



Association for
Computing Machinery

Advancing Computing as a Science & Profession

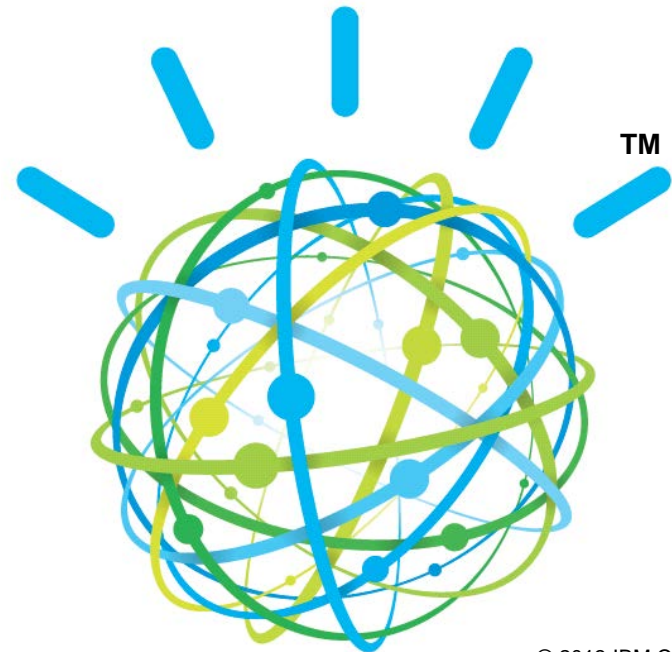


IBM Watson™: Beyond Jeopardy!™

Adam Lally

Senior Technical Staff Member

IBM T.J. Watson Research Center



<http://learning.acm.org>

- **1,350+ trusted technical books and videos by leading publishers including O'Reilly, Morgan Kaufmann, others**
- **Online courses with assessments and certification-track mentoring, member discounts on tuition at partner institutions**
- **Learning Webinars on big topics (Cloud Computing/Mobile Development, Cybersecurity, Big Data, Recommender Systems, SaaS, Agile)**
- **ACM Tech Packs on big current computing topics: Annotated Bibliographies compiled by subject experts**
- **Popular video tutorials/keynotes from ACM Digital Library, A.M. Turing Centenary talks/panels**
- **Podcasts with industry leaders/award winners**

- Use the **Facebook** widget in the bottom panel to share this presentation with friends and colleagues
- Use **Twitter** widget to Tweet your favorite quotes from today's presentation with hashtag **#ACMWebinarLally**
- Submit questions and comments via Twitter to **@acmeducation** – we're reading them!

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



The Jeopardy! Challenge: *A compelling and notable way to **drive** and **measure** the technology of automatic Question Answering and Natural Language Understanding*

**Broad/Open
Domain**

**Complex
Language**

**High
Precision**

**Accurate
Confidence**

**High
Speed**

\$200

If you're standing, it's the direction you should look to check out the wainscoting.

\$1000

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

\$600

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

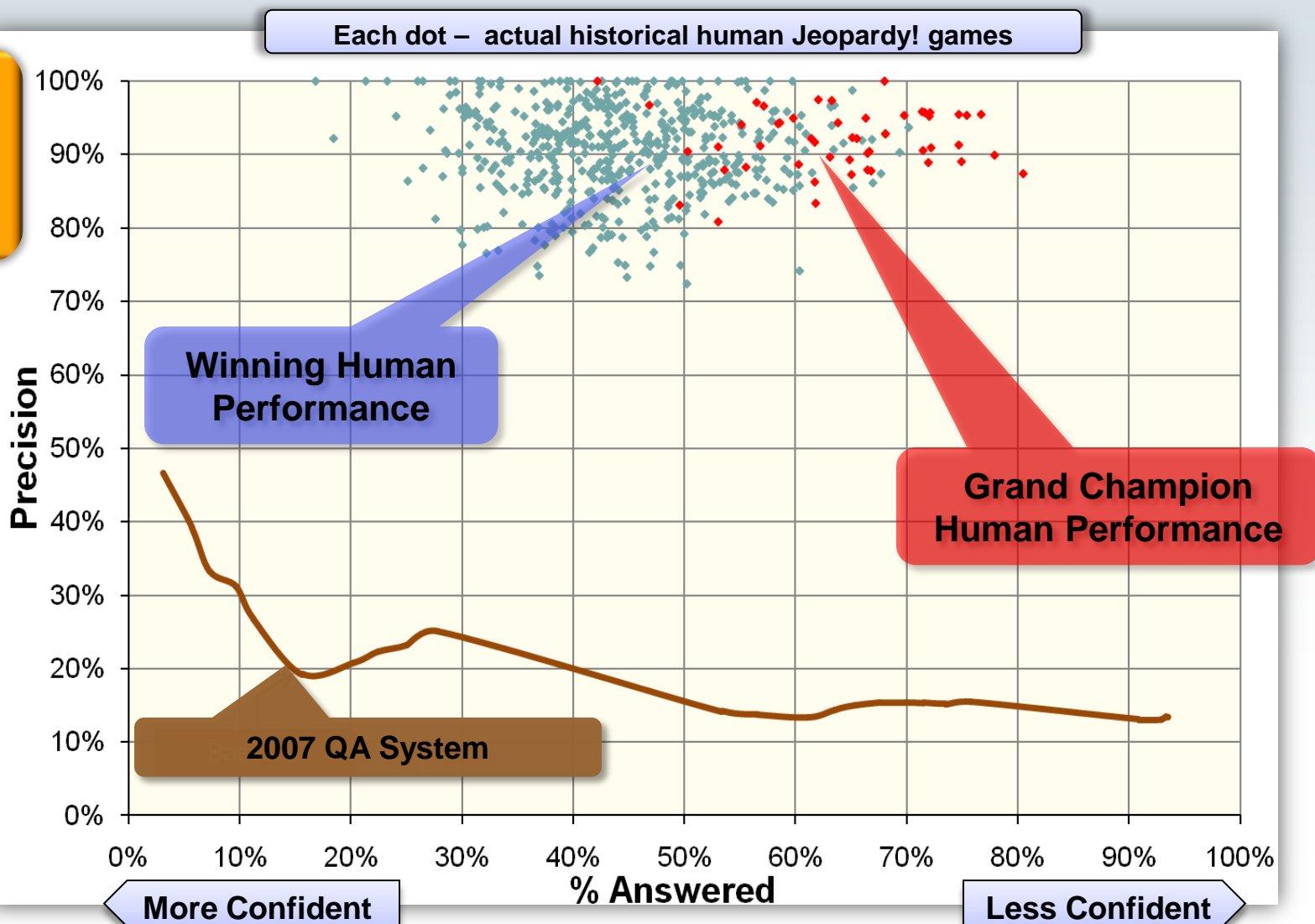
\$2000

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

Task Performance Metrics – Confidence and Precision

The Winner's Cloud

Top human players are remarkably good.



Watson answers by finding, reading, scoring and combining evidence

**IN 1698, THIS COMET
DISCOVERER TOOK A
SHIP CALLED THE
PARAMOUR PINK ON THE
FIRST PURELY
SCIENTIFIC SEA VOYAGE**

**Question
Analysis**

Important Terms: 1698, comet,
paramour, pink, ...
AnswerTypes: comet discoverer
Date(1698),
Took(discoverer, ship)
Called(ship, Paramour Pink)
...

**Content
(Structured & Unstructured)**

**Primary
Search**

Candidate Answer/Hypothesis Generation

**High-Speed
Evidence
Retrieval**

**Term Overlap
Classification
Relations
...
Temporal**

**100's of NLP Scoring
Algorithms**

Isaac Newton
Wilhelm Tempel
HMS Paramour
Christiaan Huygens
Halley's Comet
Edmond Halley
Pink Panther
Peter Sellers

[0.58	0.5	-1.3	...	0.97]
[0.71	1	13.4	...	0.60]
[0.42	0	2.0	...	0.90]
[0.84	0.5	10.6	...	0.88]
[0.33	0	6.3	...	0.83]
[0.21	1	11.1	...	0.92]
[0.91	0	-8.2	...	0.31]
[0.91	0	-1.7	...	-.20]

- 1) **Edmond Halley (0.85)**
- 2) **Christiaan Huygens (0.20)**
- 3) **Peter Sellers (0.05)**
- 4) ...

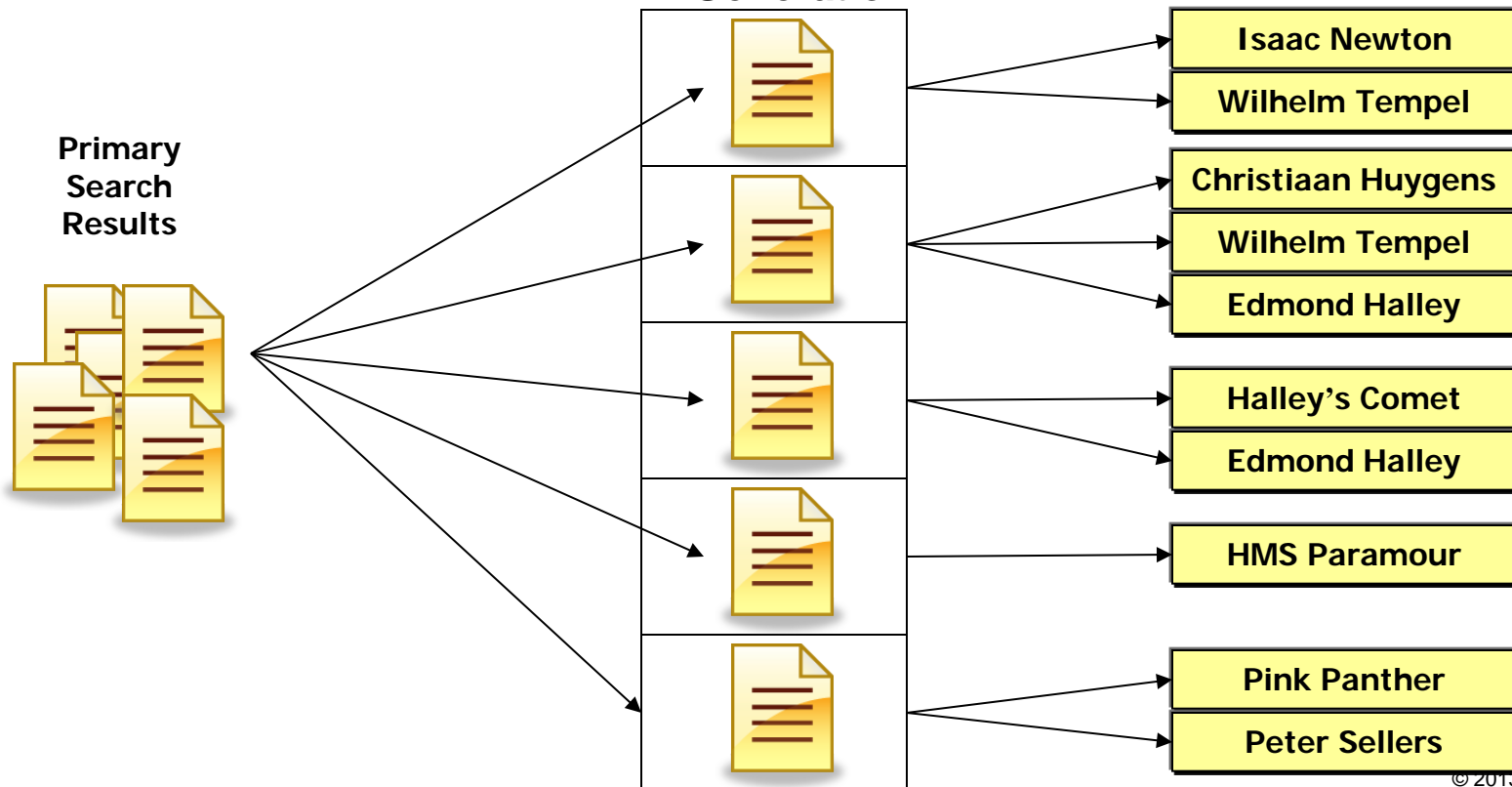
**Merging &
Ranking Based on
Statistical Machine
Learning**

