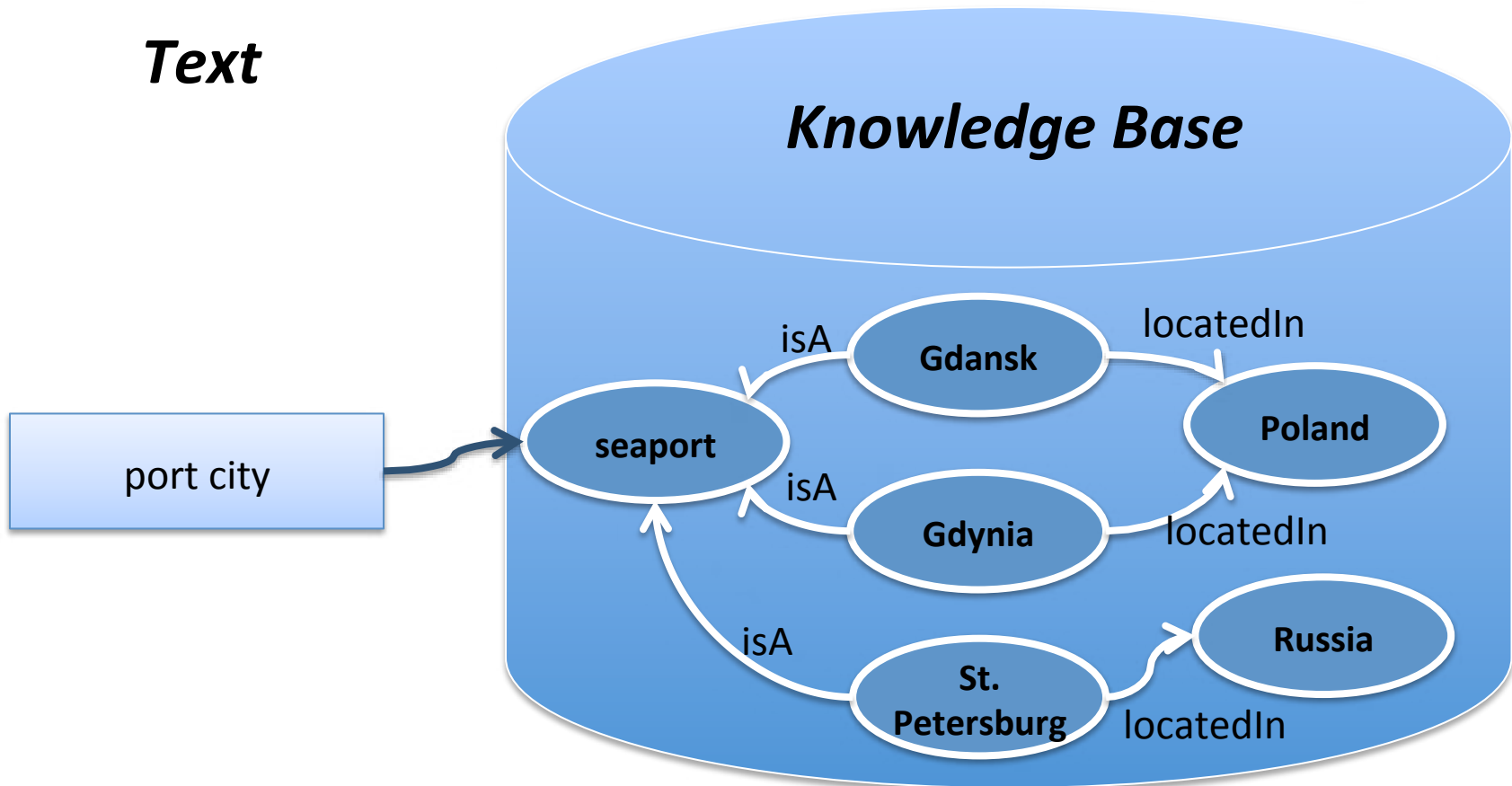


Mapping Language to Knowledge: KAFE

The Lenin shipyards in this Polish port city, where Solidarity began in the 1980s, were later renamed

KaFe Stack
Instance to Type Mapping (IceT)
Verb/Predicte to Relation Mapping (Vrappe)
Type to Type Mapping (Latte)
Instance to Instance Mapping (CHAI)

Text



Mapping Language to Knowledge: KAFE

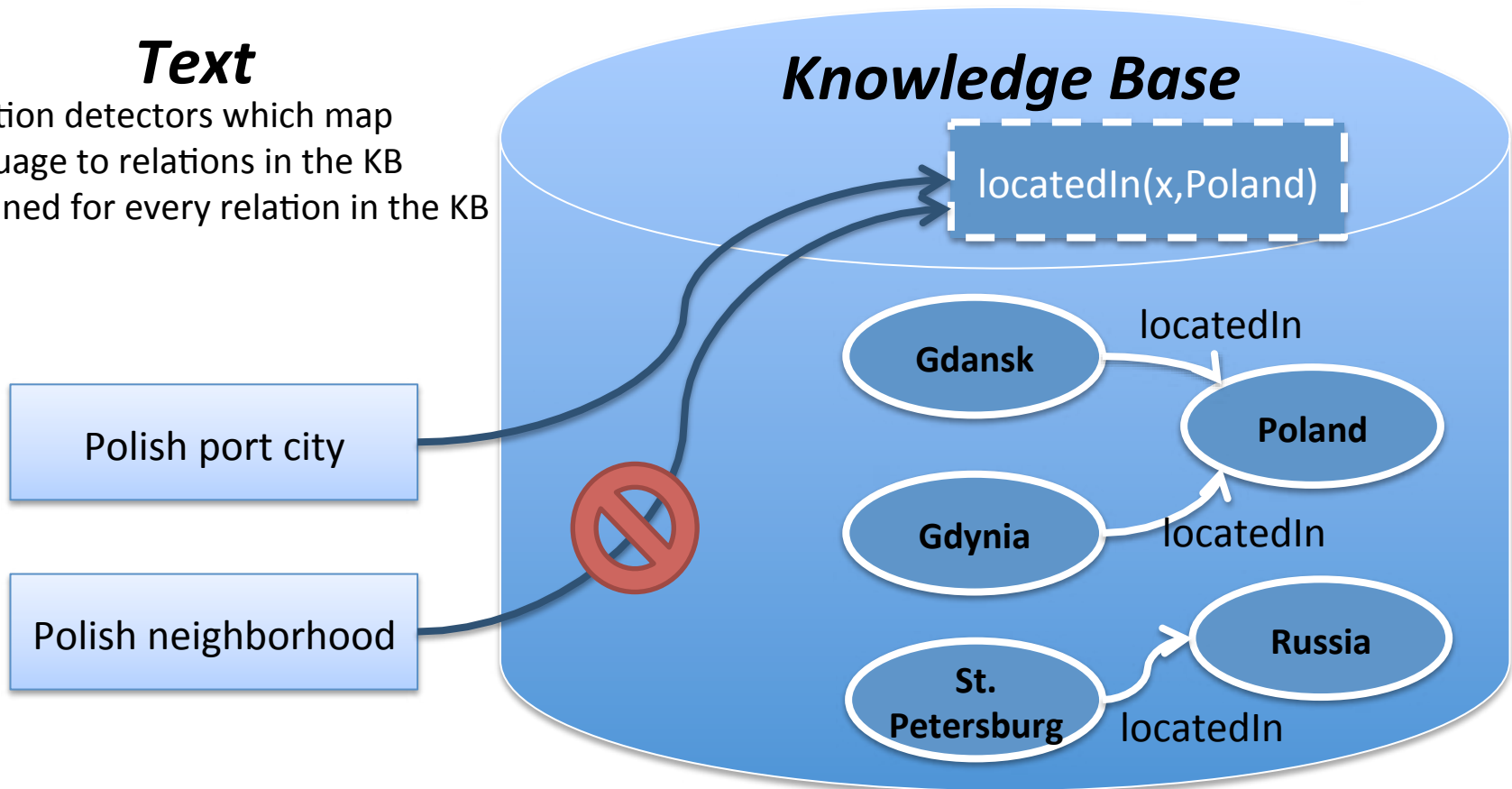
The Lenin shipyards in this Polish port city, where Solidarity began in the 1980s, were later renamed

KaFe Stack
Instance to Type Mapping (IceT)
Verb/Predicte to Relation Mapping (Vrappe)
Type to Type Mapping (Latte)
Instance to Instance Mapping (CHAI)

Text

Relation detectors which map language to relations in the KB
• trained for every relation in the KB

Knowledge Base

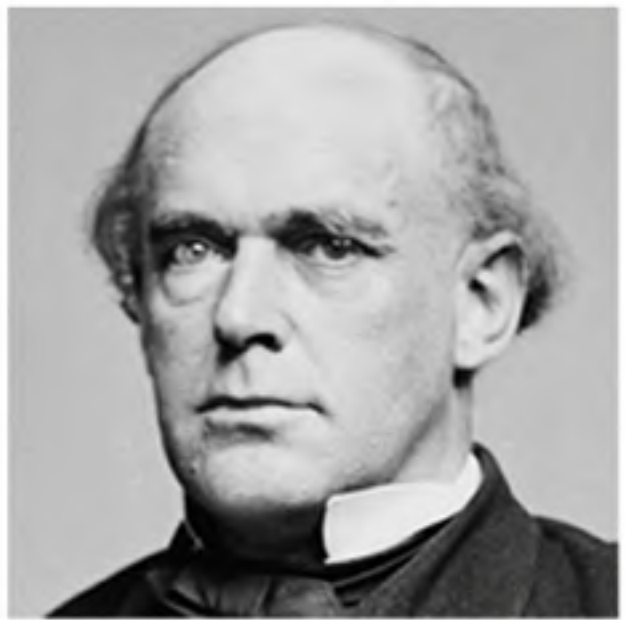


LINCOLN BLOGS

Secy. Chase just submitted this to me for the third time-- guess what, pal. This time I'm accepting it

Answer: his resignation

Treasury Secy. Chase wants to be friends with you.

