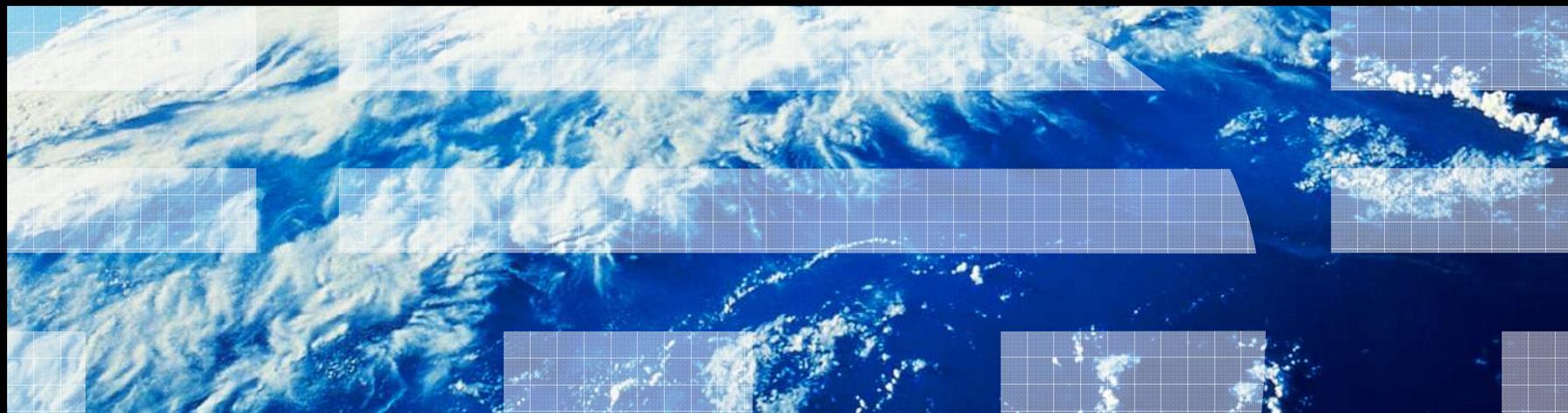




Spring 2011

What's cooking at IBM Research?

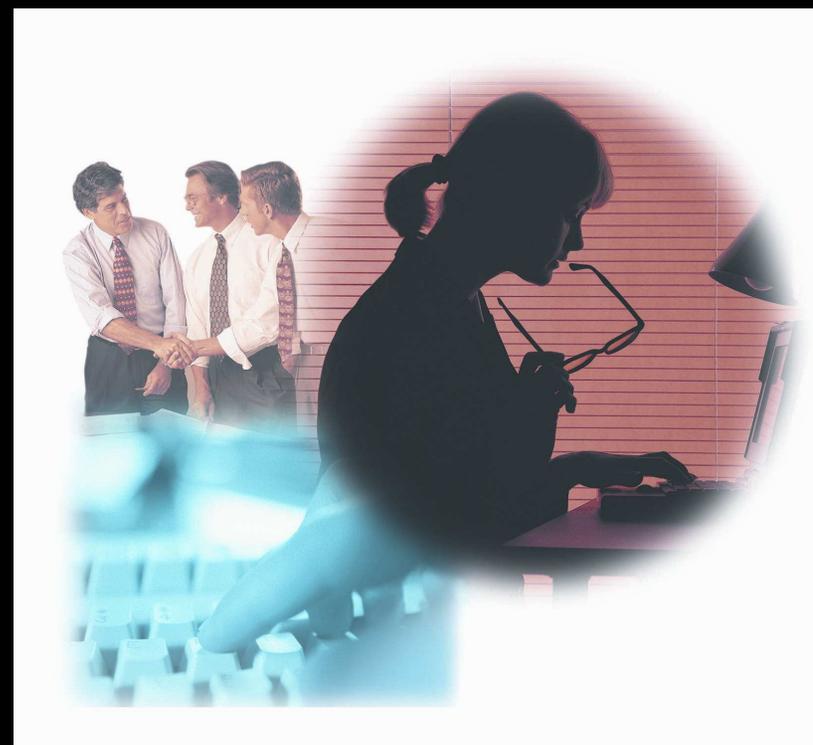


Until now the nuances of human language have posed a major challenge for computers

What if
...a **computer** system **could**
understand natural human
language?