**Diverse and Extensible Evidence
Scoring**

Each Primary Search Result is analyzed independently, in parallel, to generate candidate answers



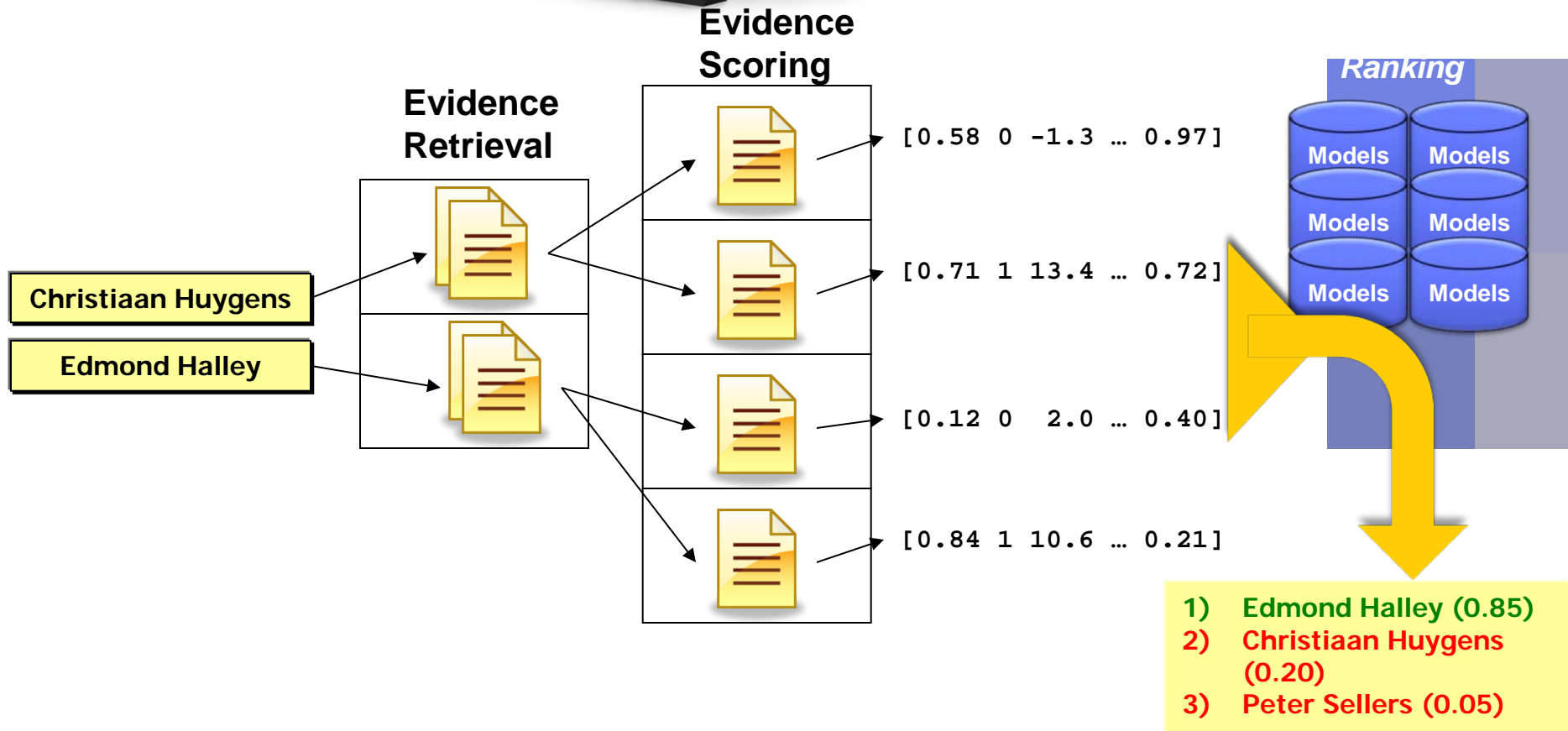
Candidate Generation



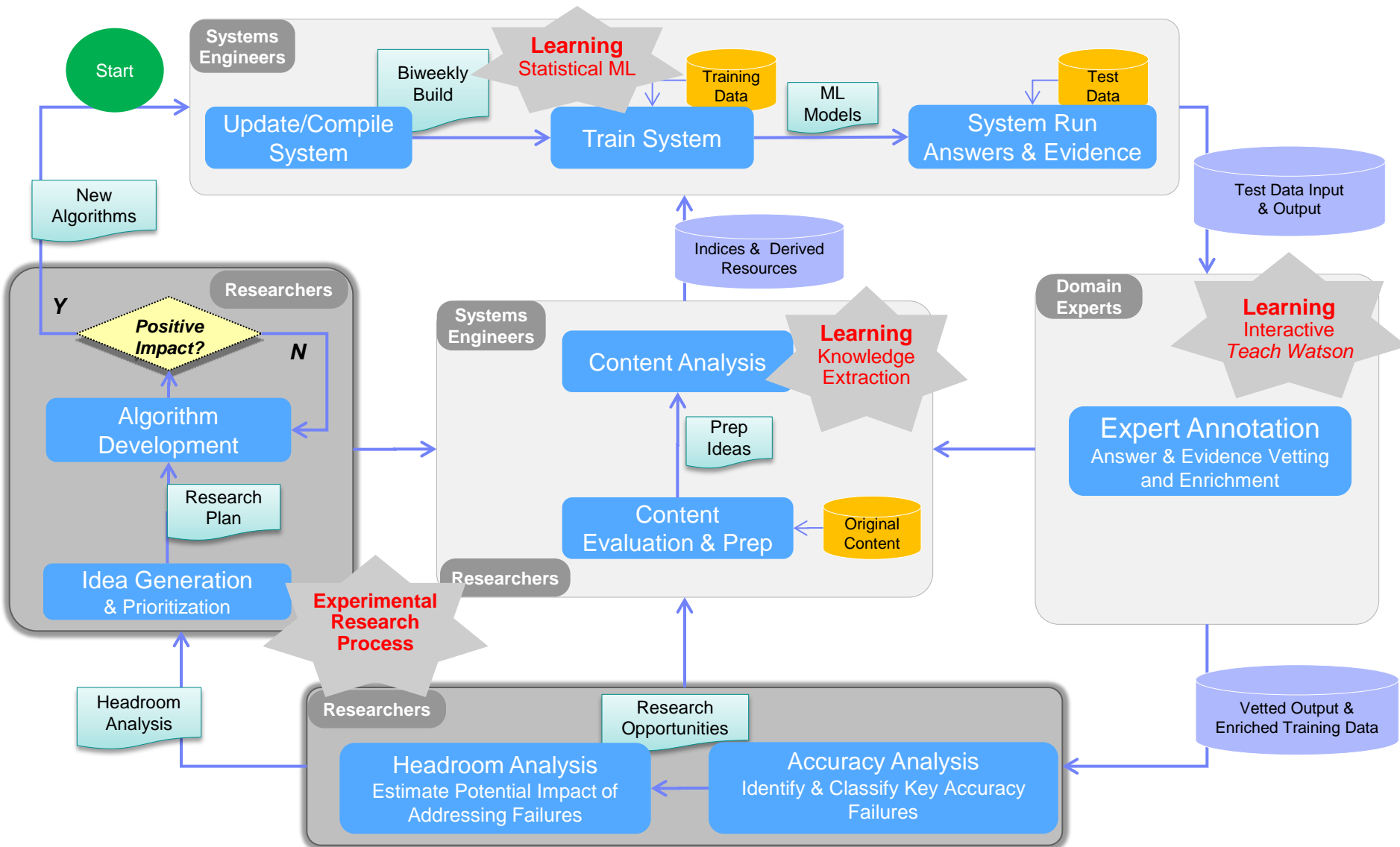
Evidence is Retrieved for each Candidate Answer in parallel