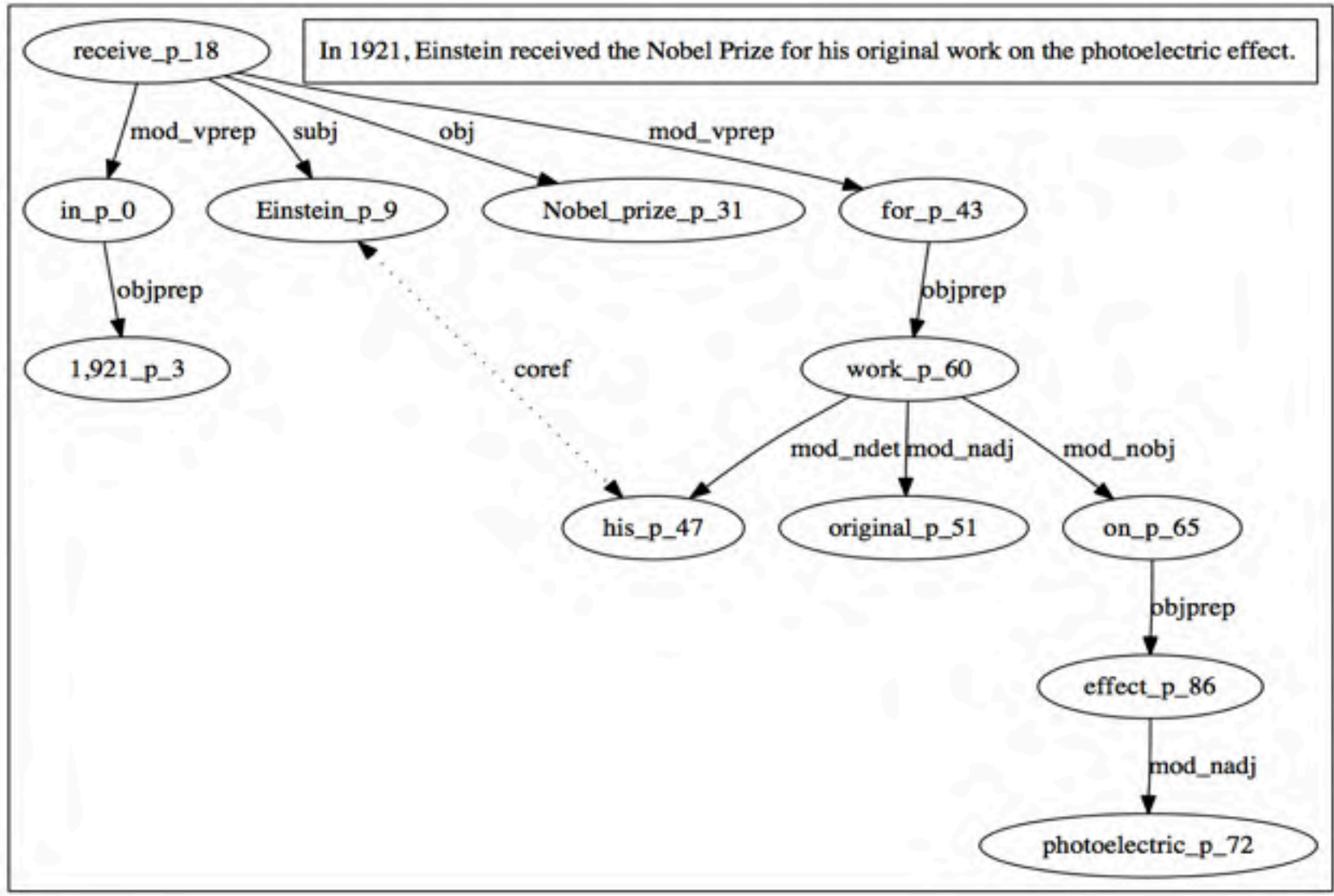


Treasury Secy. Chase

279 friends · 52 photos · 21 groups

Confirm Friend Request

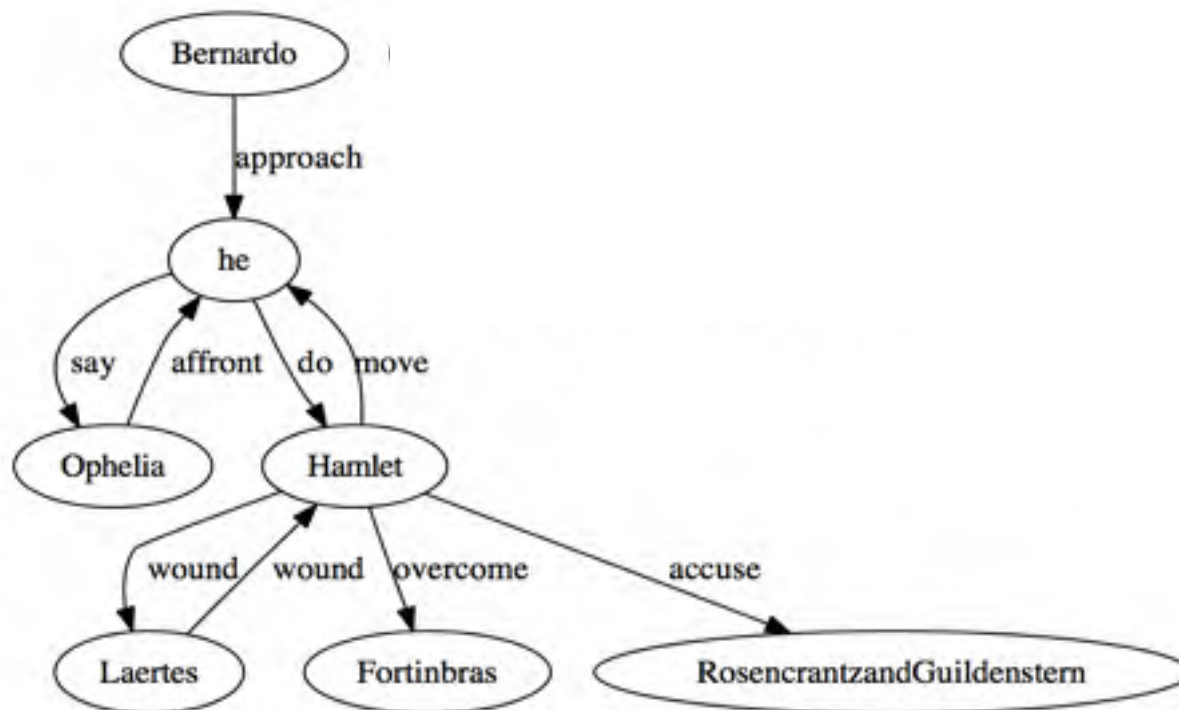
Mining over language: Dependency Parse



Extensional Knowledge

- Extensional Knowledge: Represent semantic networks extracted from text
 - Populate KBs
 - Frame inference (extract commonly occurring frame patterns)
 - Semantic relations among answers

Hamlet:
Character graph



Mining Linguistic Knowledge

Albert Einstein received his Nobel Prize...

In 1921 Einstein won the Nobel Prize for his original work on the photoelectric effect

Gerd Binnig, along with his colleague, Heinrich Rohrer, was awarded the Nobel Prize in Physics in 1986 for his work in scanning tunneling microscopy

Frame01	
verb	receive
subj	Einstein
type	PERSON/SCIENTIST
obj	Nobel Prize
mod_vprep	in

Frame02	
verb	win
subj	Einstein
type	PERSON/SCIENTIST
obj	Nobel Prize
mod_vprep	in
objprep	1921
type	YEAR
mod_vprep	For
objprep	Frame02b

Frame03	
verb	award
subj	Binnig
type	PERSON/SCIENTIST
obj	Nobel Prize in Physics
mod_vprep	in
objprep	1986

SCIENTISTS win PRIZES
 'win' = 'receive'
 'win' = 'award'
 "Nobel" isa Prize
 context: Scientists win it
 Prizes: Nobel, Turing, Macarthur, ...

Intensional Knowledge isa cuts

As with other NSAIDs, ibuprofen may be useful in the treatment of severe orthostatic hypotension

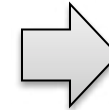
Lasix (furosemide), a diuretic, and ibuprofen, an NSAID, can be taken together

Rule-based relation detector identifies hyponymy relations in text

Frame01	
subj	Ibuprofen
type	NSAID

Frame02	
subj	Lasix
type	diuretic

Frame03	
subj	NSAID
type	drug



Ibuprofen isa NSAID

Lasix isa diuretic

NSAID isa drug

.

.

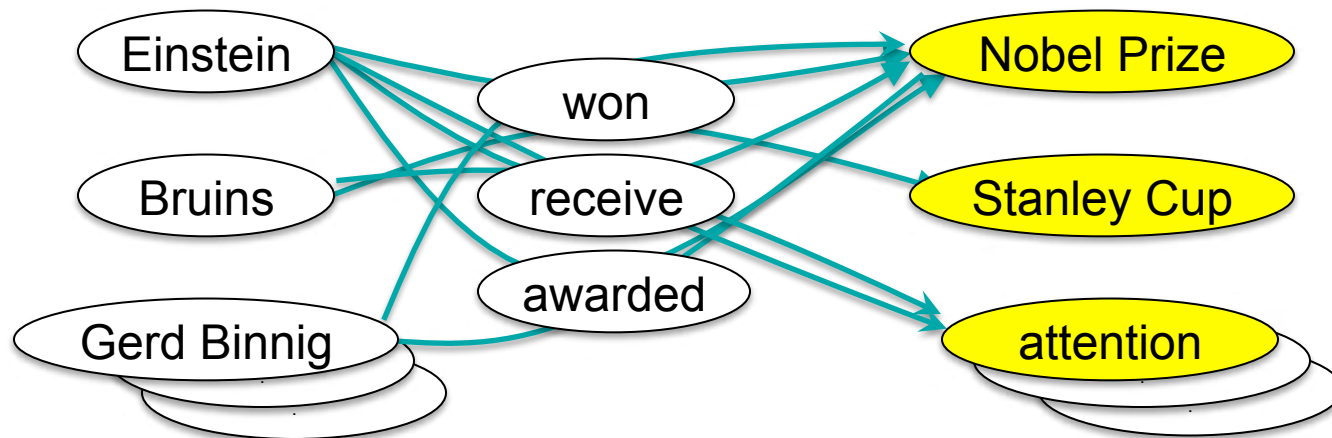
200M



Prismatic Lambda Abstractions

SVO (Subject Verb Object frame cuts)

$\lambda_o[\text{subj}=\text{Einstein}, \text{verb}=\text{won}]$: vector of counts of object filler



~800M SVO cuts...

Frequencies from λ vectors allow you to compute conditional probabilities:

$$P(o=\text{Nobel Prize} \mid \text{subj}=\text{Einstein}, \text{verb}=\text{won})$$

Type information (selectional restrictions)

$$P(o=\text{Nobel Prize} \mid \text{subj}=\text{SCIENTIST}, \text{verb}=\text{won})$$

Bootstrapping

$$\lambda_o[\text{subj}=\text{f}(\{\lambda_s[\text{verb}=\text{won}, \text{obj}=\text{Nobel Prize}]\}), \text{verb}=\text{won}]$$

Unsupervised Learning

Topics Clusters : term co-occurrence in documents - LSA



Type Clusters: terms share similar syntactic roles
(Prismatic)

explorer
firefox
safari
opera
chrome

ibm
dell
hp
acer
asus

rome
paris
london
venice
milan
florence

Inferring Lexical Answer Type (LAT) from Context

Question	LAT Inferred
ART & ARTISTS: Picasso painted this work as a protest against the bombing of a town in the Spanish Civil War	Painting
SAY "UNC"LE: Redness & itching are symptoms of this eye condition	Disease
FAMOUS LATINOS & LATINAS: From the Bronx projects & Princeton, she joined the Supreme Court in 2009	Chief Justice
AMERICAN LITERATURE: Shortly after "The House of the Seven Gables," he wrote a book of classical myths, A Wonder Book for Girls and Boys	Author

Evidence Dimensions

\$200
Keanu Reeves had a Nokia phone, but it took a land line to slip in & out of this, the title of a 1999 sci-fi flick

Evidence Dimensions

Keanu Reeves	had a Nokia Phone	took a land line to slip in & out of this	1999	Sci-fi flick
--------------	-------------------	---	------	--------------

Justification

Evidence Dimensions

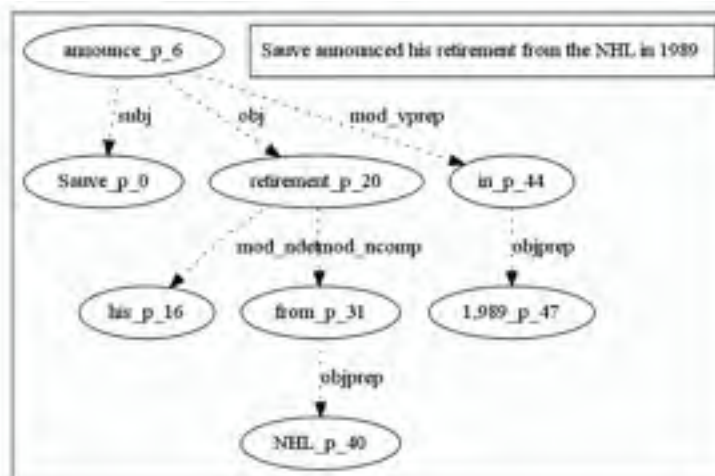
KB: ActedIn(Keanu Reeves, X) 	Text Passage: had(Keanu Reeves, Nokia) 	Text Passage: Align 	KB: OccurredIn(X, 1999) 	KB: Isa(X,sci-fi flick)
--	---	----------------------------	---------------------------------------	---------------------------------------



TWREX: Topicalized Wide-scale Relation Extraction [EMNLP 2011]

Jeopardy! Clue	Relation Detected (in italics)
<i>MOTHERS & SONS: Though only separated by one year in real life, she played mother to son Colin Farrell in "Alexander"</i>	<i>starring</i> (she, "Alexander")
<i>THE DEVIL: "The Screwtape Letters" from a senior devil to an under devil are by this man better known for children's books</i>	<i>author</i> (man, "The Screwtape Letters")
<i>THE LONE REPRESENTATIVE: Michael Castle from this state with 3 counties: New Castle, Kent and Sussex</i>	<i>residence</i> ("Michael Castle", state)

"Sauve announced his retirement from the NHL in 1989".



Evidence Dimensions






\$200
Keanu Reeves had a Nokia phone, but it took a land line to slip in & out of this, the title of a 1999 sci-fi flick

Evidence Dimensions

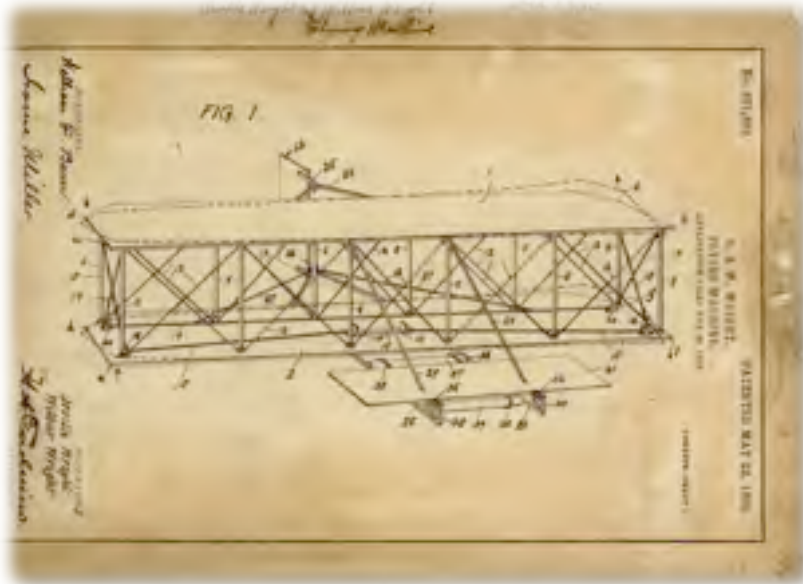
Keanu Reeves	had a Nokia Phone	took a land line to slip in & out of this	1999	Sci-fi flick
--------------	-------------------	---	------	--------------

Justification

Evidence Dimensions

KB: ActedIn(Keanu Reeves, X) 	Text Passage: had(Keanu Reeves, Nokia) 	Text Passage: Align  3	KB: OccurredIn(X, 1999) 	KB: Isa(X, sci-fi flick) 
--	--	--	---	--