The challenge: Human language is

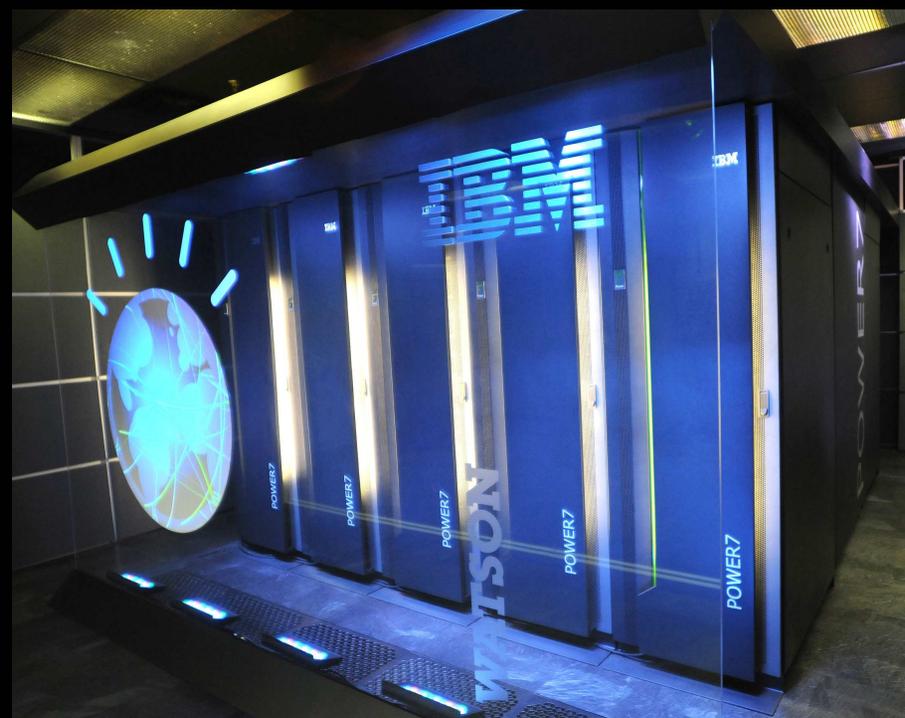
- **implicit**
- highly **contextual**
- **ambiguous**
- often **imprecise**



IBM's answer to the challenge: Watson

– An advanced question-answering computing system

- Tremendous **breakthrough** in computer's **understanding natural language**
- Rivals human's ability to answer questions with speed, accuracy and confidence
- **Stand-alone**, not connected to the Internet: **works with its own database**
– 200 million content pages
- **More than a search engine**: Gives **precise answer** – instead of a ranked list of web pages

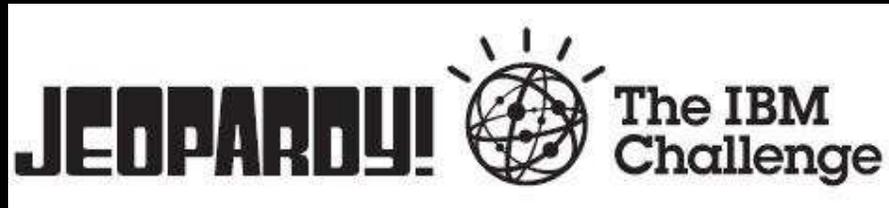


The test: Competing on Jeopardy! – America's favorite quiz show

Jeopardy! – a grand challenge for a computing system:

- **broad range of topics** (e.g. history, literature, politics, arts, science)
- analyzing **subtle meanings** in clues
- dealing with irony, riddles, and other **language complexities**
- **speed** of accurate responses (max. 3 s)
- **high** level of **confidence** in answer required





... and the winner is – Watson!