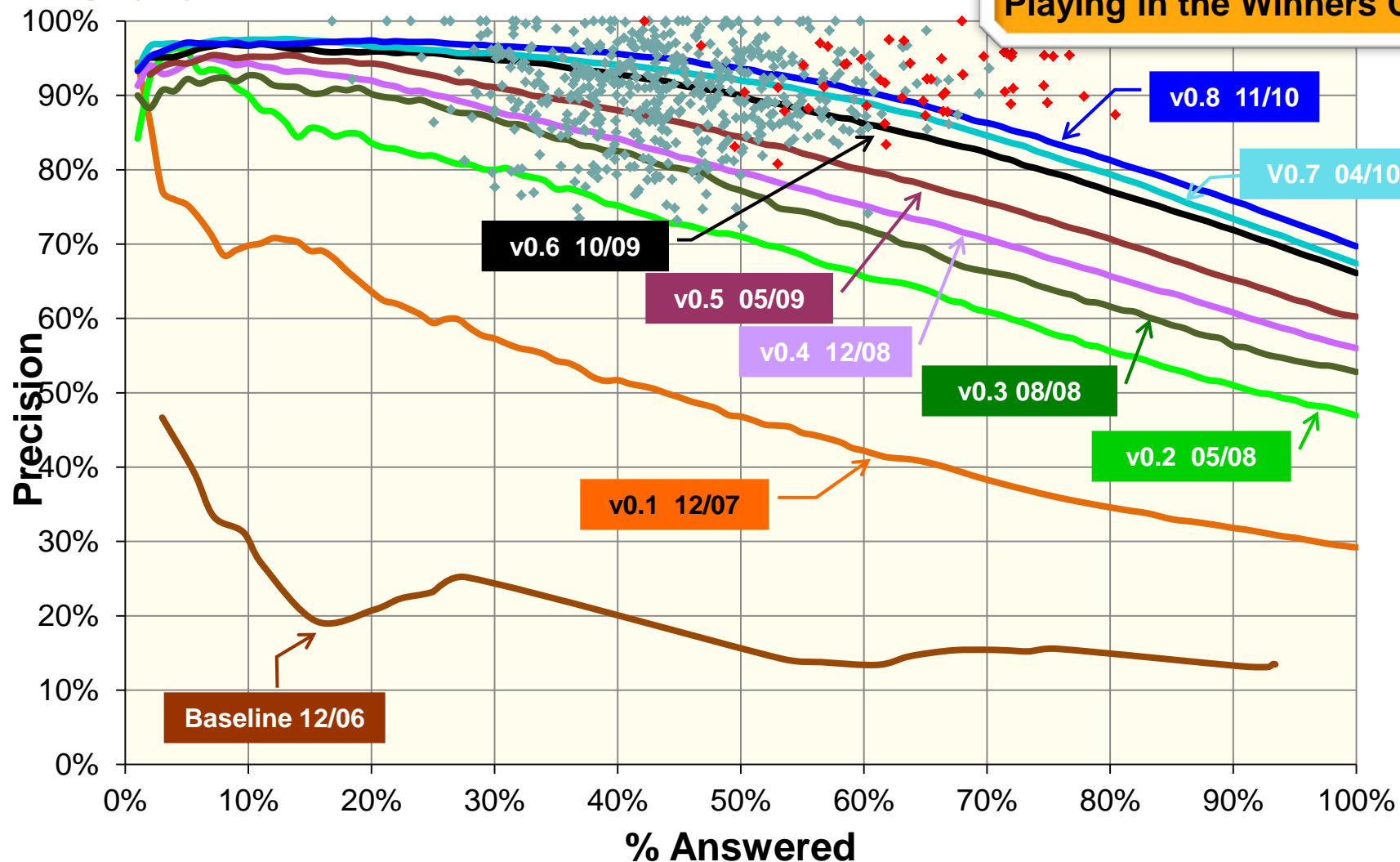
Each Piece of Evidence is scored in parallel



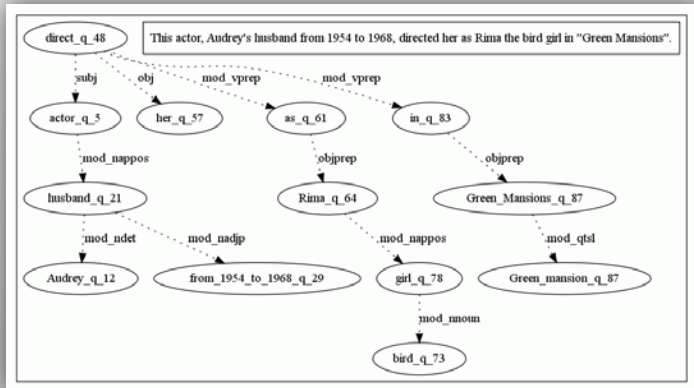
Domain and Language Adaptation Methodology



DeepQA: Incremental Progress in Answering Precision on the Jeopardy! Challenge: 6/2007 11/2010

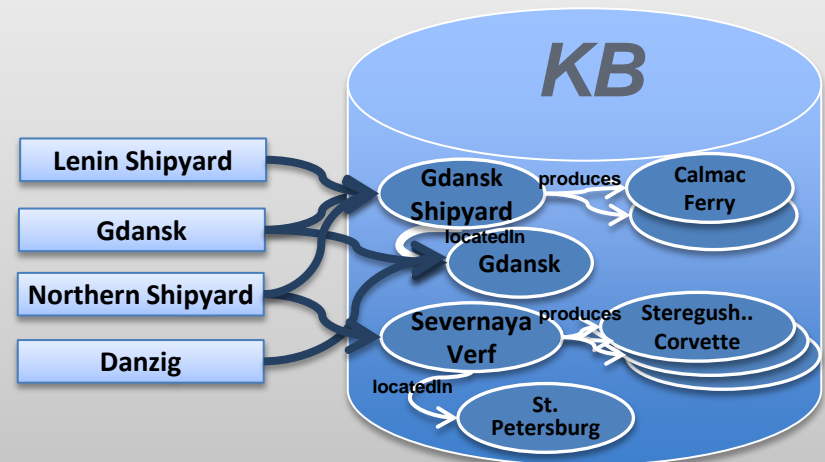


Question Processing



Dependency parse, Focus/LAT detection: 6000 rules, Decomp.
Eval on Jeopardy!: Parser: 92.4% acc, LAT detection: 96.8%

KAFE: Knowledge From Extracted Content



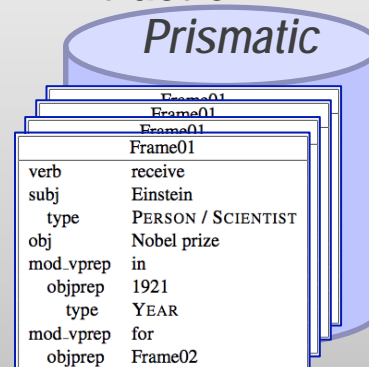
Entity Disambiguation, Entity Typing, Type Disambiguation
Eval on Wikipedia Disambig Task, F-score 92.5

Relation Extraction



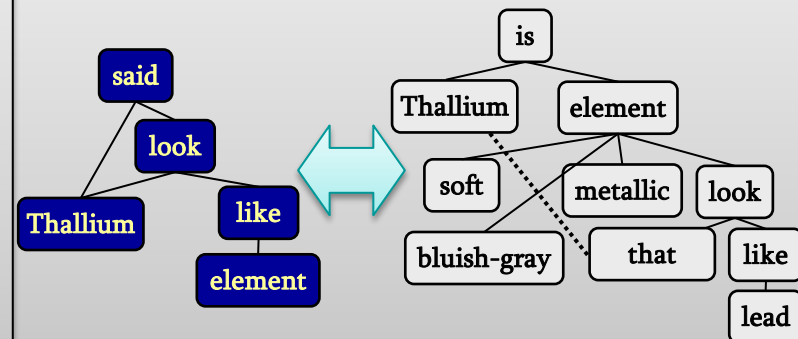
Semantic Relation Repository 7000 rels
Eval on ACE: F-score 73.2 (leading score)

Linguistic Frame Extraction



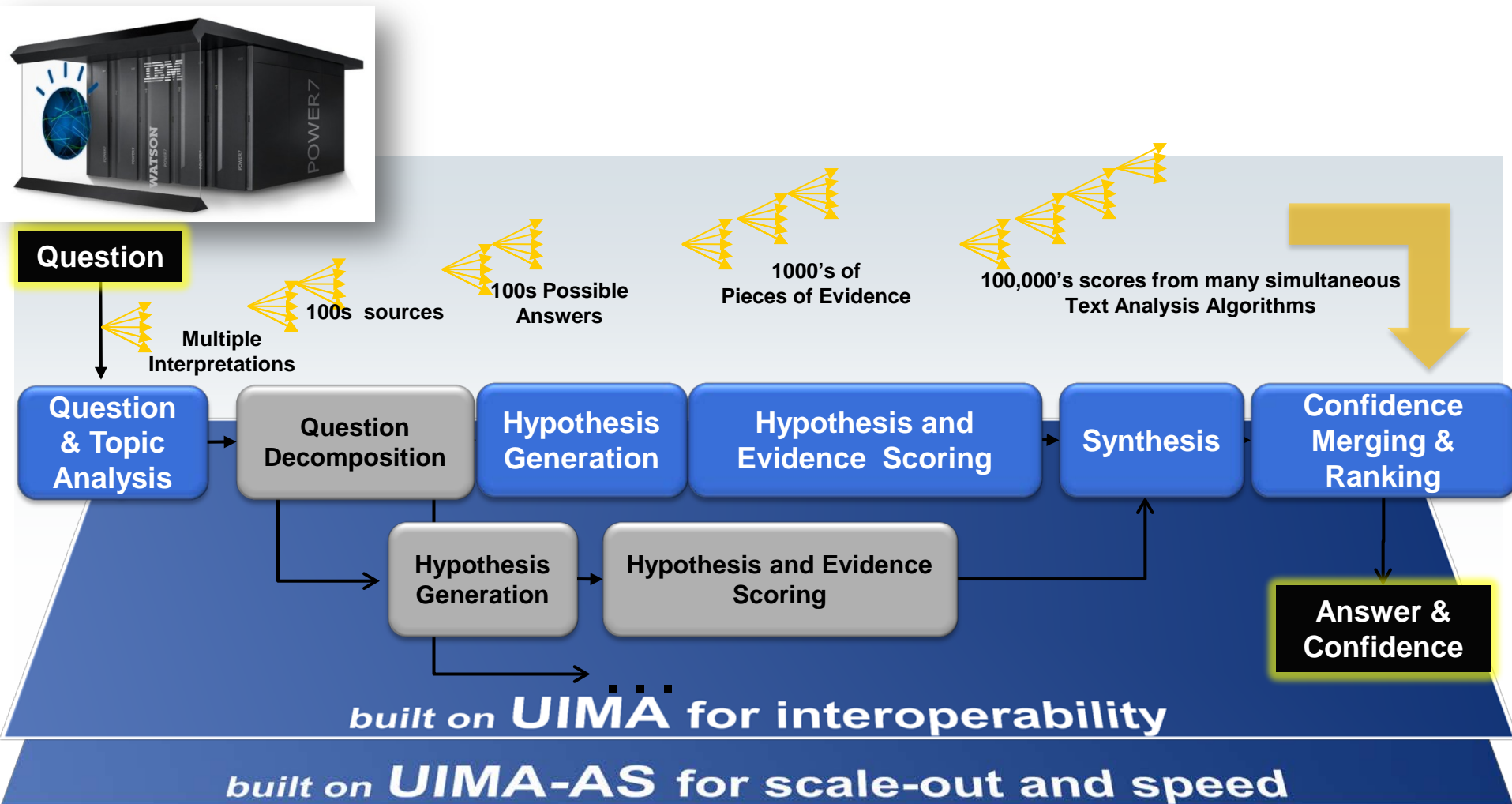
>1 Billion Frames. Mining from ClueWeb:
SVO/isa/etc. cuts, Intensional/Extensional
representation

Passage Matching Ensemble



Synonymy, Temporal/Geographic Reasoning, Linguistic Axioms
Eval on RTE 2010 Text Entailment: F-score 48.8 (leading score)

Workload Optimization: One Jeopardy! question could take 2 hours on a single 2.6Ghz Core Optimized & Scaled out on 2880-Core IBM Power750's using UIMA-AS, *Watson* is answering in 2-6 seconds.