Distance between question and justifying passage



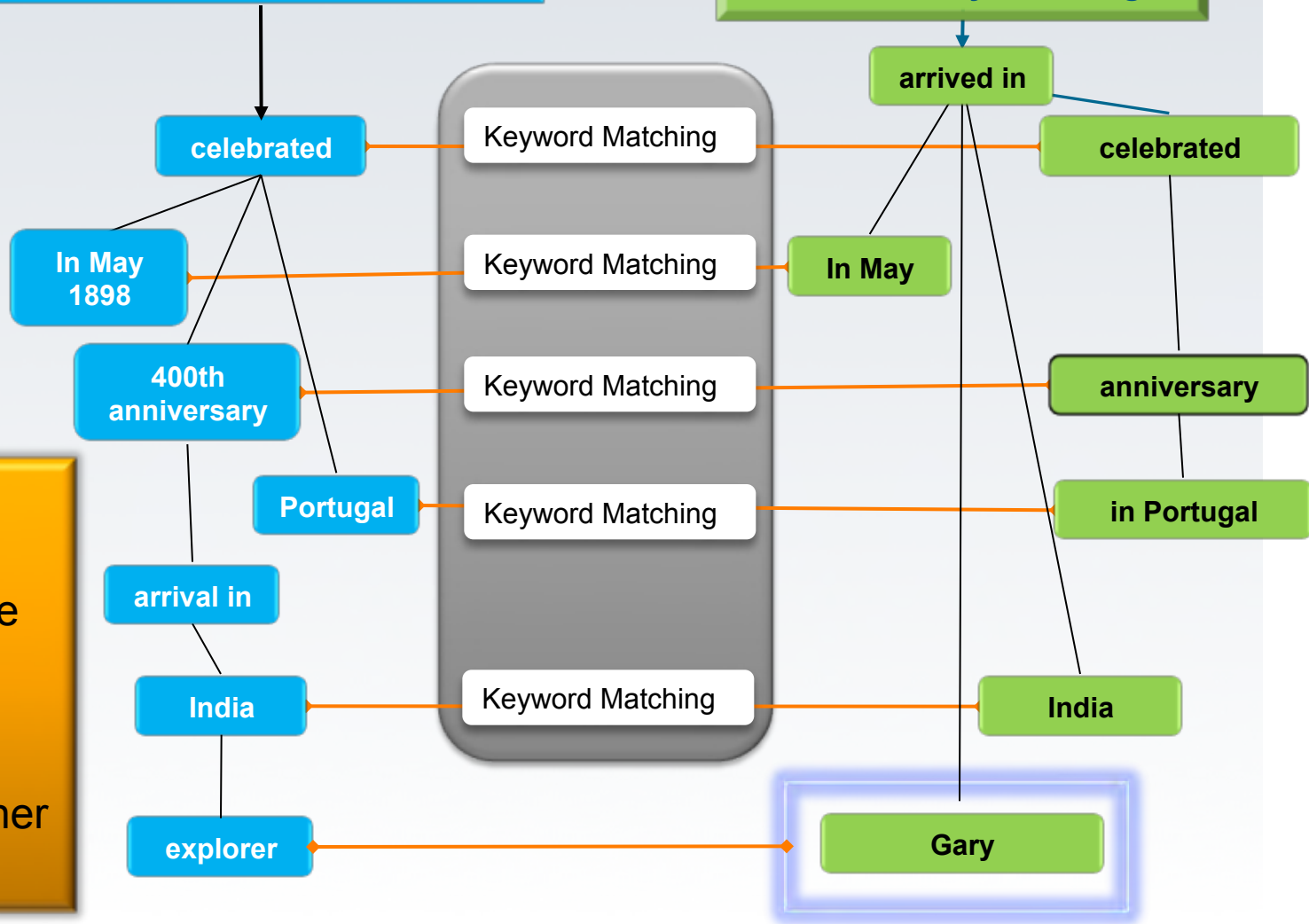
The Wright brothers' first flight was this long.

The Wright brother's first flight was 120 feet long

With Orville at the controls and Wilbur running along side to steady the wing, the plane rose 12 feet into the air and went about 120 feet on its 12-second trip. This marked the beginning of air travel for mankind.

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

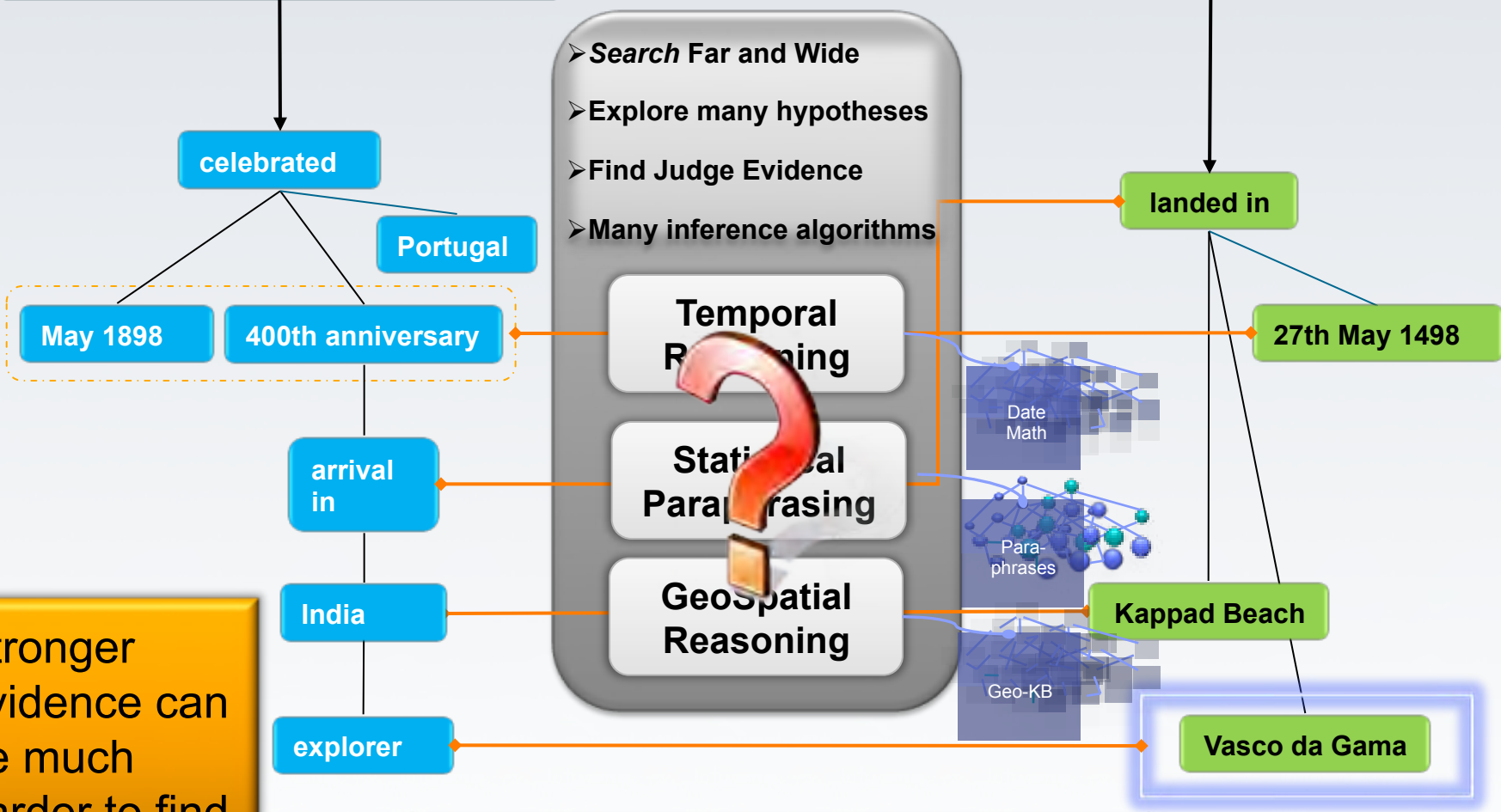
In May, Gary arrived in India after he celebrated his anniversary in Portugal.



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

Evidence Dimensions

\$200
Keanu Reeves had a Nokia phone, but it took a land line to slip in & out of this, the title of a 1999 sci-fi flick



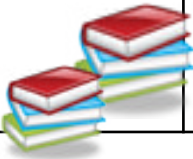


Evidence Dimensions

Keanu Reeves	had a Nokia Phone	took a land line to slip in & out of this	1999	Sci-fi flick
--------------	-------------------	---	------	--------------

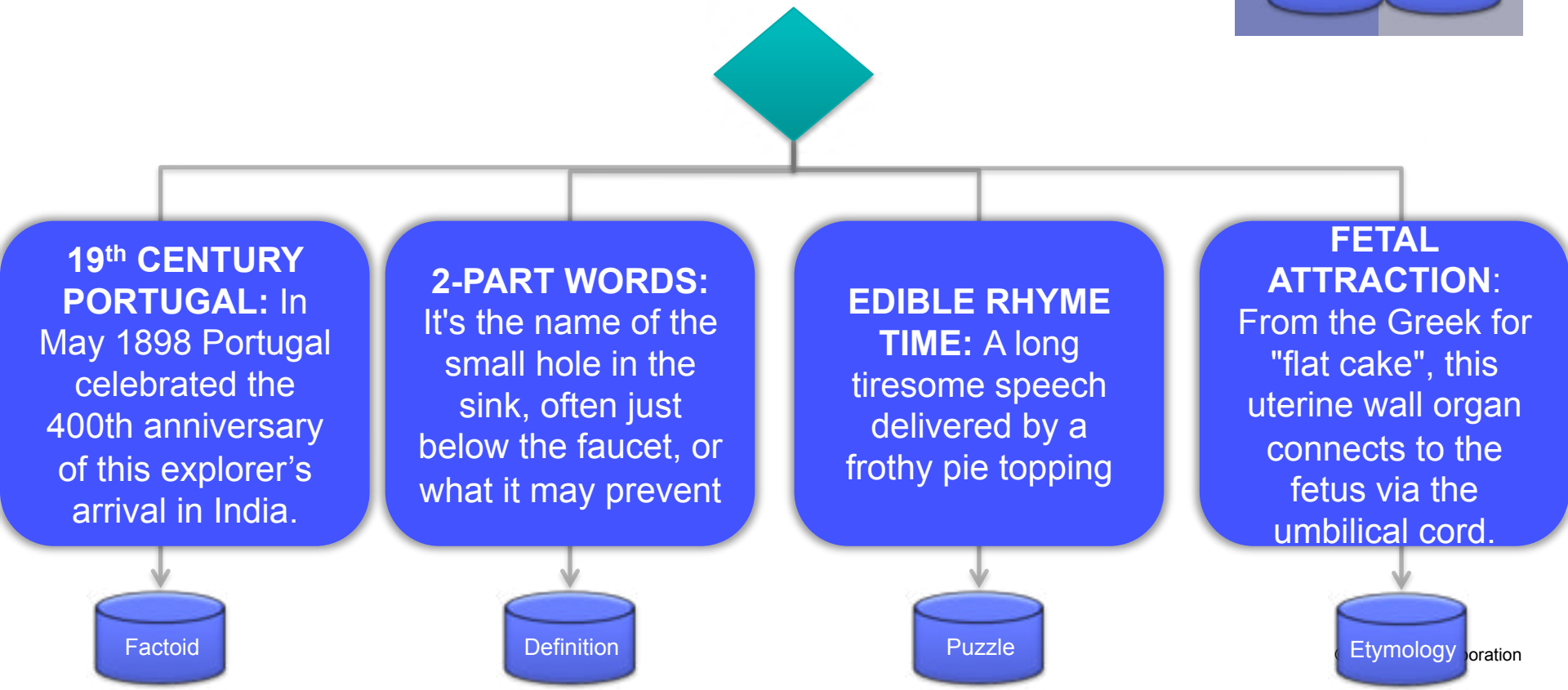
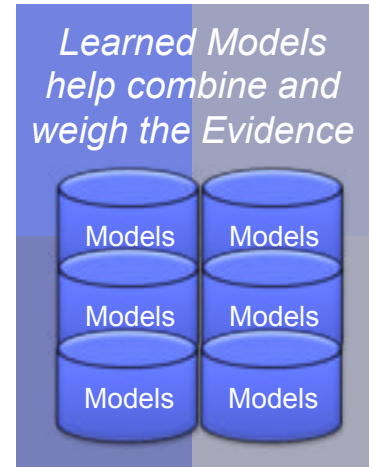


Justification

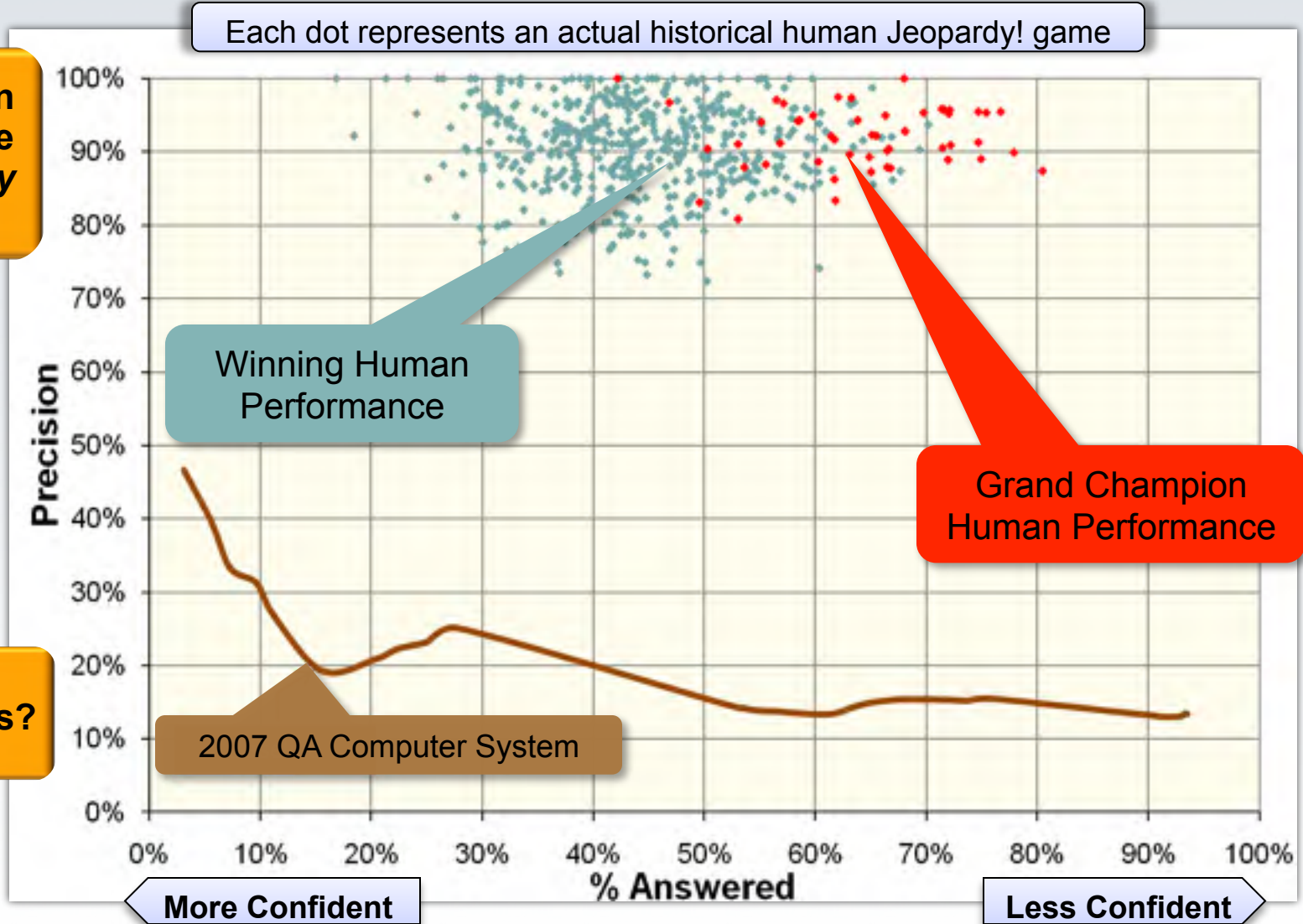
Evidence Dimensions

Evidence Sources KB: ActedIn(Keanu Reeves, X) 	Text Passage: has(Keanu Reeves, Nokia) 	Text Passage: Align 	KB: OccurredIn(X, 1999) 	KB: Isa(X,sci-fi flick) 
--	--	--	---	---

Different question classes weigh evidence differently
 Statistical and rule-based classifiers identify question class
 Partitioned Mixture of Experts trained for each question class



The Best Human Performance: *Our Analysis Reveals the Winner's Cloud*



Top human players are remarkably good.