On February 16 2011, the IBM Watson system made history:

Watson won against Ken Jennings and Brad Rutter – the two most successful Jeopardy! contestants ever.

Watson's prize money is donated to charities.



What you need to answer to become a Jeopardy! champion

Category: *Cambridge*

(sample clue featuring ambiguity)

With much “gravity” this young fellow of trinity became the Lucasian professor of mathematics in 1669.

00:00:03

Who is Isaac Newton?

What you need to answer to become a Jeopardy! champion

Category: *Don't worry about it* (sample clue featuring play on words)

You just need a nap! You don't have this sleep disorder that can make sufferers nod off while standing up.

00:00:03

What is narcolepsy?

What you need to answer to become a Jeopardy! champion

Category: *Familiar sayings*

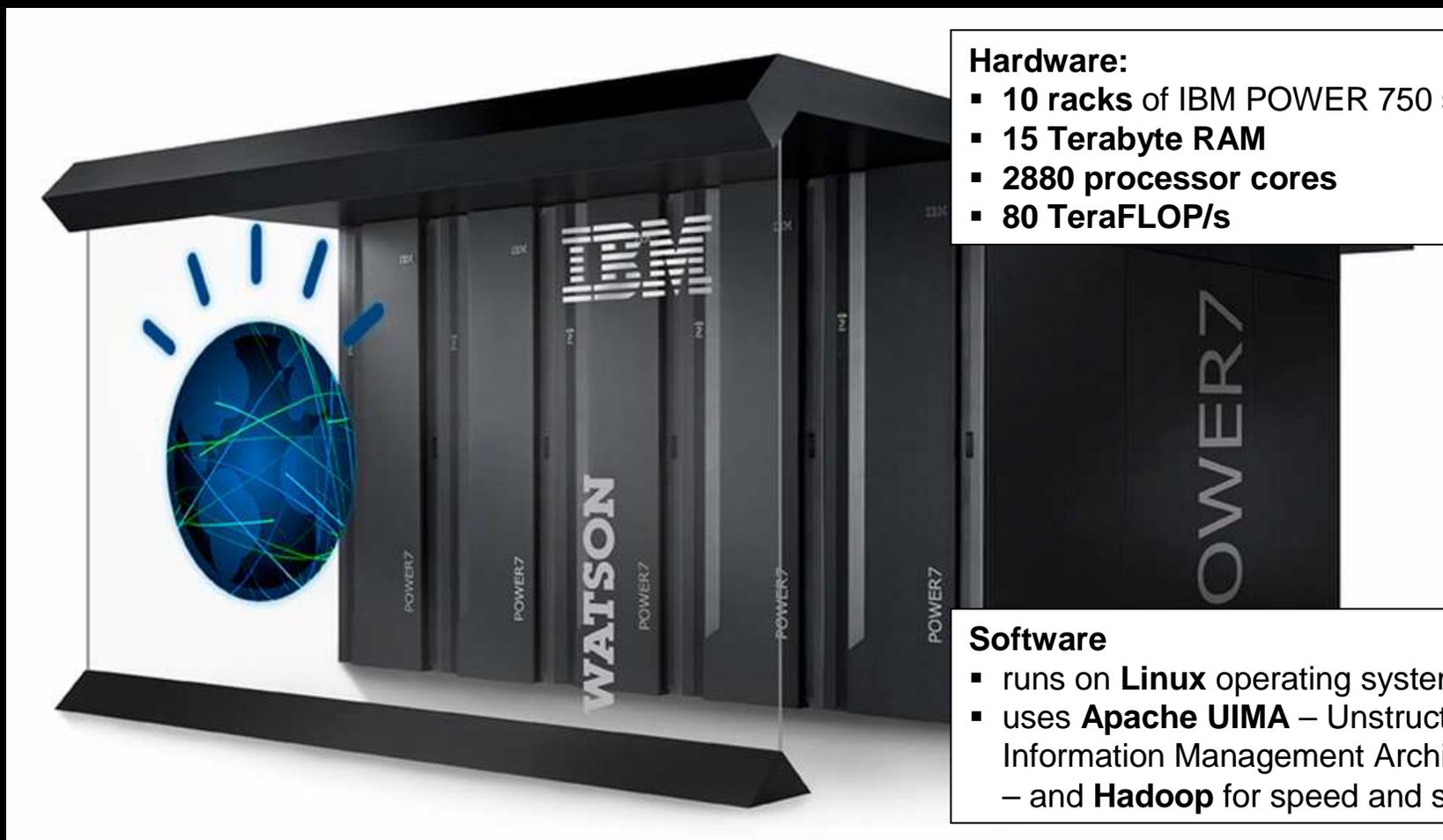
(sample clue featuring a riddle)