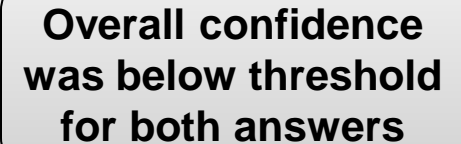


With Precision, Accurate Confidence and Speed, the rest was History



US CITIES

Its largest airport is named for a World War II hero; its second largest, for a World War II battle



**Low because
being a US City is not a strong
requirement simply based on
Jeopardy! category**

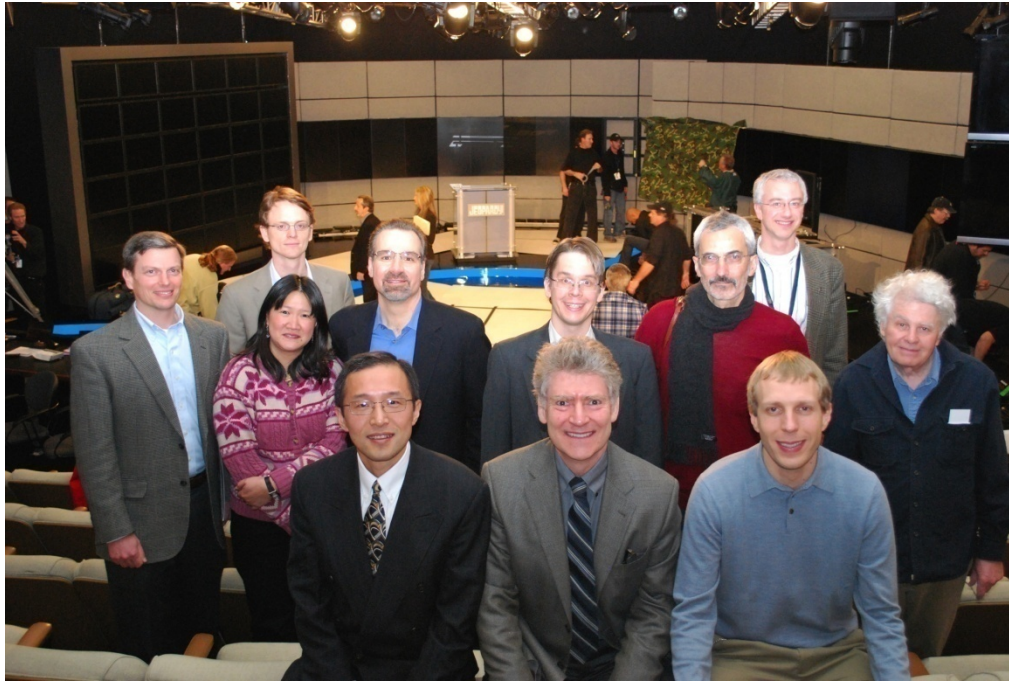
Toronto	<div><div></div></div>	14%
Chicago	<div><div></div></div>	11%
Omaha	<div><div></div></div>	10%

When being a **US City** is a more trusted requirement, **Chicago** is in first position but still weak, because Watson is not convinced about the WWII evidence it could find.

US CITIES

This US City's largest airport is named for a World War II hero; its second largest, for a World War II battle



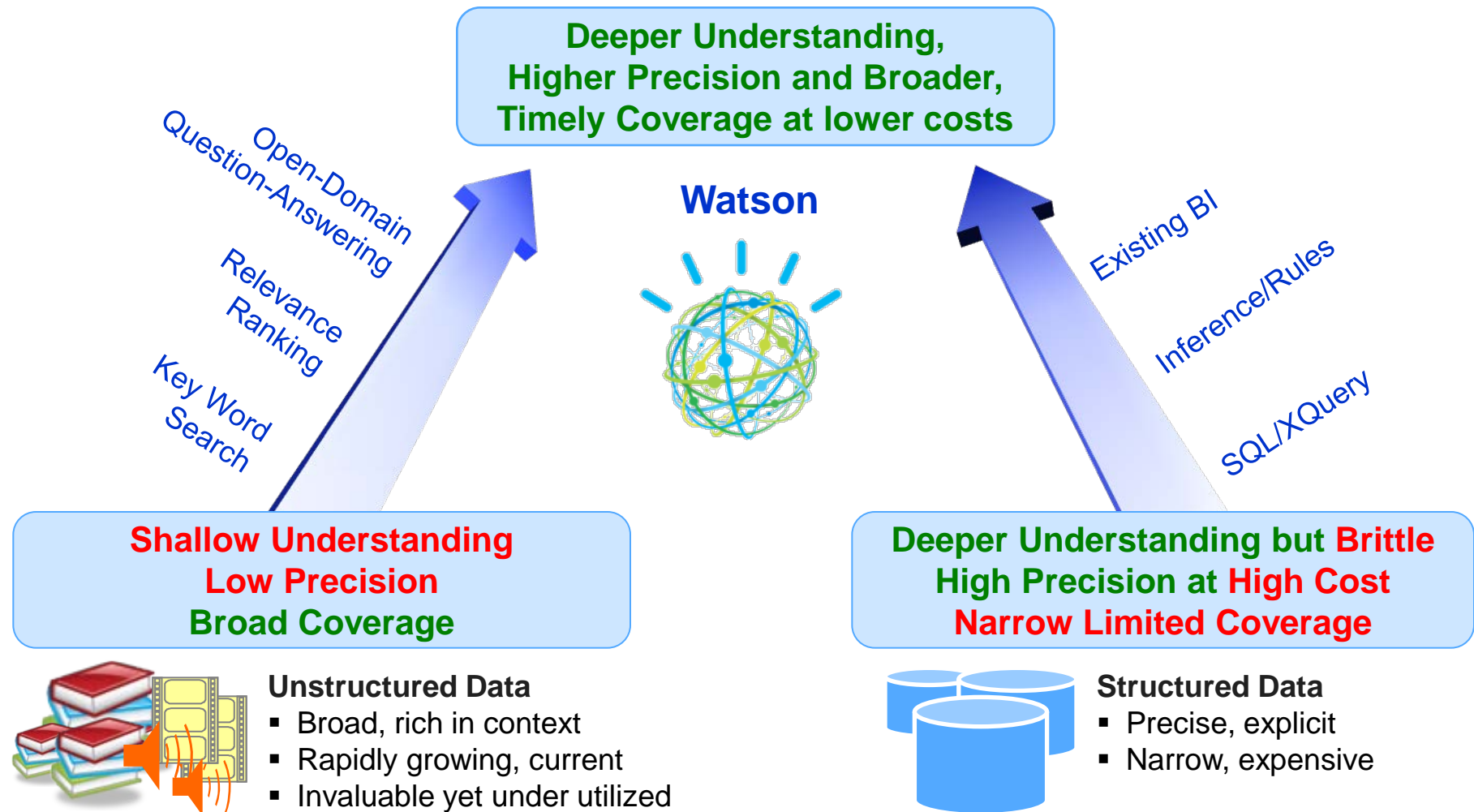


4 years
12 -> 30 people

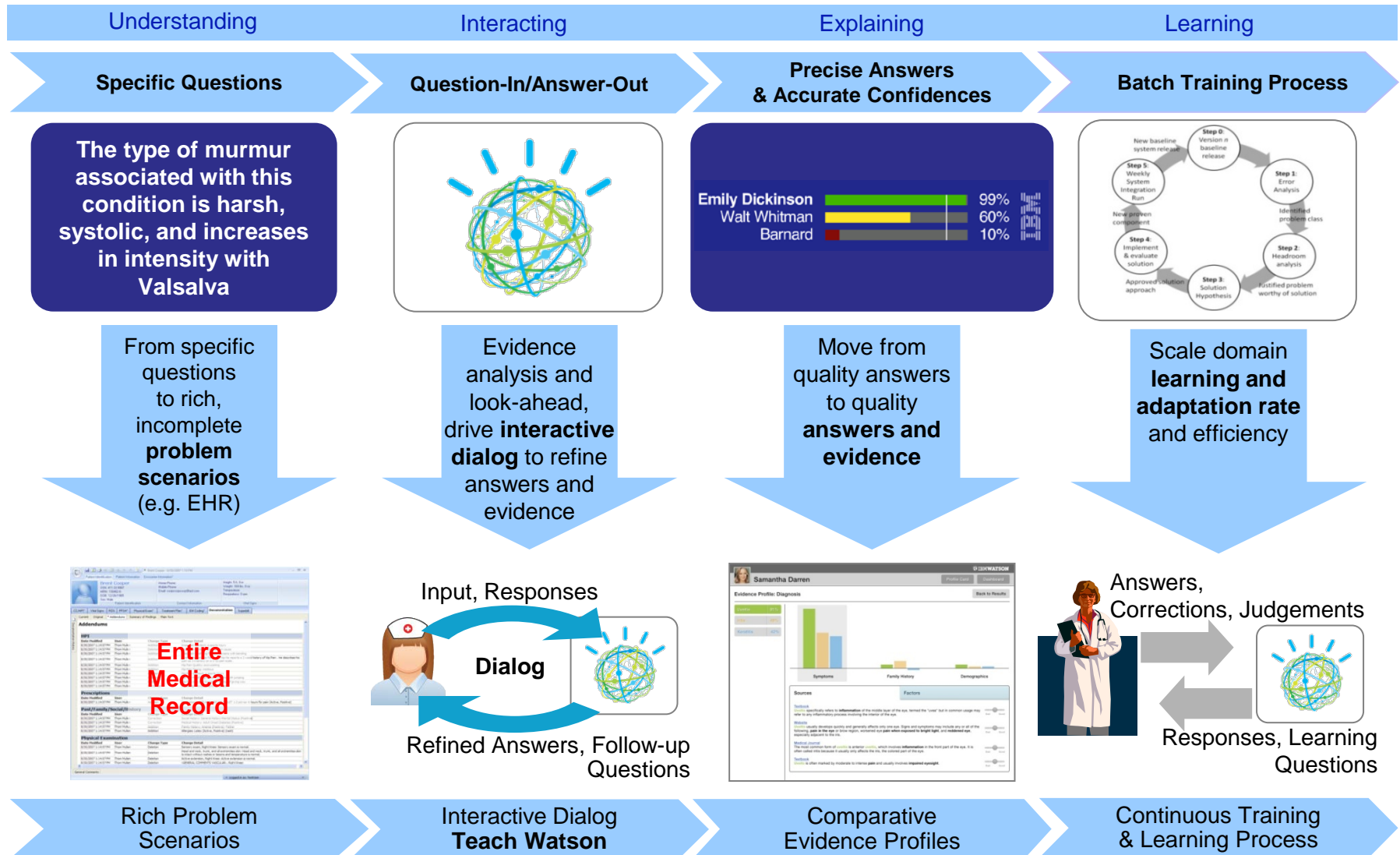


The Next Step: Beyond Jeopardy!

Watson's Real Value Proposition: Deeper Insight into Larger Volumes of Unstructured Content



Taking Watson beyond Jeopardy!