Computers?

Winning Human Performance

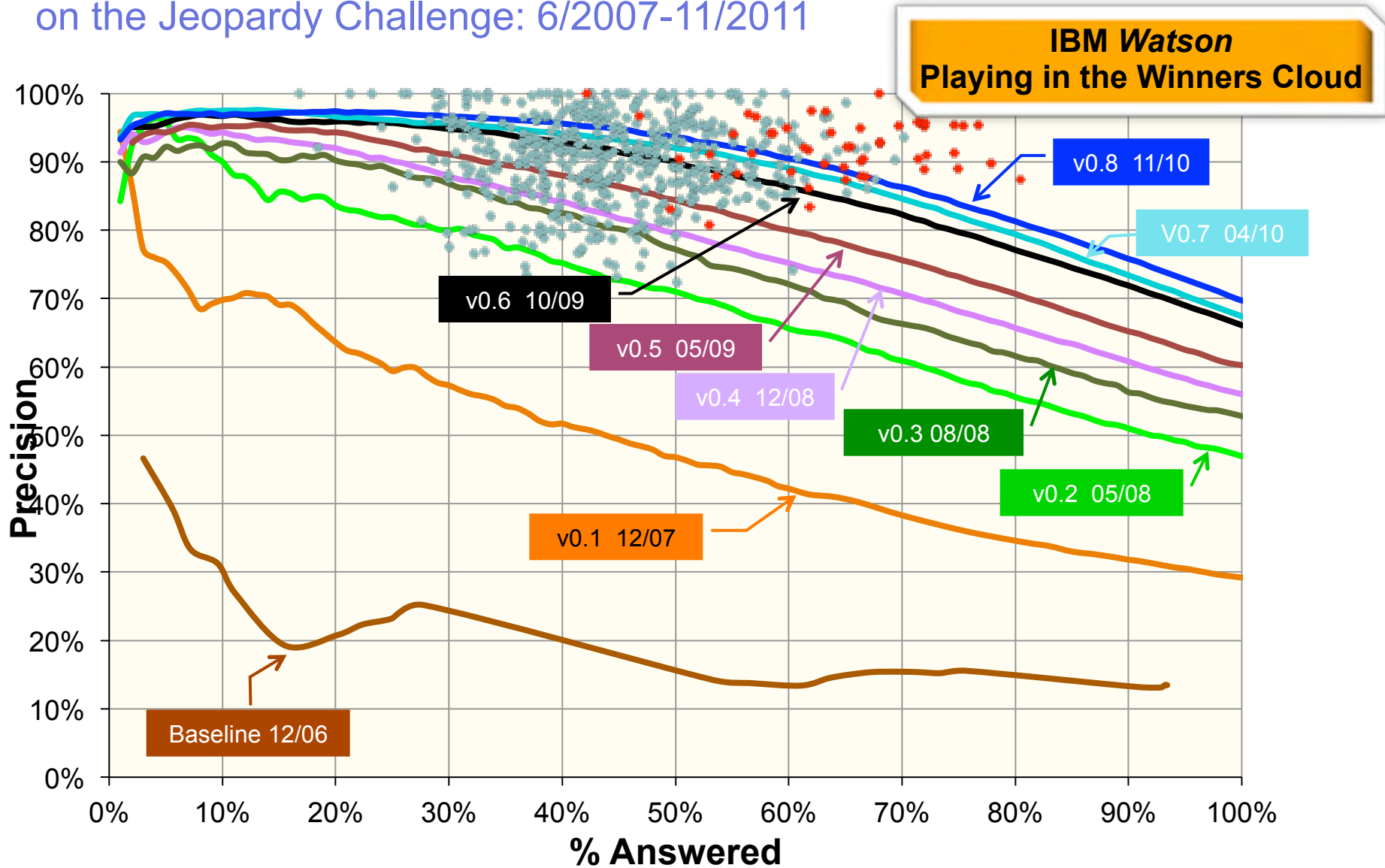
Grand Champion Human Performance

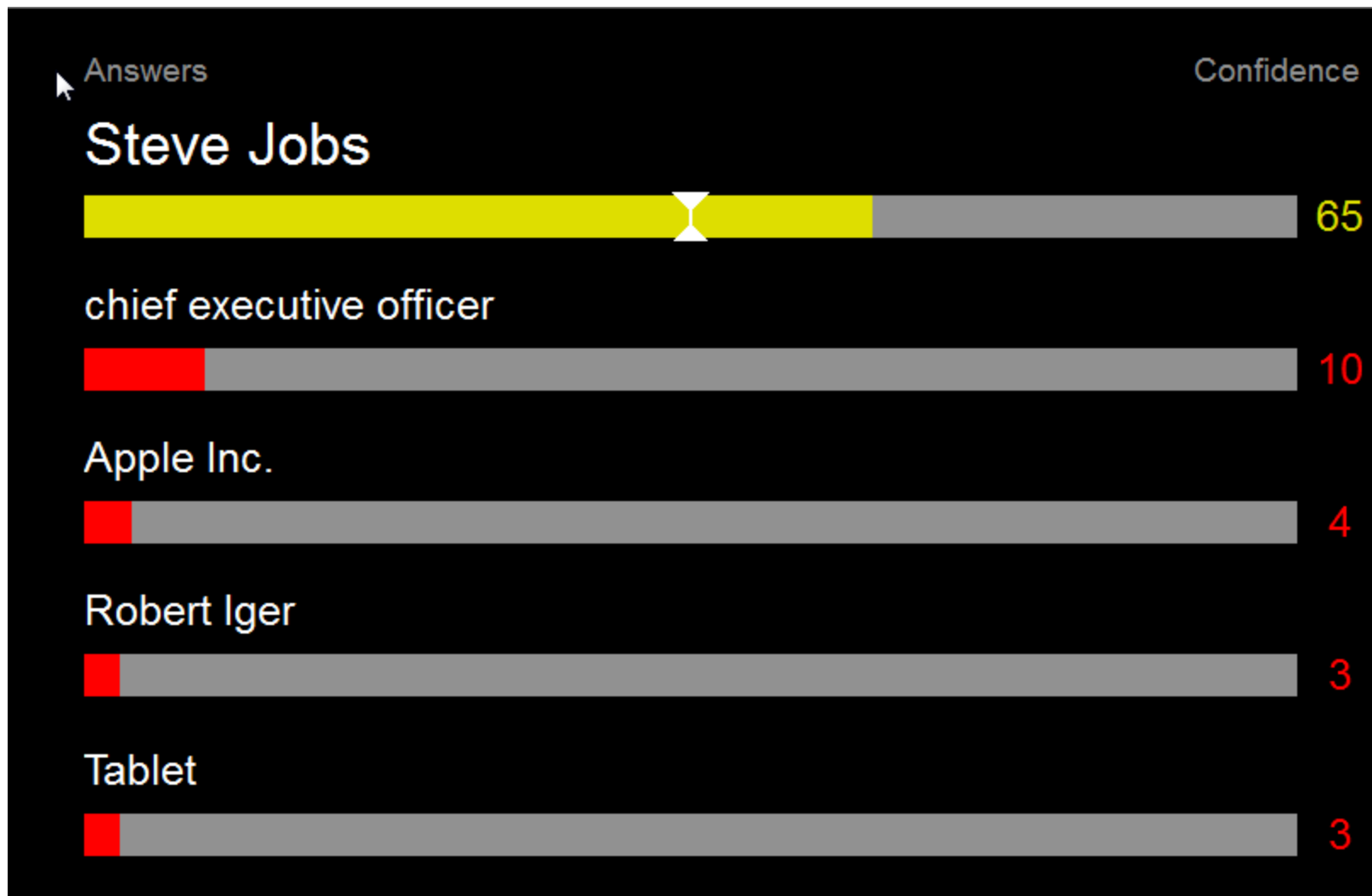
2007 QA Computer System

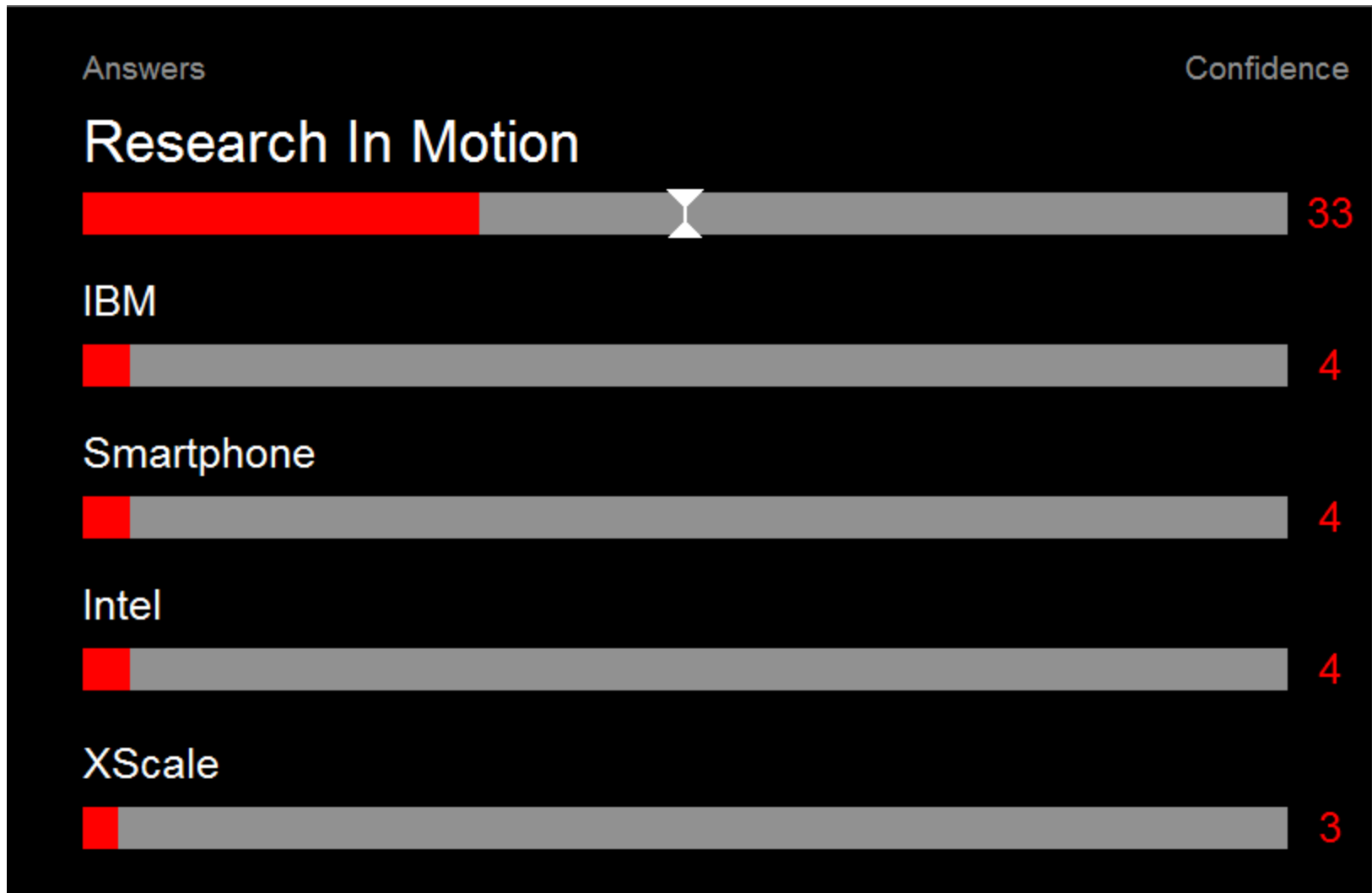
More Confident

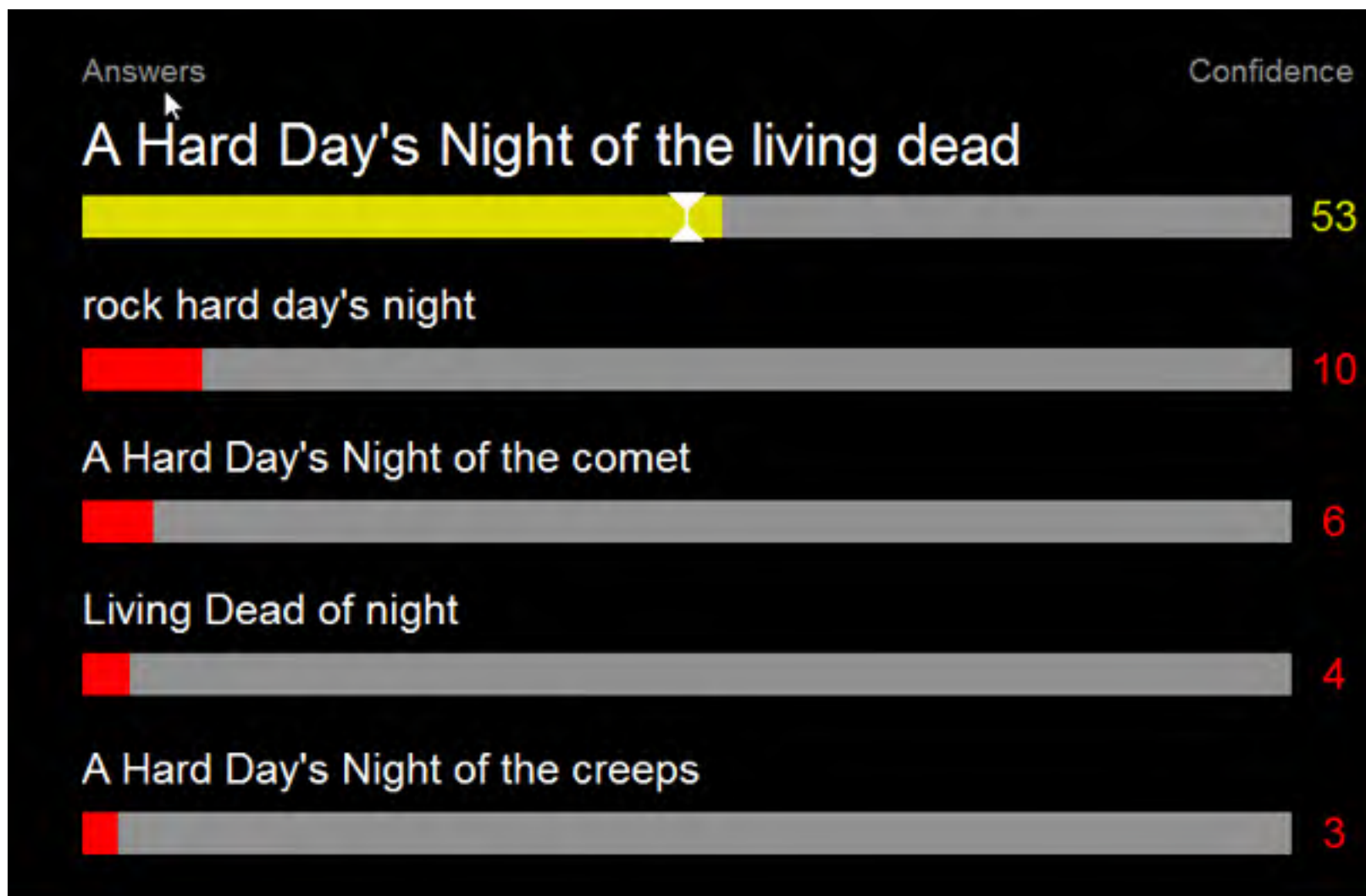
Less Confident

DeepQA: Incremental Progress in Answering Precision on the Jeopardy Challenge: 6/2007-11/2011



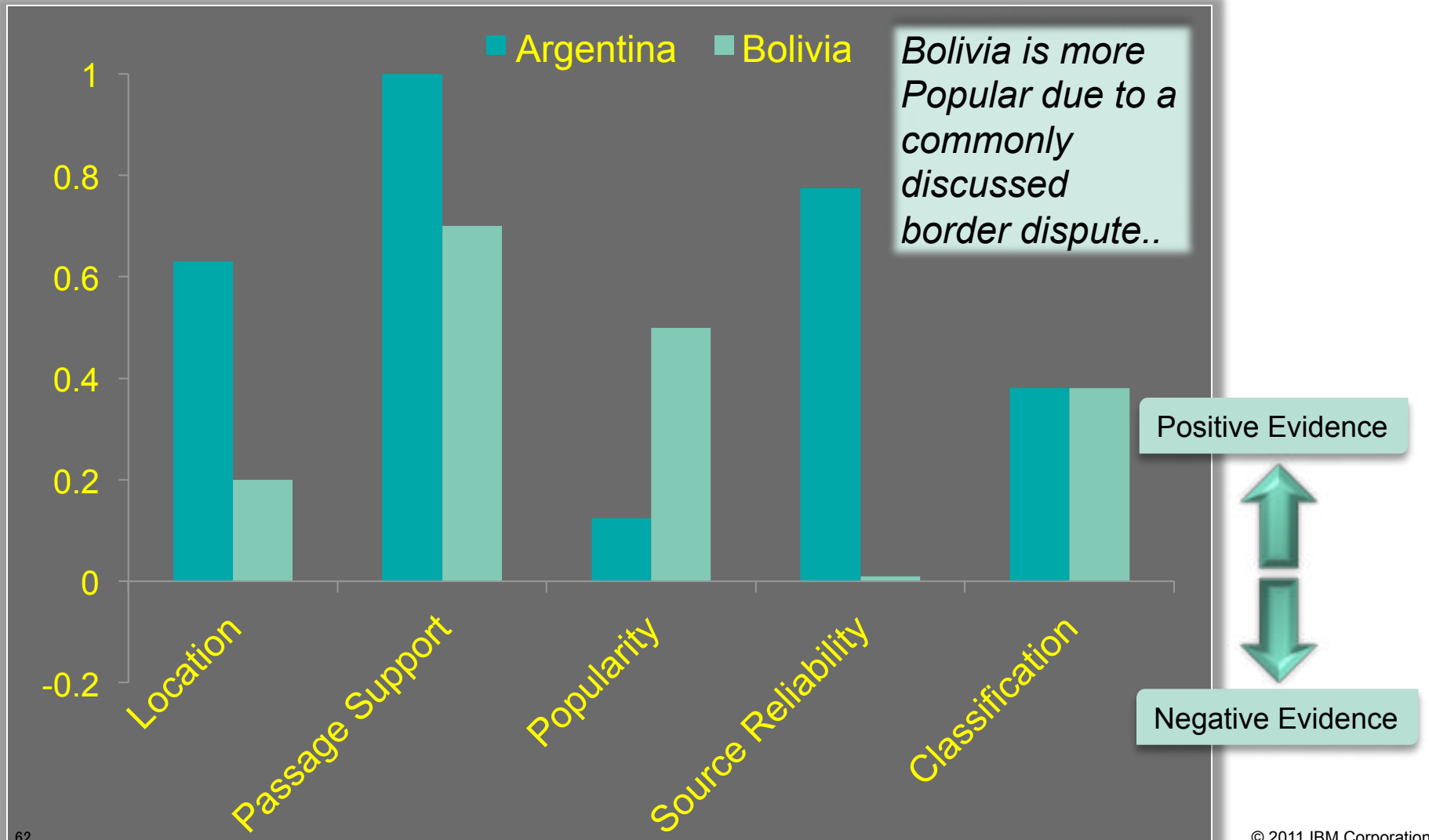






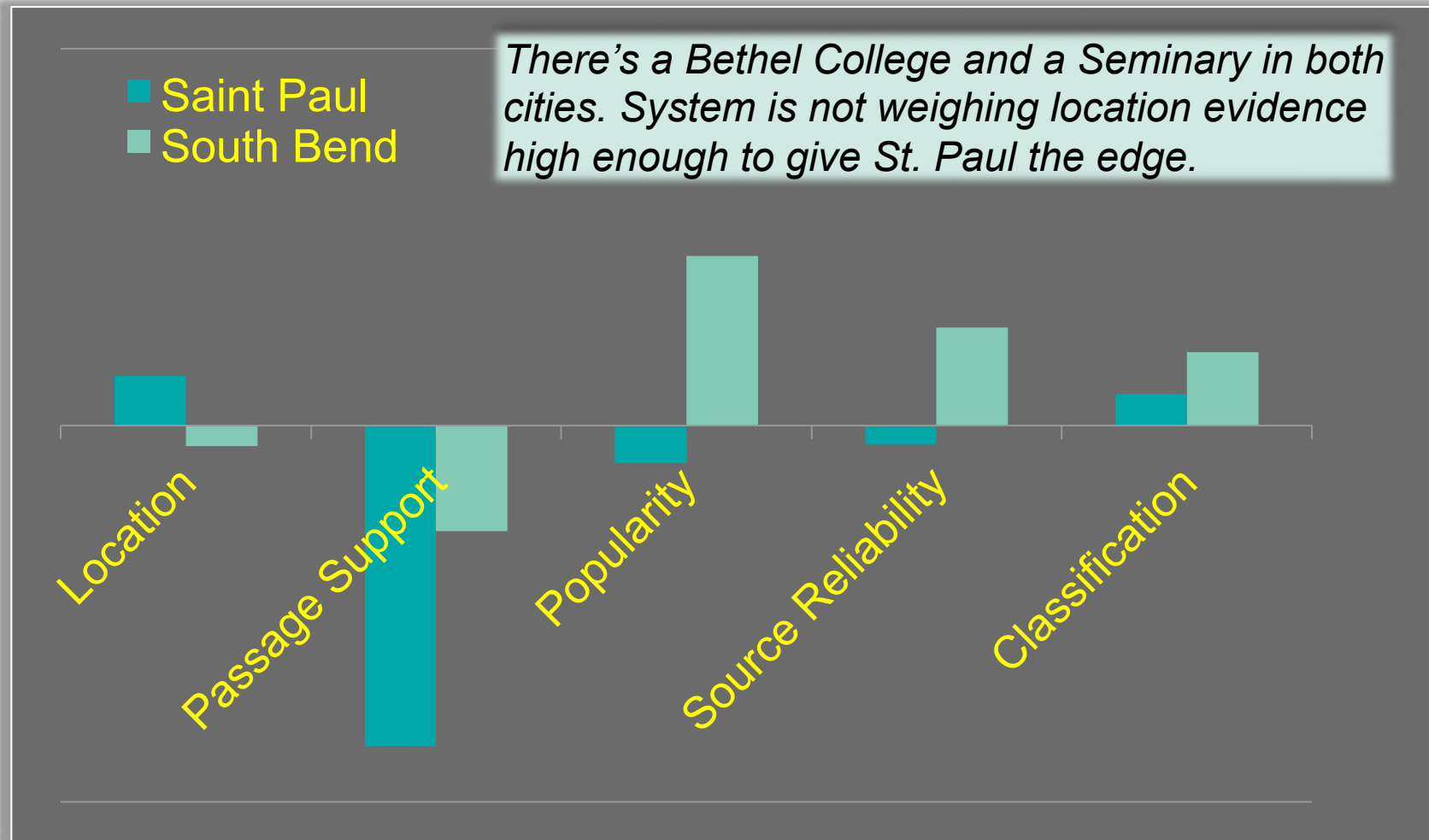
Grouping Features to produce *Evidence Profiles*