Even a broken one of these on your
wall is right twice a day.

00:00:03

What is a clock?

What powers Watson?



Hardware:

- 10 racks of IBM POWER 750 servers
- 15 Terabyte RAM
- 2880 processor cores
- 80 TeraFLOP/s

Software

- runs on **Linux** operating system
- uses **Apache UIMA** – Unstructured Information Management Architecture – and **Hadoop** for speed and scale-out

How Watson “thinks”

Category: *WORLD GEOGRAPHY*

Clue: *In 1897 Swiss climber Matthias Zurbriggen became the first to scale this Argentinean peak.*

Step 1:

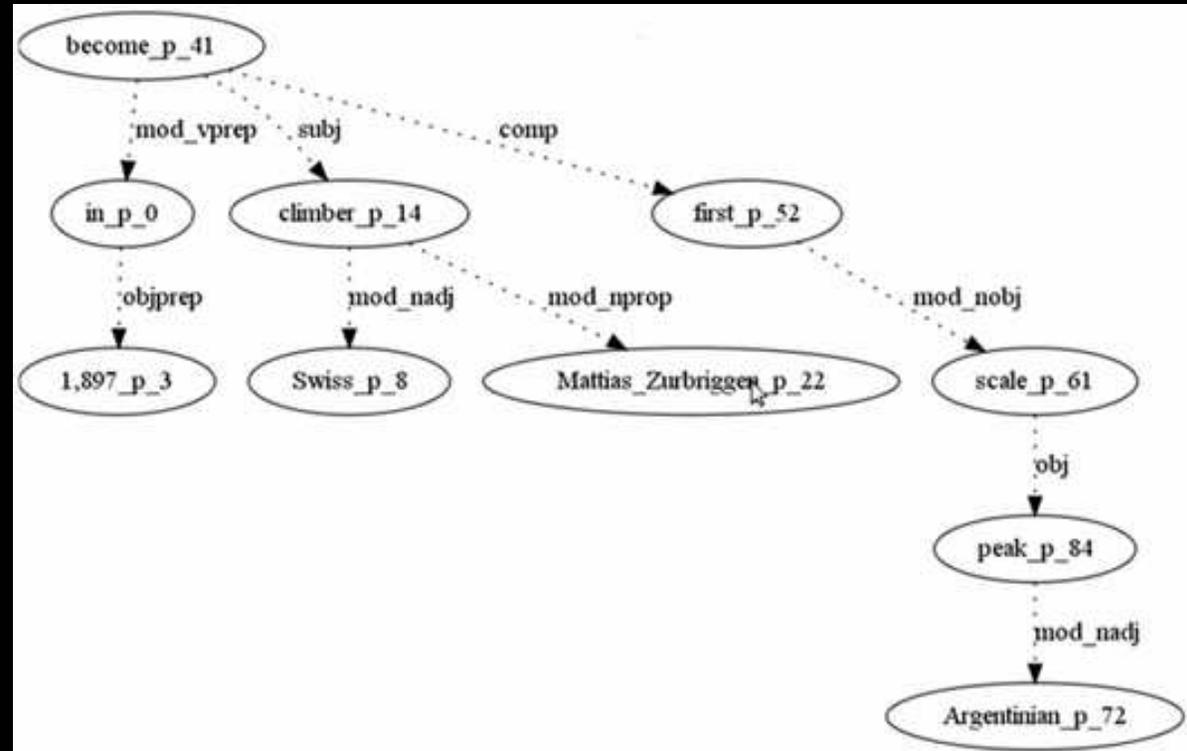
Watson dissects the clue to understand what it is asking for.