Adapting to a New Task

American College of Physicians' Doctors Dilemma Questions

The syndrome characterized by narrowing of the extra-hepatic bile duct from mechanical compression by a gallstone impacted in the cystic duct

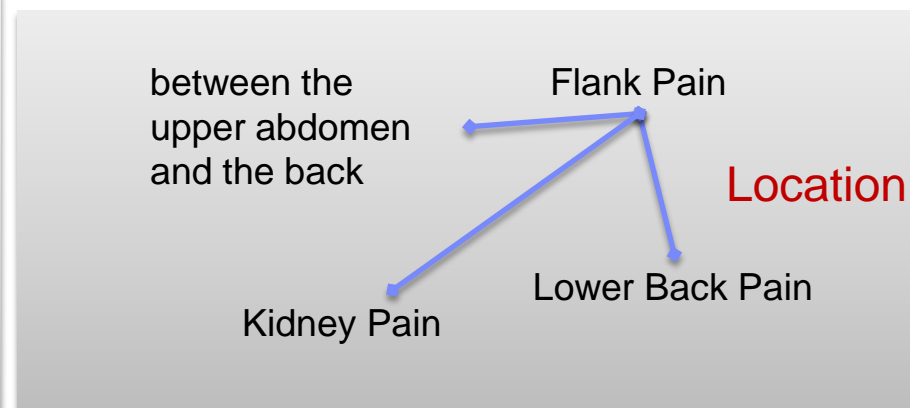
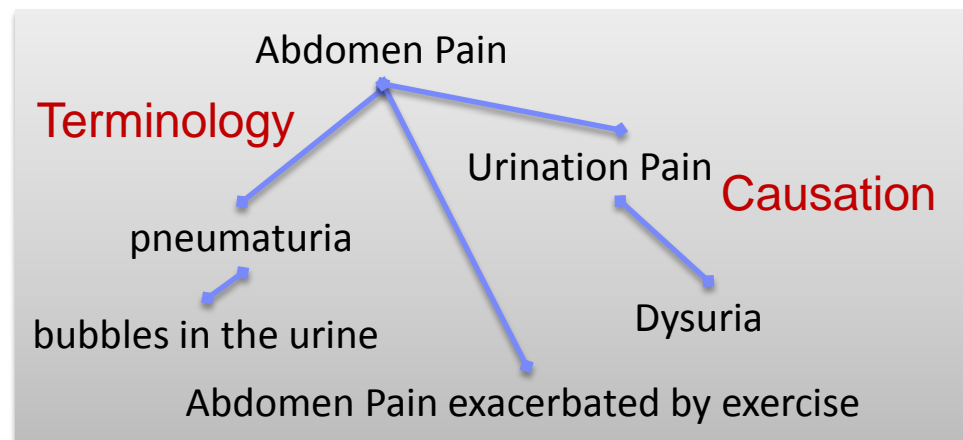
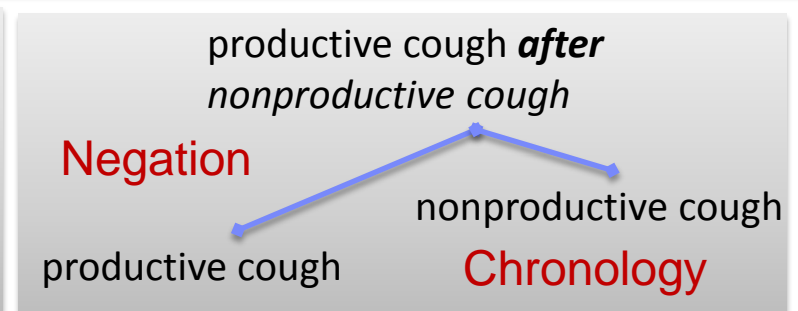
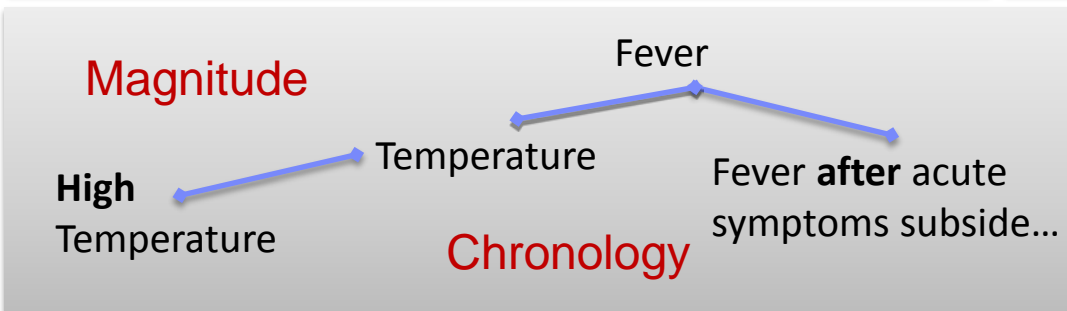
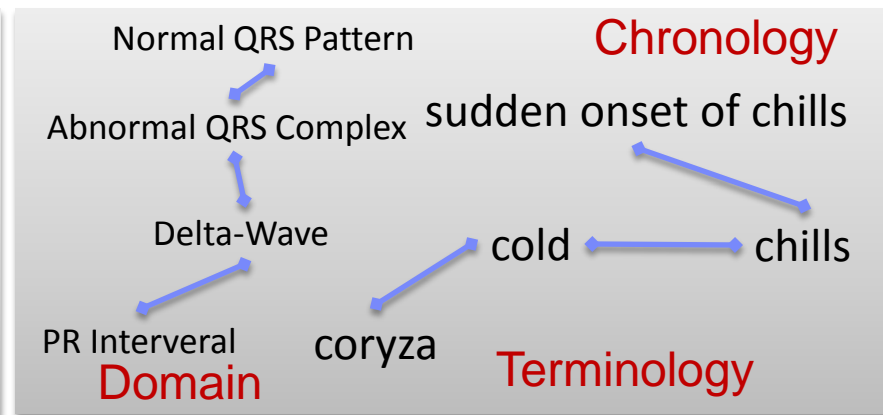
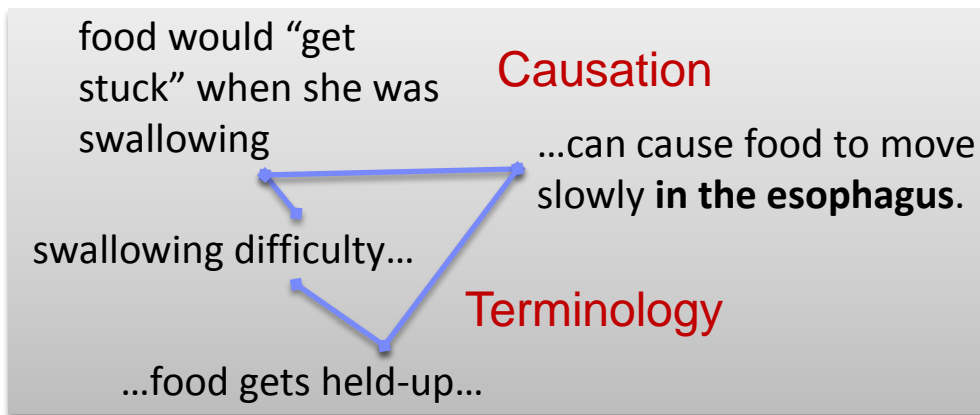
Diagnosis for fever, adenopathy, and arthralgias following drug exposure

This inflammation is characterized by nasal mucosal atrophy and foul-smelling crusts in the nasal passages

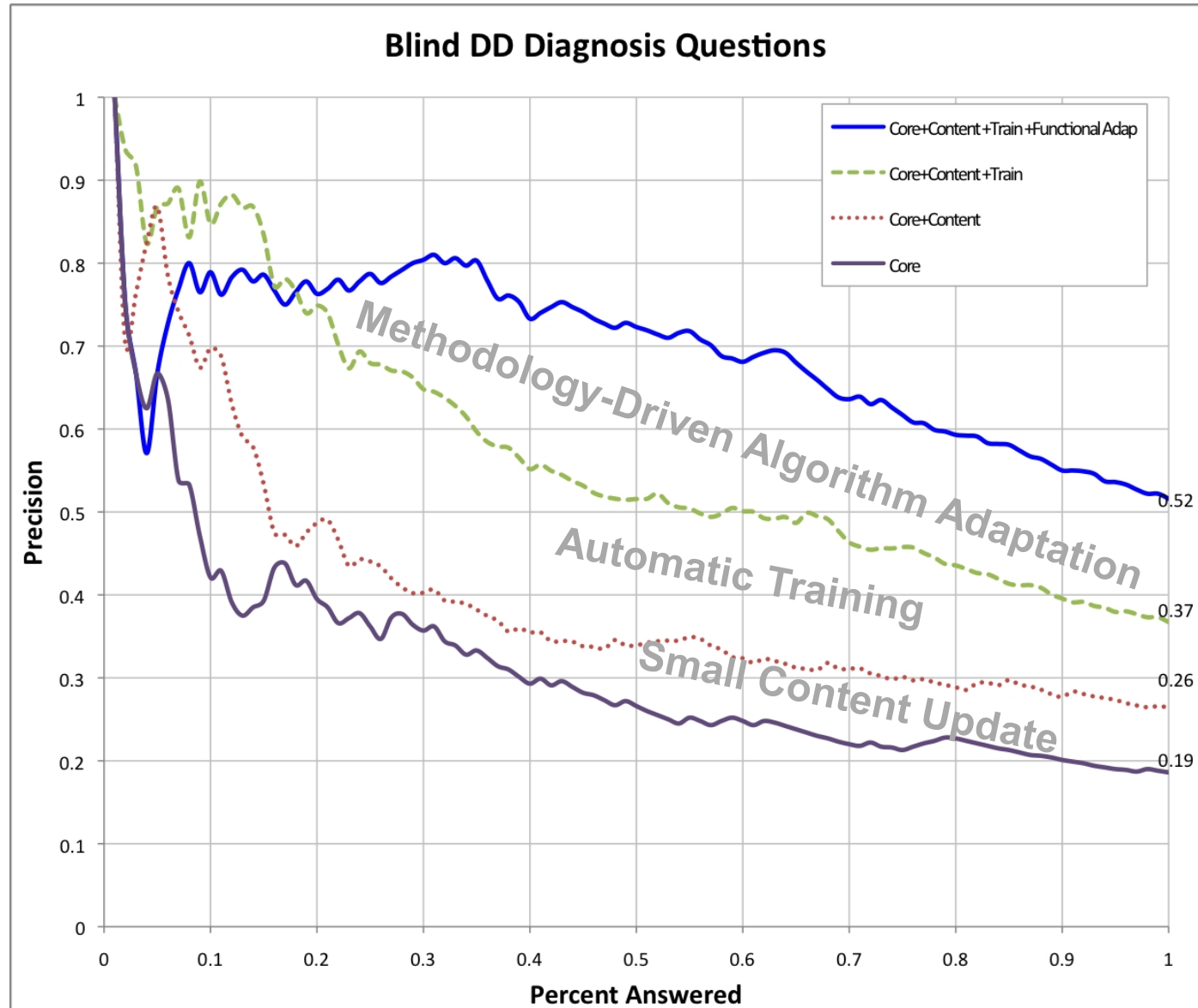
Syndrome of sore throat, fever, sepsis, and unilateral neck swelling

Neurological condition that contraindicates the use of bupropion

Watson is designed to tackle challenges in Language & Meaning



The Approach and the Methodology is Working...Again



Raising the Bar Again...Real World Medical Problems are not single sentence questions; they are rich multi-dimensional scenarios

(Step 1) A 24-year-old man presented with complaints of fever with rigors and pain in the right upper abdomen. He also complains of pain over the right shoulder. After thorough examination and investigation, he was diagnosed with an amebic liver abscess. Which nerve mediates the pain in his right shoulder?