Clue: Chile shares its longest land border with this country.



Evidence: Time, Popularity, Source, Classification etc.

Clue: You'll find Bethel College and a Seminary in this "holy" Minnesota city.



Evidence: Puns

Clue: You'll find Bethel College and a Seminary in this "holy" Minnesota city.

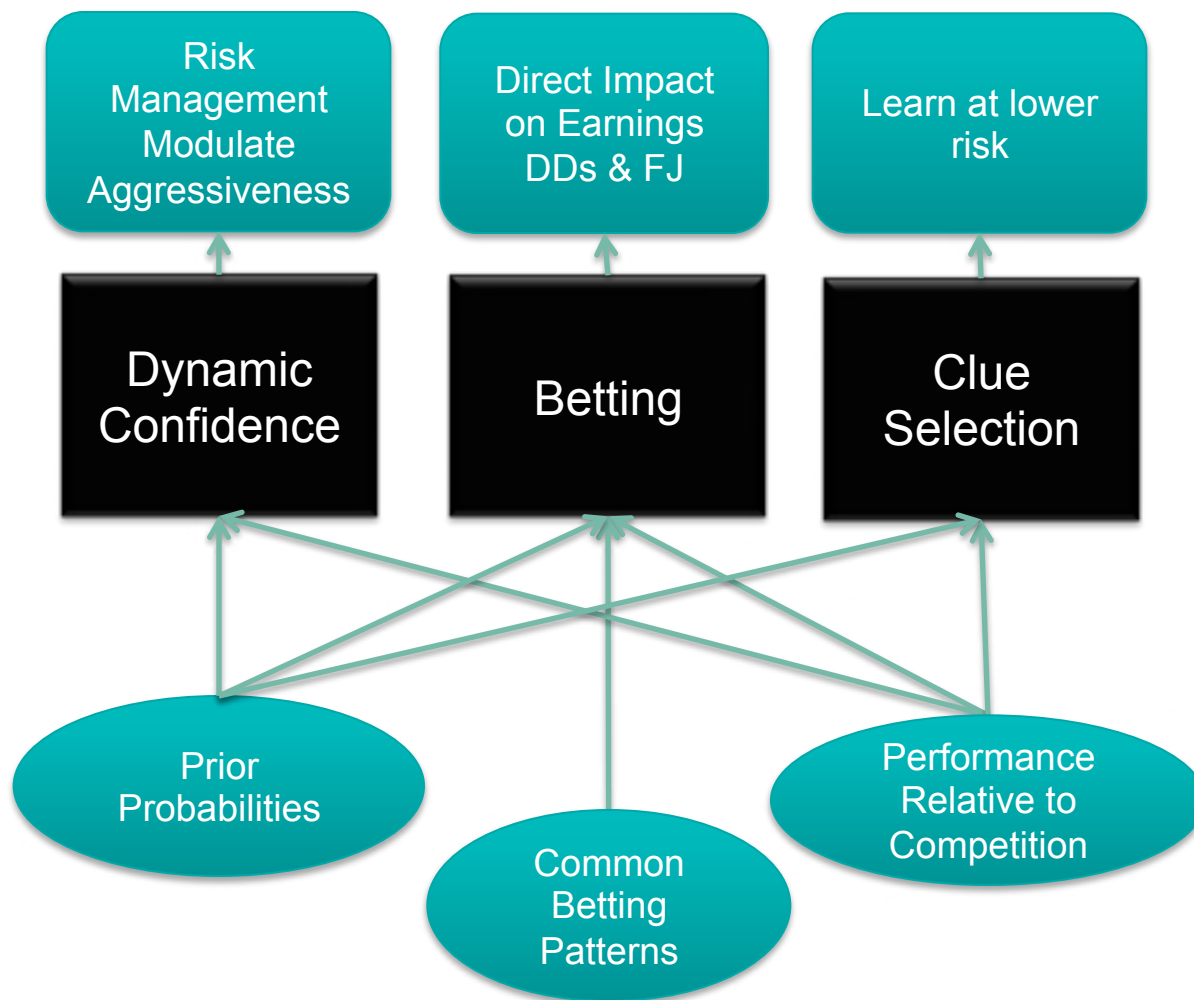
■ Saint Paul
■ South Bend

Humans may get this based on the pun since St. Paul since is a "holy" city. We added a Pun Scorer that discovers and scores Pun relationships.