Watson tokenizes and parses the clue to **identify** the **relationships between important words** and **find** the **focus of the clue**, i.e. *this Argentinean peak.*

Step 1: Analyzing the question

Category: WORLD GEOGRAPHY

Clue: In 1897 Swiss climber Matthias Zurbriggen became the first to scale this Argentinian peak.



Step 2: Search

Timeline of Climbing the Matterhorn

* August 25: H.R.H. the Duke of the Abruzzi made the ascent with Mr. A. F. Mummery and Dr. Norman Collie, and one porter, Pollinger, junior. According to Mummery the weather was threatening, and, the Prince climbing very well, they went exceedingly fast, so that their time was probably the quickest possible. They left the bivouac at the foot of the snow ridge at 3.40 a.m., and reached the summit at 9.50. A few days afterwards the first descent of the ridge was accomplished by Miss Bristow, with the guide **Matthias Zurbriggen**, of Macugnaga.

The first known ascent of Aconcagua was during an expedition was during an expedition led by Edward Fitz Gerald in the summer of **1897**. **Swiss climber Matthias Zurbriggen** reached the summit alone on January 14 via today's Normal Route. A few days later Nicholas Lanti and Stuart Vines made the second ascent. These were the highest ascents in the world at that time. It's possible that the mountain had previously been climbed by Pre-Columbian Incans.

Step 2:

Watson **searches its databases for text passages that contain the key words.**

Using important terms from the clue, Watson performs a search over millions of documents to find relevant passages.

Step 3: Hypothesis & candidate generation

Timeline of Climbing the **Matterhorn**

* August 25: H.R.H. the Duke of the **Abruzzi** made the ascent with Mr. A. F. Mummery and Dr. Norman Collie, and one porter, Pollinger, junior. According to Mummery the weather was threatening, and, the Prince climbing very well, they went exceedingly fast, so that their time was probably the quickest possible. They left the bivouac at the foot of the snow ridge at 3.40 a.m., and reached the summit at 9.50. A few days afterwards the first descent of the ridge was accomplished by **Miss Bristow**, with the guide Matthias Zurbriggen, of Macugnaga.

The first known ascent of **Aconcagua** was during an expedition led by Edward Fitz Gerald in the summer of 1897. Swiss climber Matthias Zurbriggen reached the summit alone on January 14 via today's **Normal Route**. A few days later Nicholas Lanti and Stuart Vines made the second ascent. These were the highest ascents in the world at that time.
It's possible that the mountain had previously been climbed by Pre-Columbian Incans.

Answer candidates

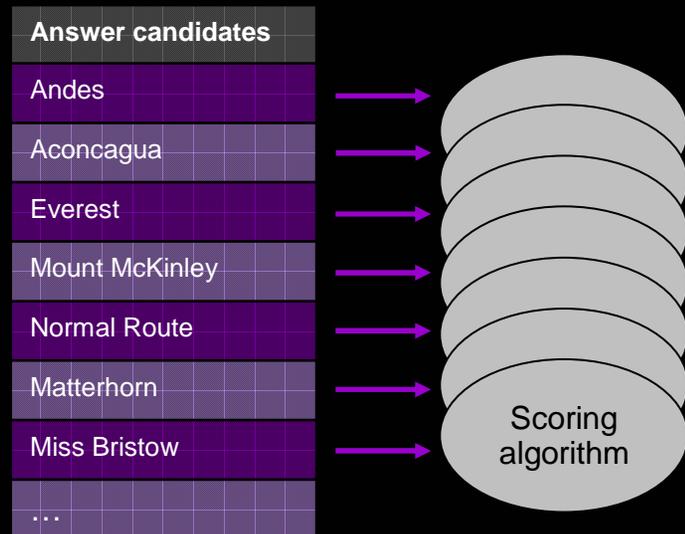
Andes
Aconcagua
Everest
Mount McKinley
Normal Route
Matterhorn
Miss Bristow
...