- (A) Axillary
- (B) Descending cervical nerve
- (C) Intercostobrachial
- (D) Right phrenic
- (E) Right sympathetic chain
- (F) Right vagus
- (G) Transverse cutaneous

(Step 2) A 50-year-old obese female underwent a cholecystectomy and T-tube drainage of the common bile duct. On the seventh postoperative day, she developed sudden epigastric and left-sided chest pain. She was short of breath and was sweating profusely. A review of the chart showed that her temperature was between 99 F and 100 F for the last 2 days. The patient's vitals: pulse- 90/min, BP- 110/60 mmHg, and RR- 24/min. The ECG reading showed tachycardia and the ventilation-perfusion lung scan showed a V/Q mismatch.

- (A) Pulmonary embolectomy
- (B) Urokinase therapy
- (C) Exploratory laparotomy and drainage of abscess
- (D) Heparinization
- (E) Inferior vena caval filter

The initial management of this patient is:

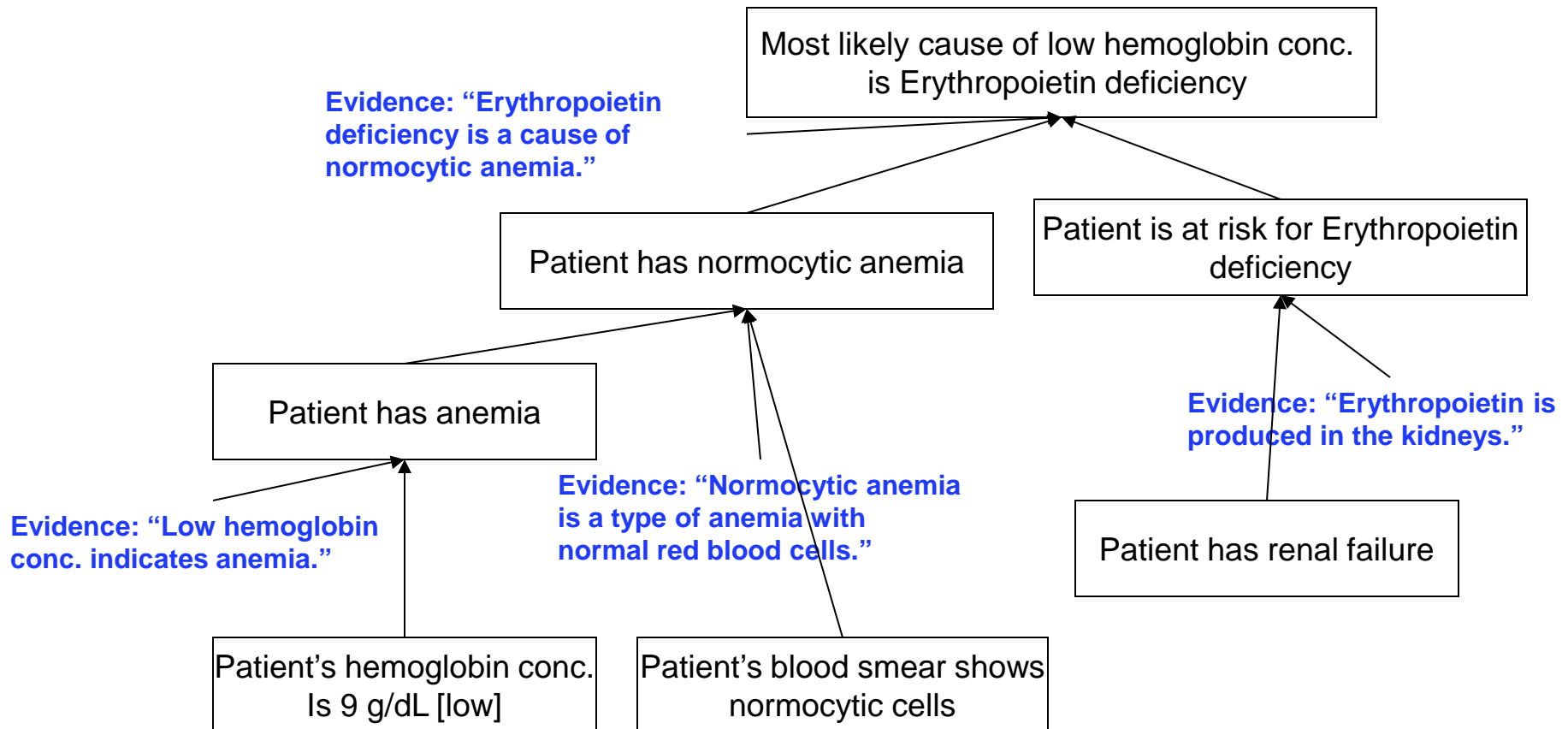
(Step 3) A 45-year-old African American male patient with bipolar disorder and congestive heart failure (CHF) complains of fatigue at a recent office visit. On physical exam his blood pressure is 145 /85 mm Hg, his lungs are clear, and he is non-edematous. His medications include lisinopril, lithium, and furosemide. Later that day his lab work comes back. His serum potassium is 3.4mmol/l, his bicarbonate is 26mmol/l, his creatinine is 1mg/dl. How should you proceed?

- (A) Discontinue lisinopril, monitor blood pressure and potassium
- (B) Discontinue lisinopril, monitor volume status and calcium
- (C) Discontinue lisinopril, monitor volume status and phosphorus
- (D) Discontinue lisinopril, monitor volume status and potassium
- (E) Reduce furosemide, monitor volume status and potassium

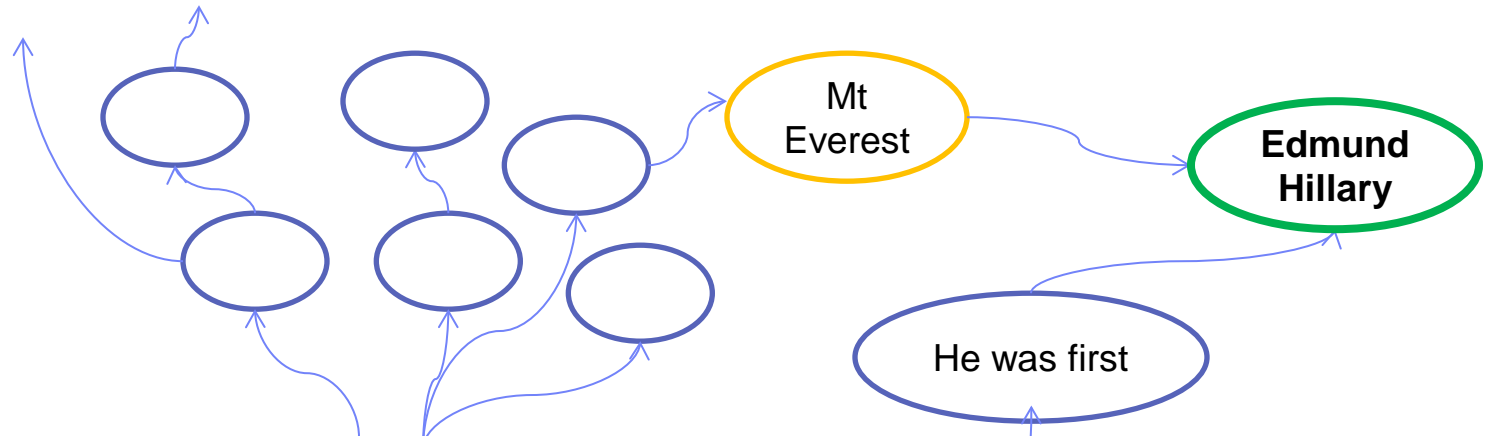
The answers are not one step away!

➤ Discovering rationalized paths through the content becomes the primary value.

“A 32-year-old woman with type 1 diabetes mellitus has had progressive renal failure... Her hemoglobin concentration is 9 g/dL... A blood smear shows normochromic, normocytic cells. What is the problem?”