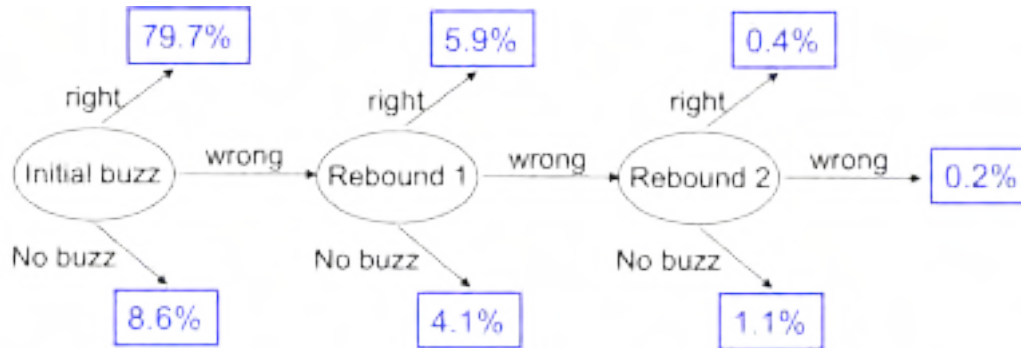
Game Strategy

Managing the Luck of the Draw



Buzz-in: Stochastic Process Model of Regular Clues

Statistics over 150K regular-clue (no DD) outcomes:



Using historical priors and game state information, Watson uses function approximation and Monte Carlo simulation to identify optimal buzzing for its confidence:

Buzz Threshold (Earnings-max)

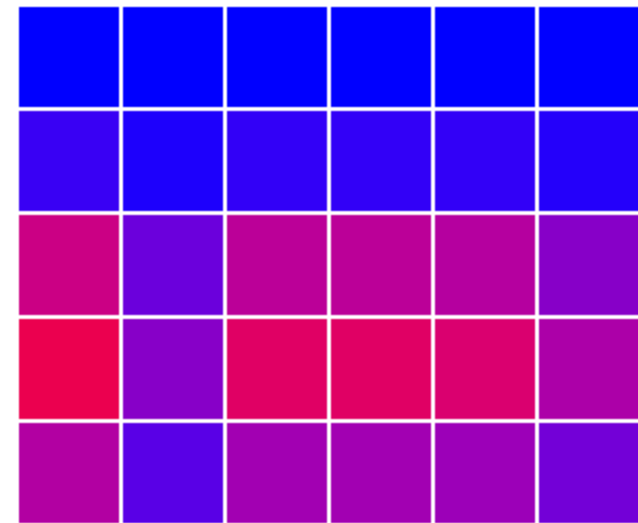


Buzz Threshold (Conservative)



Modeling Daily Double Placement

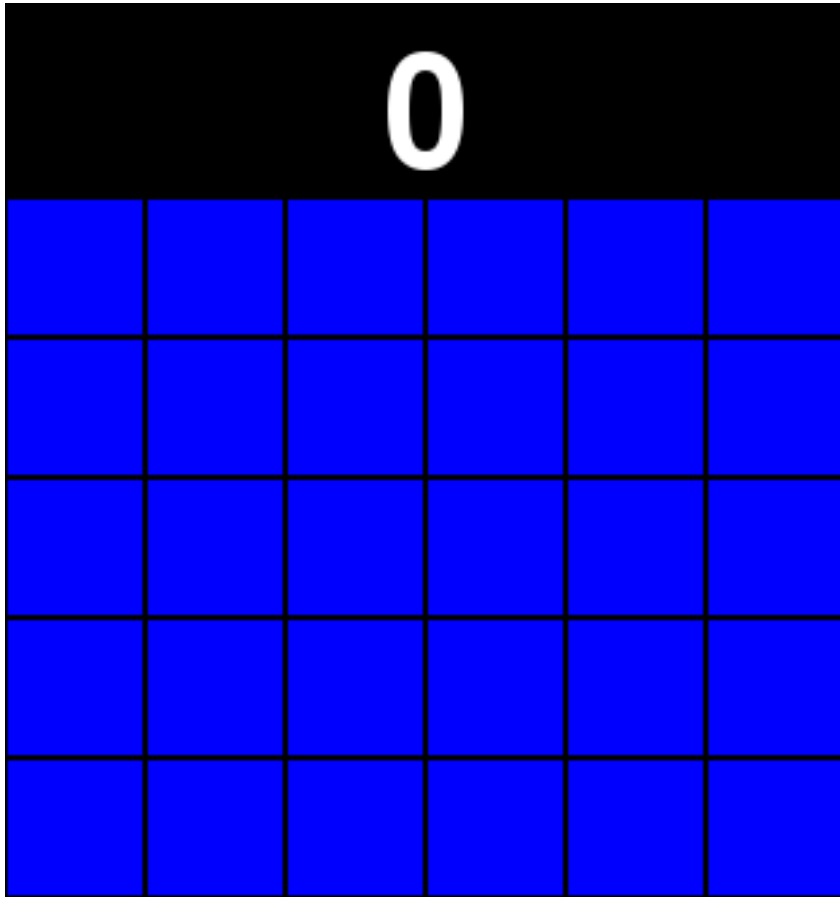
- ❑ **Statistics over 9k DDs** (3k Round1, 6k Round2):
 - (Widely known) DDs most frequent in the high-value rows (third, fourth, fifth) with harder clues
 - Row frequencies published on J! Archive
 - Some columns are more likely than others to have a DD!
 - First column most likely to have a DD
 - Second column least likely to have a DD
 - row-column frequencies used to randomly place DDs in simulated games; Watson uses them as
 - ❑ Bayesian prior
 - Columns are conditionally dependent! Watson performs a Bayesian update based on observed DDs



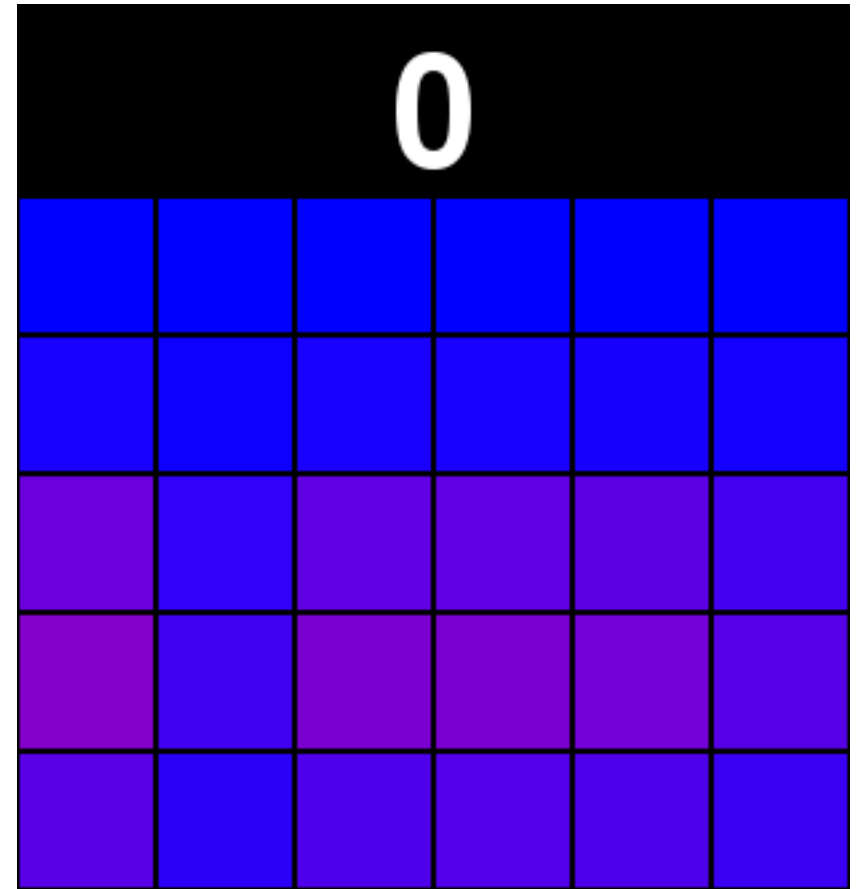
Unlikely Likely

Square Selection: Daily Double Prediction

Human



Watson

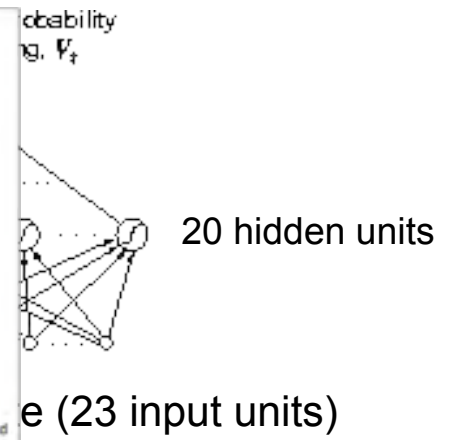
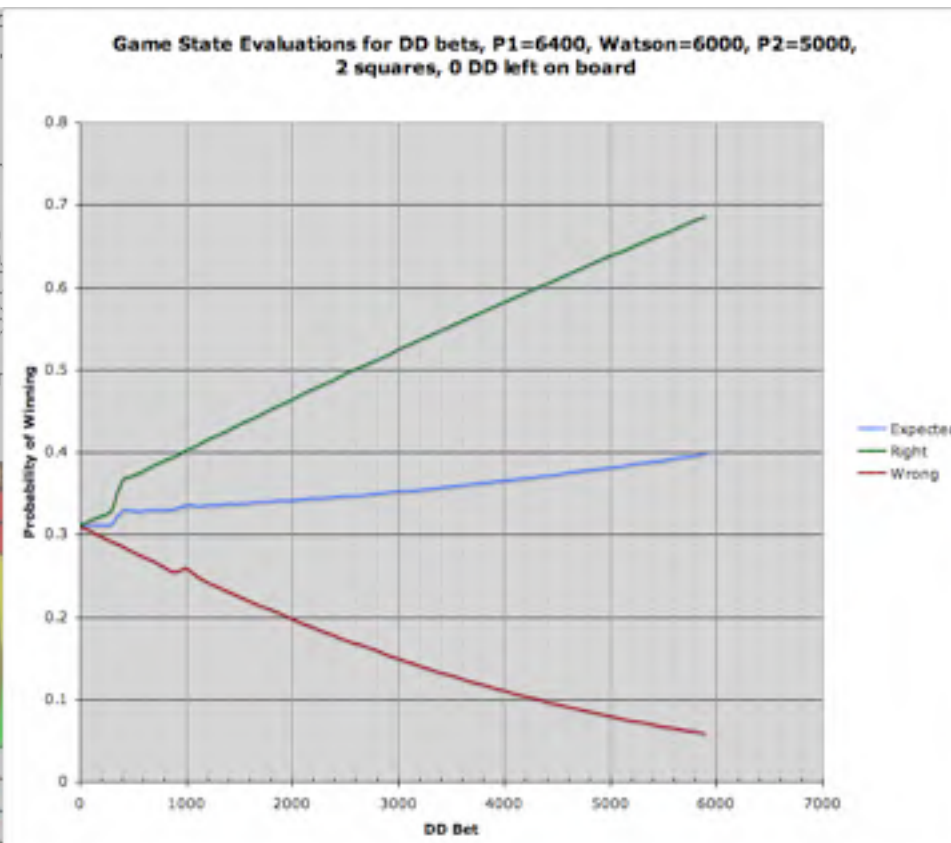
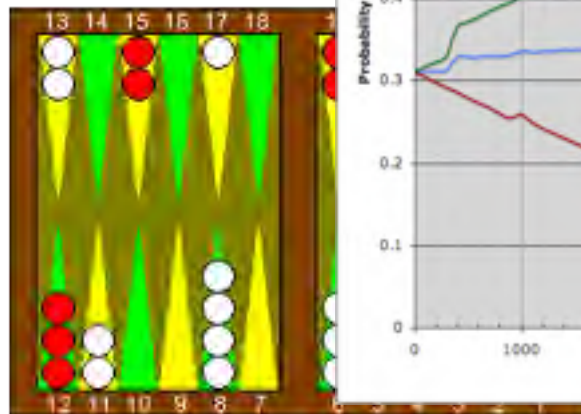
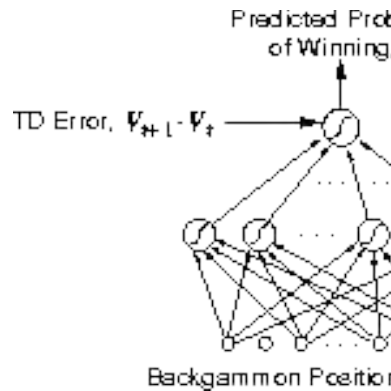