Step 3:

Watson analyzes the text passages and generates possible “candidate answers”.

Watson extracts important entities (“candidate answers”) from the documents. **Focus is on coverage**: as much as possible is added (e.g., peaks, mountain ranges, people).

Step 4: Answer scoring



Type coercion scorers: Estimate the likelihood of a candidate answer being the right type (here a peak).

Temporal match: If question includes a reference to a date, this scorer will look for evidence in the passage that matches that date.

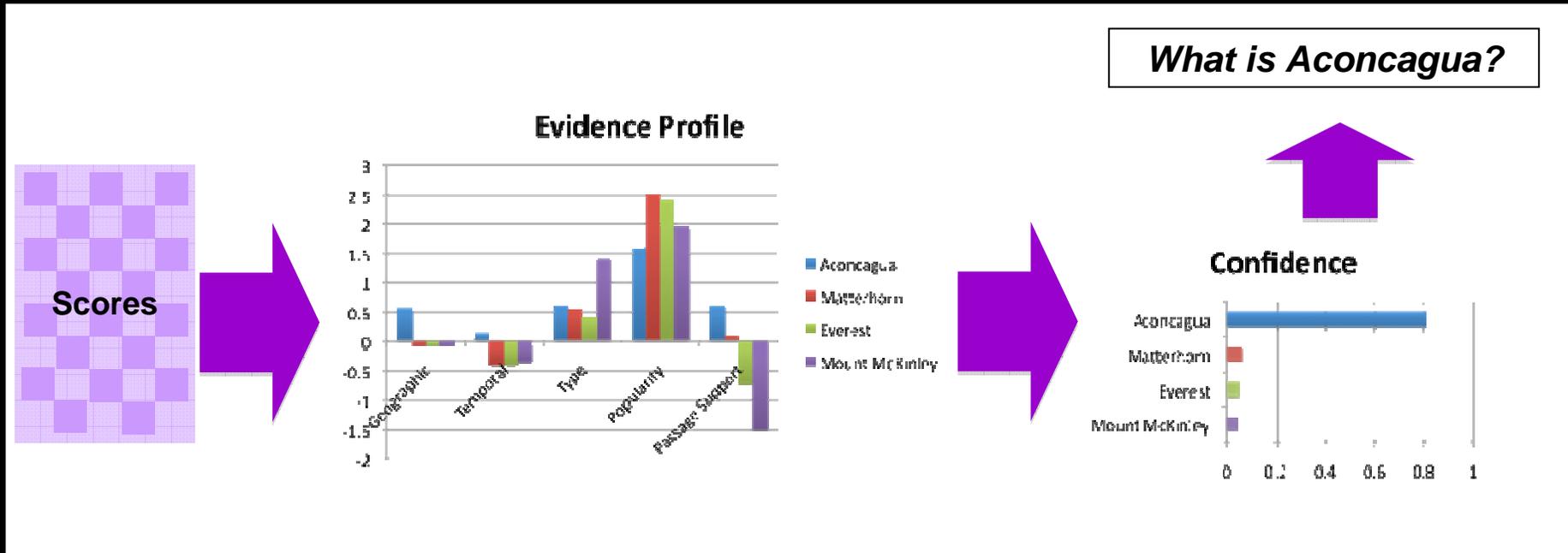
Answer scorers depending on different relations or constraints detected in question.
Modifier “Argentinean” peak is detected as geospatial relation, indicating correct answer must be located in Argentina.