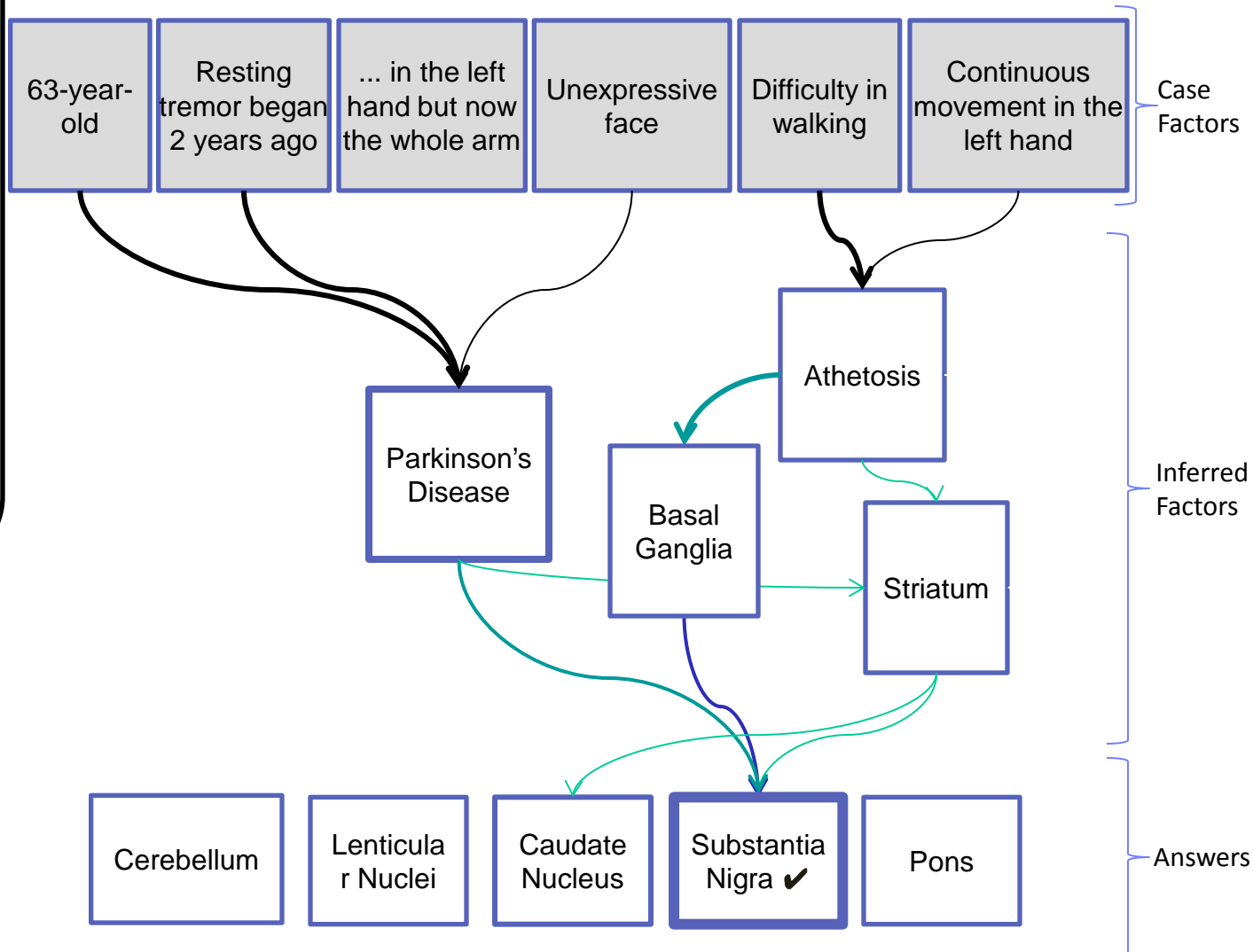


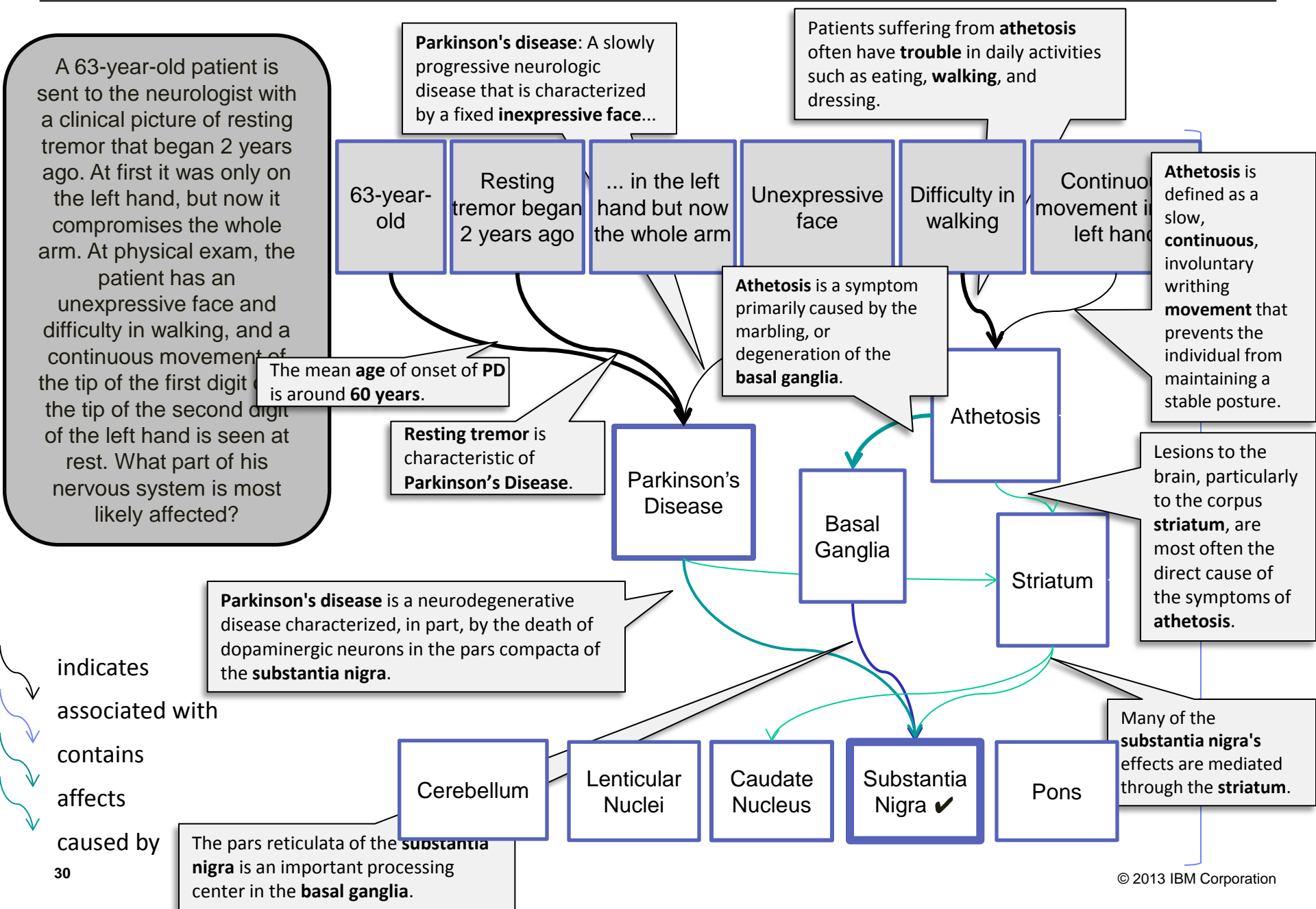
Missing Links in Jeopardy



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

A 63-year-old patient is sent to the neurologist with a clinical picture of resting tremor that began 2 years ago. At first it was only on the left hand, but now it compromises the whole arm. At physical exam, the patient has an unexpressive face and difficulty in walking, and a continuous movement of the tip of the first digit over the tip of the second digit of the left hand is seen at rest. What part of his nervous system is most likely affected?





WatsonPaths

- Given real or pedagogical medical Scenarios...
 - Starting in the Medical school and inspired by “Problem-based Learning” techniques
 - Watson develops and cites an “Inference Graph” – paths from initial factors to candidate conclusions
 - Proposes answers, finds/links concepts, provides confidence and evidence
- Powerful Decision Support
 - Over large volumes of current, Natural Language content
 - Trades-off precision of “if-then” for breadth, currency & flexibility from volumes of content
 - Reduces otherwise costly, narrow, slow and brittle rule development and maintenance
- Collaborative Learning and Decision Support – Users Want...
 - Collaboratively researching and finding solutions to diagnostic and treatment challenges
 - Not just answers.... Missing factors, competing hypotheses, ranked evidence and Confidence
 - Provenance – justifying explanations from all available content in language they understand
 - User has ownership over the rationalization process through interactive collaboration
- TeachWatson
 - As part of the collaboration with WatsonPaths, users can help **Train** and **Teach** Watson
 - Users correct, refine, annotate results and opportunistically answer Watson’s questions
 - Providing invaluable training data and filling Watson’s knowledge gaps
 - Can demonstrate that this technology can incrementally scale over time

Factor Analysis and Identification

The screenshot displays the IBM WatsonPaths application interface. At the top, the 'WatsonPaths' title is visible, along with the IBM Watson logo. Below the title, there are two tabs: 'Question' and 'Inference Graph'. The 'Question' tab is active, showing a medical case description. The 'Inference Graph' tab is also visible, showing a graph of the case. The 'Question' tab contains a text area with the following text: 'A 63-year-old patient is sent to the neurologist with a clinical picture of resting tremor that began 2 years ago. At first it was only on the left hand, but now it compromises the whole arm. At physical exam, the patient has an unexpressive face and difficulty in walking, and a continuous movement of the tip of the first digit over the tip of the second digit of the left hand is seen at rest. What part of his nervous system is most likely affected?'. Below the text area, there is a 'Teach Watson' button. The 'Inference Graph' tab shows a graph of the case, with a node labeled 'Athetosis' circled in red. A callout box points to this node, stating: 'Simple ways to state the case to Watson including using Watson's multidimensional input analysis.' Another callout box points to the 'Athetosis' node, stating: 'Athetosis'. The interface also includes a 'Zoom' slider and a 'Show shortest paths only' checkbox.

WatsonPaths

Question Inference Graph

Factors Hypotheses Inferences Teach Watson Z

Question
A 63-year-old patient is sent to the neurologist with a clinical picture of resting tremor that began 2 years ago. At first it was only on the left hand, but now it compromises the whole arm. At physical exam, the patient has an unexpressive face and difficulty in walking, and a continuous movement of the tip of the first digit over the tip of the second digit of the left hand is seen at rest. What part of his nervous system is most likely affected?

Other Factor
resting tremor Assert

Athetosis

Watson Suggests
A 63-year-old patient is sent to the neurologist with a clinical picture of resting tremor that began 2 years ago.
At first it was only on the left hand but now it compromises the whole arm.
At physical exam the patient has an unexpressive face and difficulty in walking and a continuous movement of the tip of the first digit over the tip of the second digit of the left hand is seen at rest.

Speak now
Cancel

A 63-year-old patient is sent to the neurologist with a clinical picture of resting tremor that began 2 years ago

Athetosis

Simple ways to state the case to Watson including using Watson's multidimensional input analysis.

Athetosis

Zoom Show shortest paths only.