Unlikely

Likely

Daily Double Betting: Strategy

- Train an Artificial Neural Net over millions of simulated games pitting Watson vs. two simulated human opponents
- Use TD() reinforcement learning algorithm just as in TD-Gammon



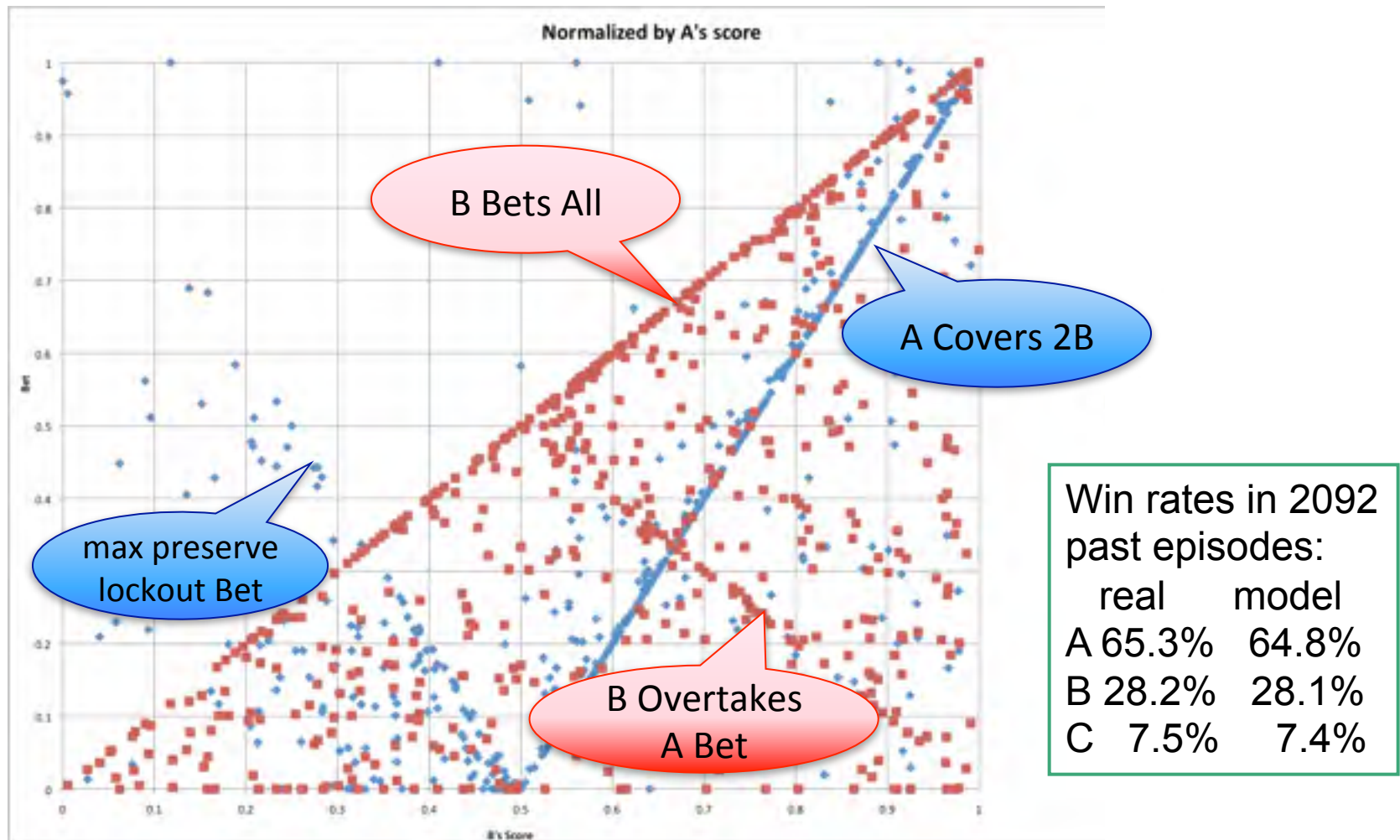
THE WHISKEY TRIAL	*ARD* STUFF	RHETT-ORX
\$200	\$200	\$200
\$400	\$400	\$400
\$600	\$600	\$600
\$800	\$800	\$800
\$1000	\$1000	

erry	Jon
\$1000	\$1000

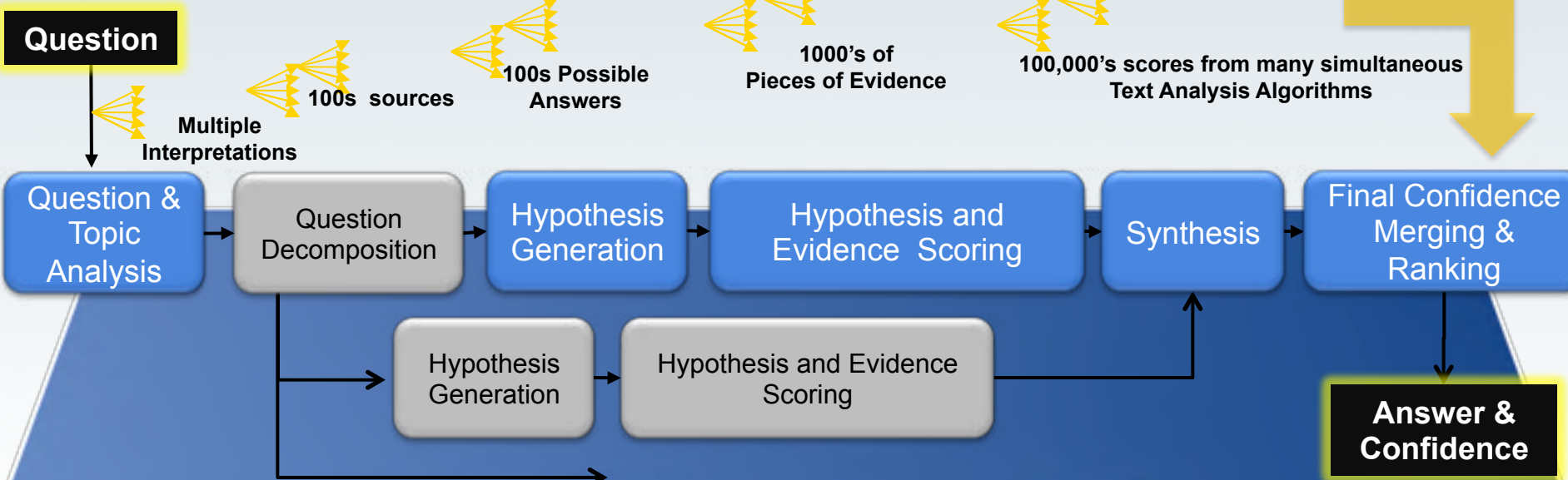
Final Jeopardy: Modeling Human Betting

Average FJ accuracy r **50%** FJ accuracy correlation r **0.3**

Bets depend on score positioning: 1st place (**"A"**), 2nd place (**"B"**), 3rd place (**"C"**)



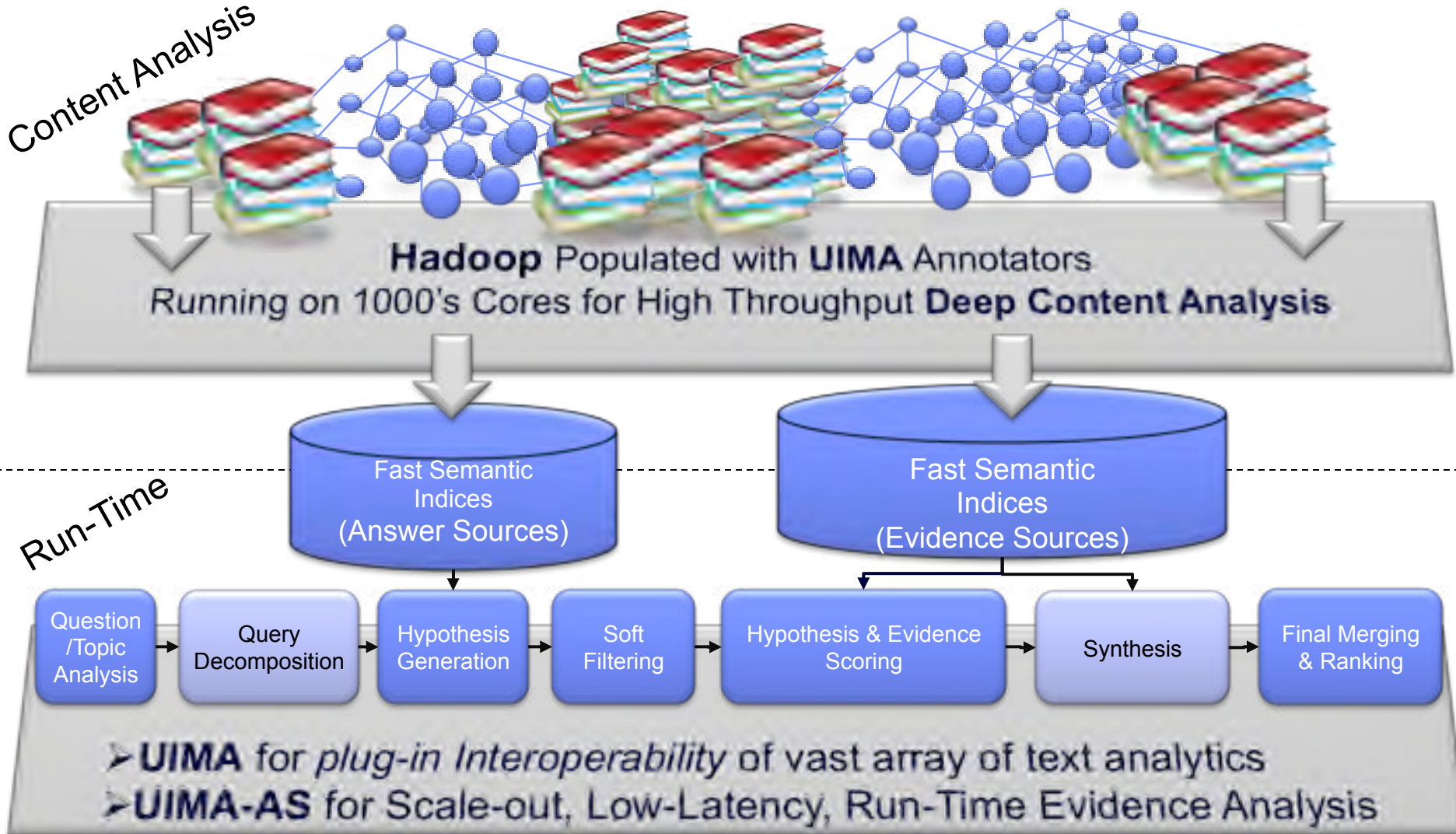
One Jeopardy! question can take 2 hours on a single 2.6Ghz Core
 Optimized & Scaled out on 2880-Core IBM workload optimized
 POWER7 HPC using UIMA-AS,
Watson answers in 2-6 seconds.



built on UIMA for interoperability

built on UIMA-AS for scale-out and speed

UIMA in Hadoop is used for high-throughput content analysis to generate fast semantic indices. These are used for DeepQA's rapid evidence evaluation in the context of the question at Run-Time



Watson – a Workload Optimized System

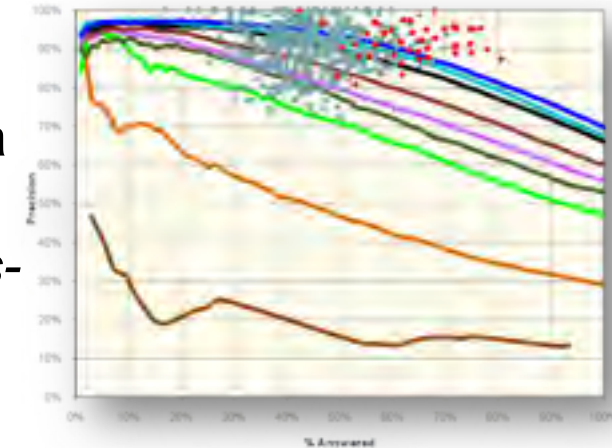
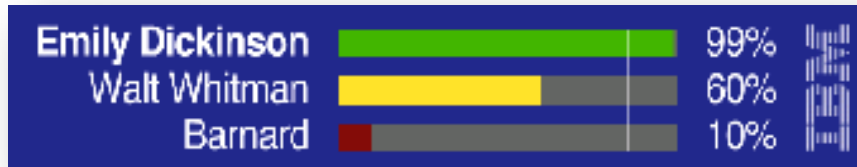
- 90 x IBM Power 750¹ servers
- 2880 POWER7 cores
- POWER7 3.55 GHz chip
- 500 GB per sec on-chip bandwidth
- 10 Gb Ethernet network
- 15 Terabytes of memory
- 20 Terabytes of disk, clustered
- Can operate at 80 Teraflops
- Runs IBM DeepQA software
- Scales out with and searches vast amounts of unstructured information with UIMA & Hadoop open source components
- Linux provides a scalable, open platform, optimized to exploit POWER7 performance
- 10 racks include servers, networking, shared disk system, cluster controllers



¹ Note that the Power 750 featuring POWER7 is a commercially available server that runs AIX, IBM i and Linux and has been in market since Feb 2011

Precision, Confidence & Speed

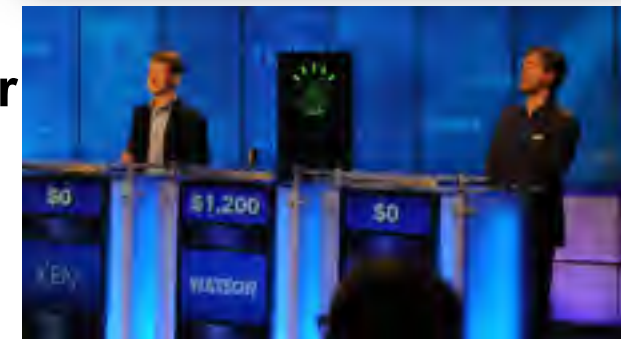
- Deep Analytics** – Combining many analytics in a novel architecture, we achieved very high levels of **Precision** and **Confidence** over a huge variety of *as-is* content.