Passage scorer: One of the most important answer scores. A single supporting text passage is scored using a variety of complex passage scoring analytics to measure the evidence provided by the passage.

Step 4:

Candidate answers are scored using a large number of **answer-scoring analytics**. In a massively parallel manner, Watson uses >100 answer and deep evidence scoring algorithms to determine how well a candidate answer matches what the clue is asking for.

Step 5: Summarizing all evidence



Step 5:

Watson **summarizes all evidence and determines its confidence** in the answers.

Scores are grouped into meaningful **evidence dimensions**. Their plot yields the evidence profile for the candidate answer.

Watson statistically combines the scores to produce a **final confidence score**.

Beyond the game: Real-world applications for Watson

- **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



Watson's ability to understand the meaning and context of human language and to process information rapidly to find precise answers holds enormous potential to drive progress across many industries.

A closer look at healthcare



Watson has the potential

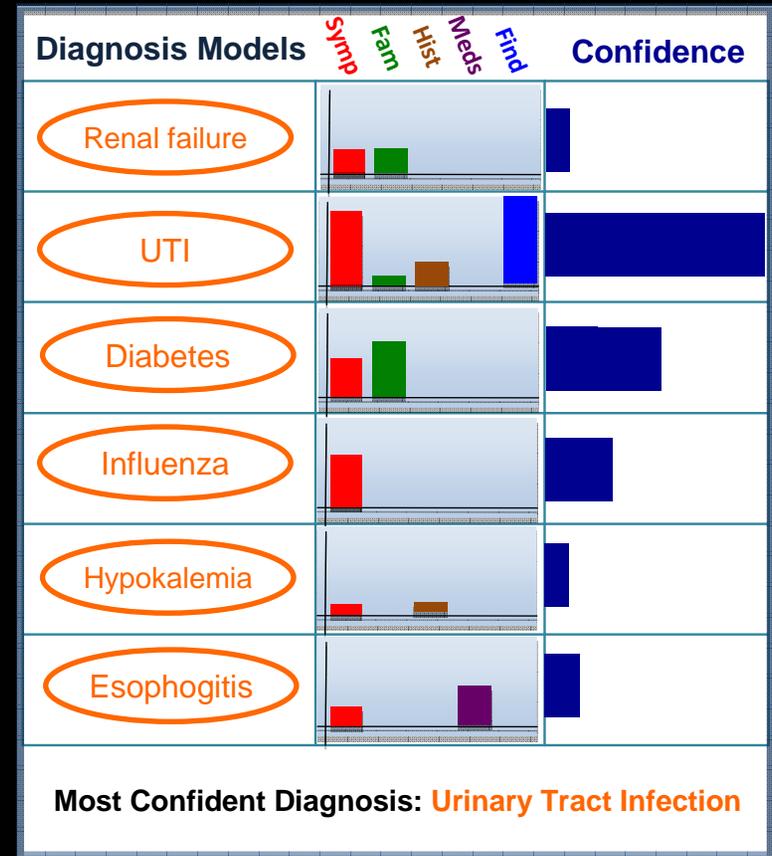
- **to provide critical and timely information** to doctors **to help diagnose** and treat patients, leading to significant improvements in healthcare
- to **establish evidence profiles from disparate data** (e.g. latest knowledge in medical literature, reference material, prior cases, tests and findings, individual medical data of patient and symptoms)

Watson's Deep Q&A system proposes a differential diagnosis to help personalize medical treatment

Continuous evidence-based diagnostic analysis

The system considers and synthesizes a broad range of evidence:

- Symptoms
- Family history
- Patient history
- Medications
- Tests & findings
- Notes & hypotheses
- Huge volumes of texts:
journals, references, databases etc.



... and the winner is: Humankind



The sun never sets at IBM Research