Hypothesis (or Candidate Answer) Generation

WatsonPaths

Question Inference Graph

Factors Hypotheses Inferences TeachWatson Z

Extracted Candidate Answers

- Cerebellum
- Lenticular nuclei
- Pons
- Caudate nucleus
- Substantia nigra

Hypothesize Answer

Striatum Add

Striatum

Watson Suggests

- Caudate nucleus
- Substantia nigra

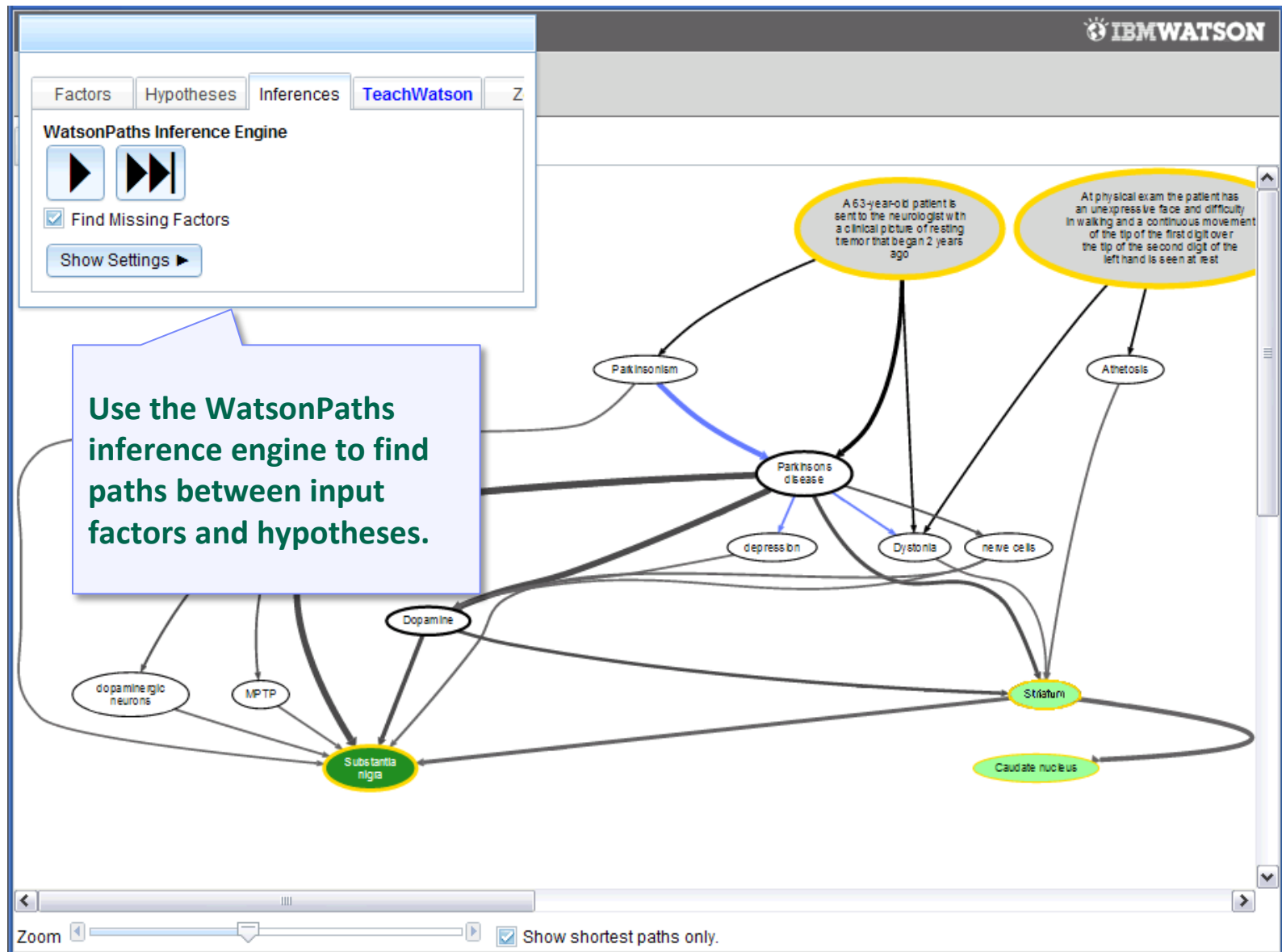
Speak now

Cancel

Substantia nigra Caudate nucleus Striatum

Zoom ☒ Show shortest paths only.

Discover and Evidence Paths from Input Factors to Hypotheses



Explore the Evidence and Teach Watson

The screenshot shows the IBM Watson interface for a diagnostic reasoning task. On the left, a sidebar contains tabs for 'Factors', 'Hypotheses', 'Inferences', and 'TeachWatson'. The 'TeachWatson' tab is active, displaying a 'Judge Connection' slider (weak to strong), 'Watson's Confidence' slider, and 'Judge Supporting Evidence'. The evidence section shows two paragraphs of text with highlighted phrases like 'Resting tremor' and 'typifies Parkinson's disease'. Below this, a 'TeachWatson' section asks for an explanation of the word 'typifies' and provides a 'Correct' or 'Wrong' feedback option. The main area displays a network diagram with nodes representing anatomical structures: 'Substantia nigra', 'Striatum', 'Caudate nucleus', and 'Athetosis'. A large text box with a green border and a speech bubble icon contains the text: 'See Watson's evidence for making a connection. Judge it to improve the inference and Teach Watson in the process.' The interface also includes a 'Zoom' slider and a 'Show shortest paths only' checkbox at the bottom.

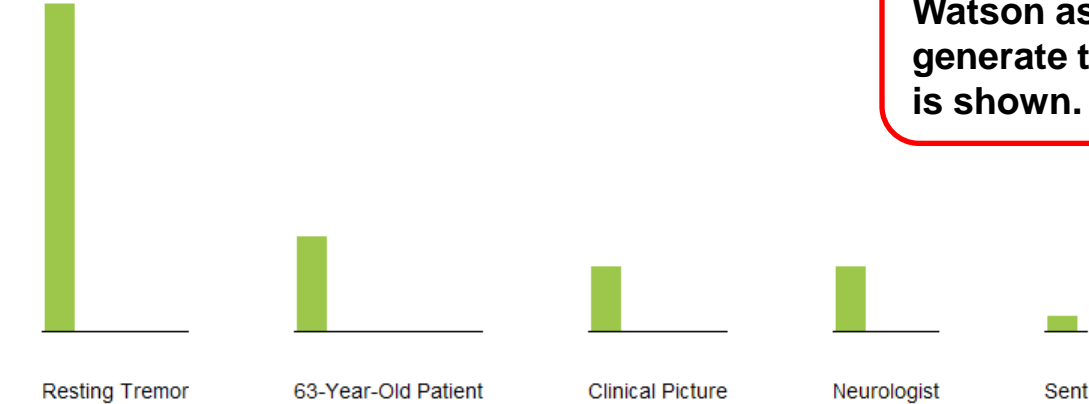
Open Evidence

www.localhost.com:8080/open_evidence_wksp/com.ibm.watson2.openEvidence.question_client/www/?dataSourceType=wics&dataSource=/users/ak

IBM WATSON

2000000 . What causes: A 63-year-old patient is sent to the neurologist with a clinical picture of resting tremor that began 2 years ago. [Back to Inferences View](#)

Evidence Profile:

Parkinson's disease	10%
Show <input type="button" value="factors"/> Order by <input type="button" value="delta"/> <input type="checkbox"/> Show Non-Contributors	
	
Evidence	
<div><div></div> 47% Tremor</div> <p>Details</p> <p>Essential tremor is typically a mixture of postural and kinetic tremors although one form may predominate over the other. Rest tremor occurs rarely in elderly patients with advanced essential tremor, and physicians must always consider the possibility of coexistent Parkinson disease or some other cause of parkinsonism in patients with rest tremor. In most instances, the "rest tremor" in essential tremor is actually a postural tremor that is caused by incomplete muscle relaxation. Patients are often examined while seated in a chair that does not provide complete support for the head, neck, torso, and proximal limb muscles.</p> <div><div></div> 46% Tremor</div> <p>Details</p> <p>In general, one particular tremor type is predominant and sometimes the only tremor present in a defined clinical condition,</p>	

And the evidence it found to support the answer is shown.

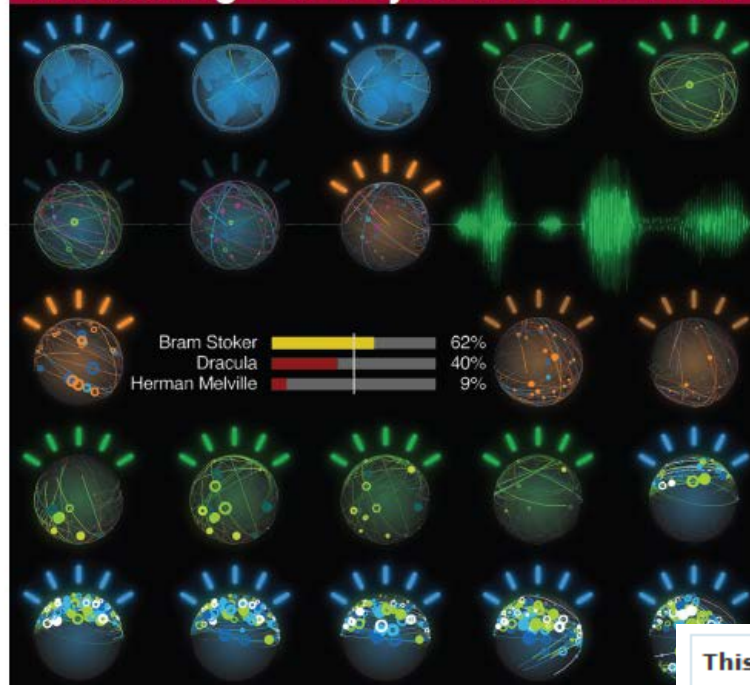
The question Watson asked to generate the edge is shown.

VOLUME 56, NUMBER 3/4, MAY/JUL. 2012

IBM

Journal of Research and Development

Including IBM Systems Journal



This Is Watson

THANK YOU

IBM
Systems Journal

Vol. 45, No. 3, 2004

Unstructured
Information
Management

Unstructured information represents the vast majority of the data collected and accessible to enterprises. Creating order from an abundance of sources in many formats, unstructured information management (UIM) aims to make this data accessible and searchable via management systems and applications. The interplay of shapes and colors on the cover conveys the diversity of unstructured data sources and the great potential of UIM for harnessing these sources by providing the needed structure. The visualization output of a cluster analysis of a group of genes is shown as an example of the ability of UIM to bring order out of chaos.

This is Watson

In 2007, IBM Research took on the grand challenge of building a computer system that could compete with champions at the game of Jeopardy!. In 2011, the open-domain question-answering system dubbed Watson beat the two highest ranked players in a nationally televised two-game Jeopardy! match. This special issue provides a deep technical overview of the ideas and accomplishments that positioned our team to take on the Jeopardy! challenge, build Watson, and ultimately triumph. It describes the nature of the question-answering challenge represented by Jeopardy! and details our technical approach. The papers herein describe and provide experimental results for many of the algorithmic techniques developed as part of the Watson system, covering areas including computational linguistics, information retrieval, knowledge representation and reasoning, and machine learning. The papers offer component-level evaluations as well as their end-to-end contribution to Watson's overall question-answering performance.

Links

- “Artificial Intelligence: Learning Through Interactions and Big Data,” An *IBM Redbooks®* Point-of-View publication by IBM:
<http://www.redbooks.ibm.com/redpapers/pdfs/redp4974.pdf>
- “This is Watson” issue of *IBM Journal of Research & Development* , 2012:
<http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=6177717>
- “Unstructured Information Management” issue of *IBM Systems Journal*, 2004:
<http://researchweb.watson.ibm.com/journal/sj43-3.html>

ACM: The Learning Continues...

- Questions about this webcast? learning@acm.org
- ACM Learning Webinars:
<http://learning.acm.org/multimedia.cfm>
- ACM Learning Center: <http://learning.acm.org>
- ACM SIGSOFT: <http://www.sigsoft.org/>
- FSE 2013: <http://esec-fse.inf.ethz.ch/>



Association for
Computing Machinery

Advancing Computing as a Science & Profession

IBM, the IBM logo, ibm.com, Smarter Planet and the planet icon are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Jeopardy! Elements TM & © 2013 Jeopardy Productions, Inc. All Rights Reserved. A current list of IBM trademarks is available on the Web at www.ibm.com/legal/copytrade.shtml. © International Business Machines Corporation 2013.