- Speed** – By optimizing Watson's computation for Jeopardy! on over **2,800 POWER7** processing cores we went **from 2 hours per question on a single CPU to an average of just 3 seconds.**



- Results** – in 55 real-time sparring games against former **Tournament of Champion Players last year** Watson put on a very competitive performance in all games -- placing 1st in 71% of the them!



Potential Business Applications



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

Tech Support: Help-desk, Contact Centers



Enterprise Knowledge Management and Business Intelligence

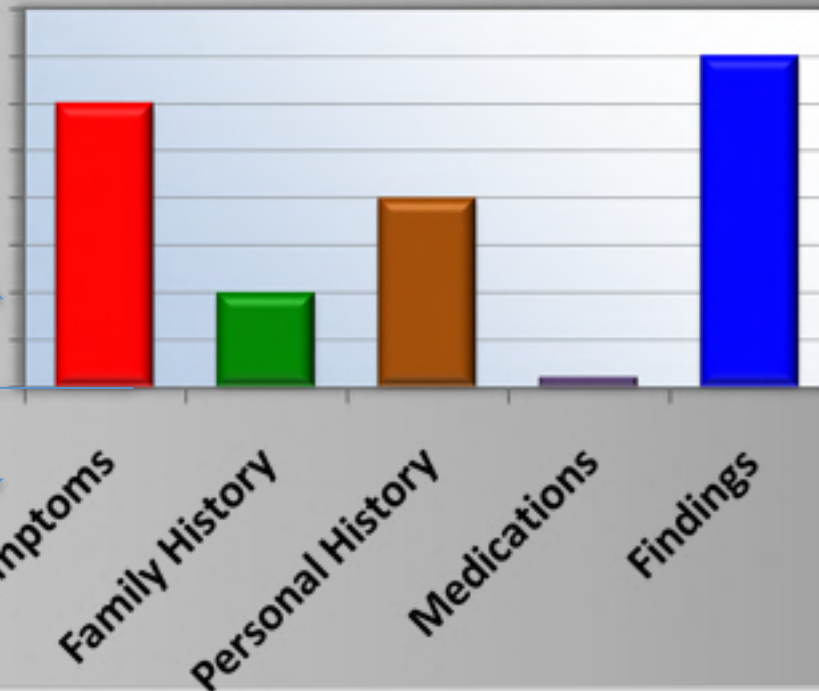
Government: Improved Information Sharing and Security



Evidence Profiles from disparate data is a powerful idea

- Each dimension contributes to supporting or refuting hypotheses based on
 - **Strength of evidence**
 - **Importance of dimension for diagnosis** (learned from training data)
- Evidence dimensions are combined to produce an overall confidence

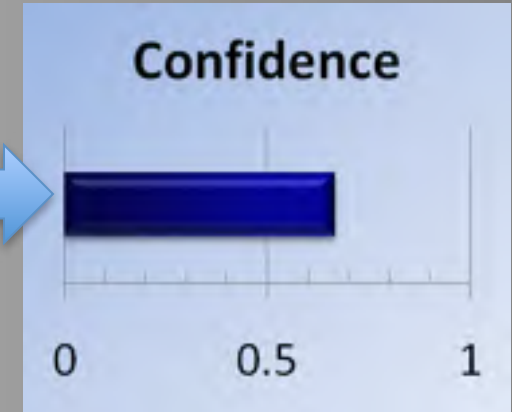
Evidence Profile for UTI Diagnosis



Positive Evidence

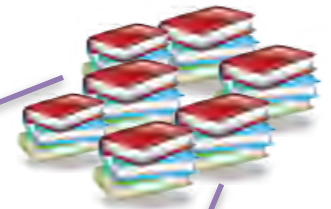
Negative Evidence

Overall Confidence



Mapping from Language to Language

Flexible Matching Features allows us to gather evidence that Dysphagia is a element of the diagnosis

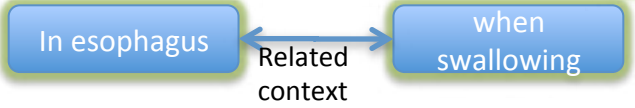
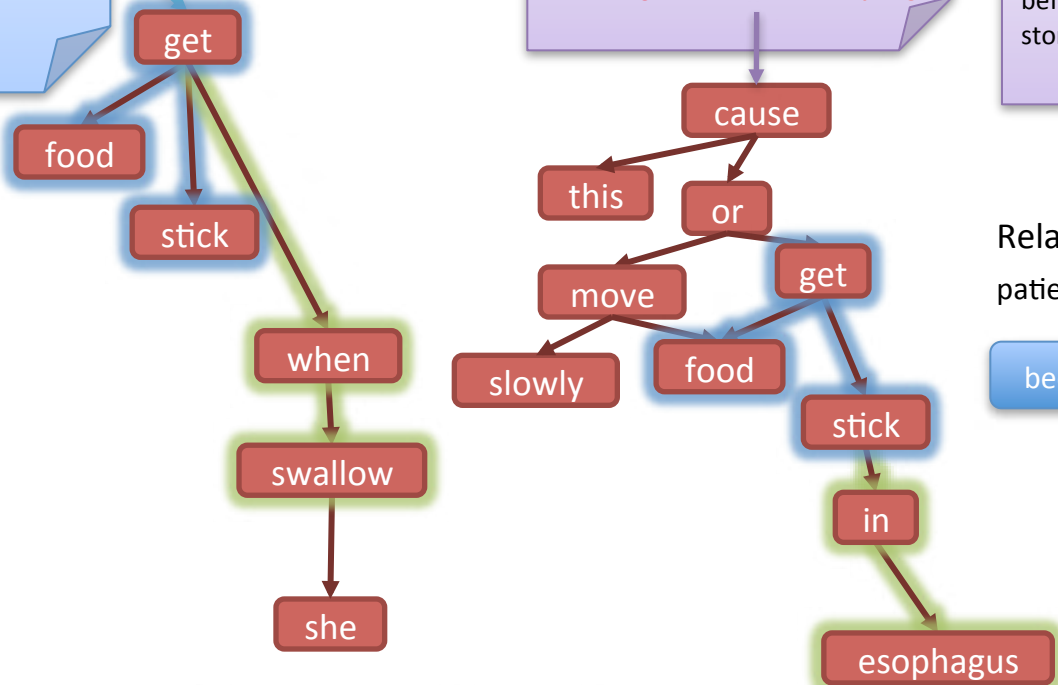


A 58-year-old woman presented to her primary care physician after several days of **dizziness**, **anorexia**, **dry mouth**, **increased thirst**, and **frequent urination**. She had also had a **fever** and reported that **food would "get stuck" when she was swallowing**.

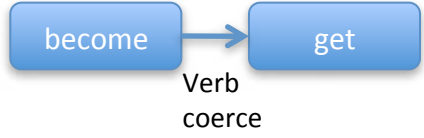
She reported no **pain** in her abdomen, back, or flank and no **cough**, **shortness of breath**, **diarrhea**, or **dysuria**.

Dysphagia
These disorders can stop the nerves and muscles in your esophagus (the tube that runs from your mouth and throat down to your stomach) from working right. This can **cause food to move slowly or even get stuck in the esophagus**.

Dysphagia
The most common symptom of esophageal dysphagia is the inability to swallow solid food, which the patient will describe as '**becoming stuck**' or 'held up' before it either passes into the stomach or is regurgitated.



Relation Detection:
patientDesc("becoming stuck")





Samantha Darren

Profile Card

Dashboard

Factors

SYMPTOMS

- Red eye CC NEW
- Eye pain CC NEW
- Eye inflammation CC NEW
- Blurred vision NEW
- Floating spots NEW
- Sensitivity to light NEW

FAMILY HISTORY

Arthritis

ALLERGIES

Penicillin

DEMOGRAPHICS

34 years old

Female

SCRATCHPAD

Diagnosis

Treatment

Ask Watson

Uveitis	<div style="width: 90%;"></div>	90%
Iritis	<div style="width: 50%;"></div>	50%
Keratitis	<div style="width: 42%;"></div>	42%
Anterior Uveitis	<div style="width: 31%;"></div>	31%
Photophobia	<div style="width: 27%;"></div>	27%
Behcet's Disease	<div style="width: 13%;"></div>	13%
Rhumatoid Arthritis	<div style="width: 11%;"></div>	11%
Ankylosing Spondylitis	<div style="width: 10%;"></div>	10%
Sarcoidosis	<div style="width: 4%;"></div>	4%

Evidence

Consider



Samantha Darren

IBM WATSON

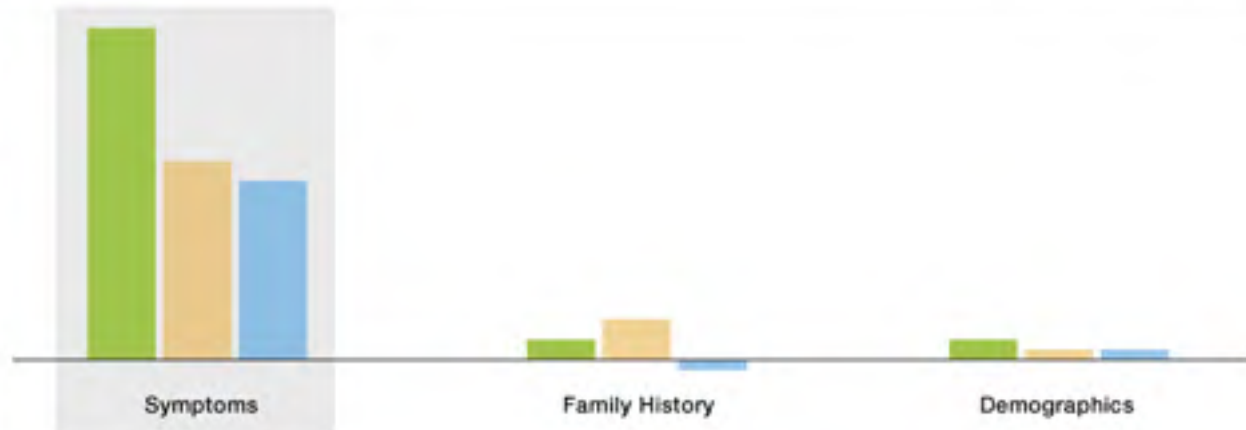
Profile Card

Dashboard

Evidence Profile: Diagnosis

Back to Results

- Uveitis 91%
- Iritis 48%
- Keratitis 42%



Sources

Factors

Textbook

Uveitis specifically refers to **inflammation** of the middle layer of the eye, termed the "uvea" but in common usage may refer to any inflammatory process involving the interior of the eye.



Website

Uveitis usually develops quickly and generally affects only one eye. Signs and symptoms may include any or all of the following. **pain in the eye** or brow region, worsened eye **pain when exposed to bright light**, and **reddened eye**, especially adjacent to the iris.



Medical Journal

The most common form of **uveitis** is anterior **uveitis**, which involves **inflammation** in the front part of the eye. It is often called **iritis** because it usually only affects the iris, the colored part of the eye.



Textbook

Uveitis is often marked by moderate to intense **pain** and usually involves **impaired eyesight**.



SONY PICTURES TELEVISION, IBM, and THE GHOST OF MERV GRIFFIN PRESENT

THIS IS JEOPARDY!

AN UNPRECEDENTED MAIN EVENT:

THREE NIGHTS ONLY!

FEB

14TH 15TH 16TH

CHECK YOUR
LOCAL LISTINGS

MAN VS MAN VS MACHINE

★ ★ ★ ★ TONIGHT'S ★ ★ ★ ★
CONTESTANTS

KEN "THE BLOCK" JENNINGS



"HULLIGAN-DOLLAR" BRAD RUTLER



WATSON

10202 WEST WASHINGTON BLVD
CULVER CITY, CALIFORNIA

Watson: Always Improving..

NEW YORK TIMES HEADLINES: An exclamation point was warranted for the "end of" this! In 1918

A sentence (.48)

WORLD FACTS: The Denmark Strait separates these 2 islands by about 200 miles

Greenland and Taiwan (0.31)

THE AMERICAN DREAM: Decades before Lincoln, Daniel Webster spoke of government "made for", "made by" & "answerable to" them

No one (0.29)

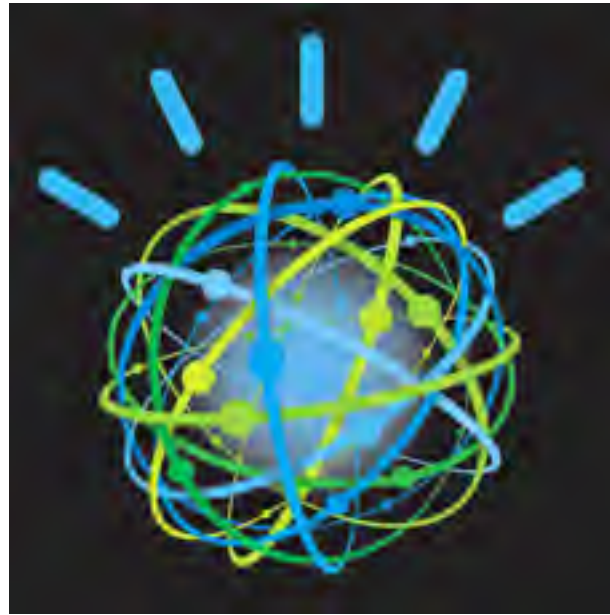


THE QUEEN'S ENGLISH : Give a Brit a tinkle when you get into town & you've done this

Urinate (.56)

BOTTOMS UP!: Often served at a brunch, it's made with equal amounts of champagne and orange juice

Breakfast (0.11)



THANK YOU