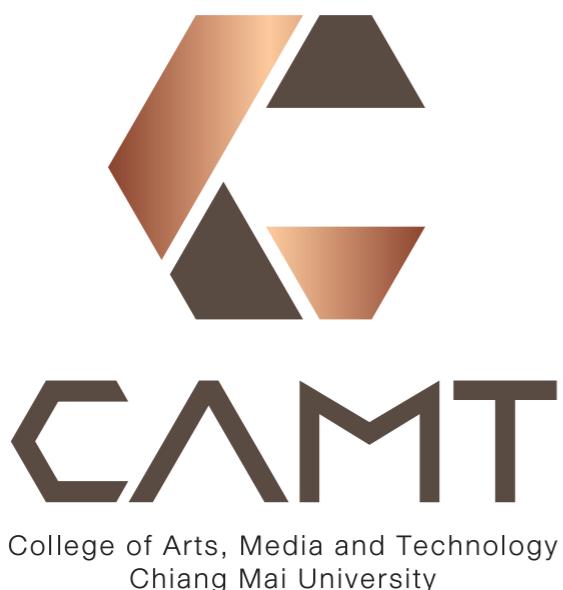


SE 481 Introduction to Information Retrieval

Module #8 — Question-answering system



Passakorn Phannachitta, D.Eng.

passakorn.p@cmu.ac.th

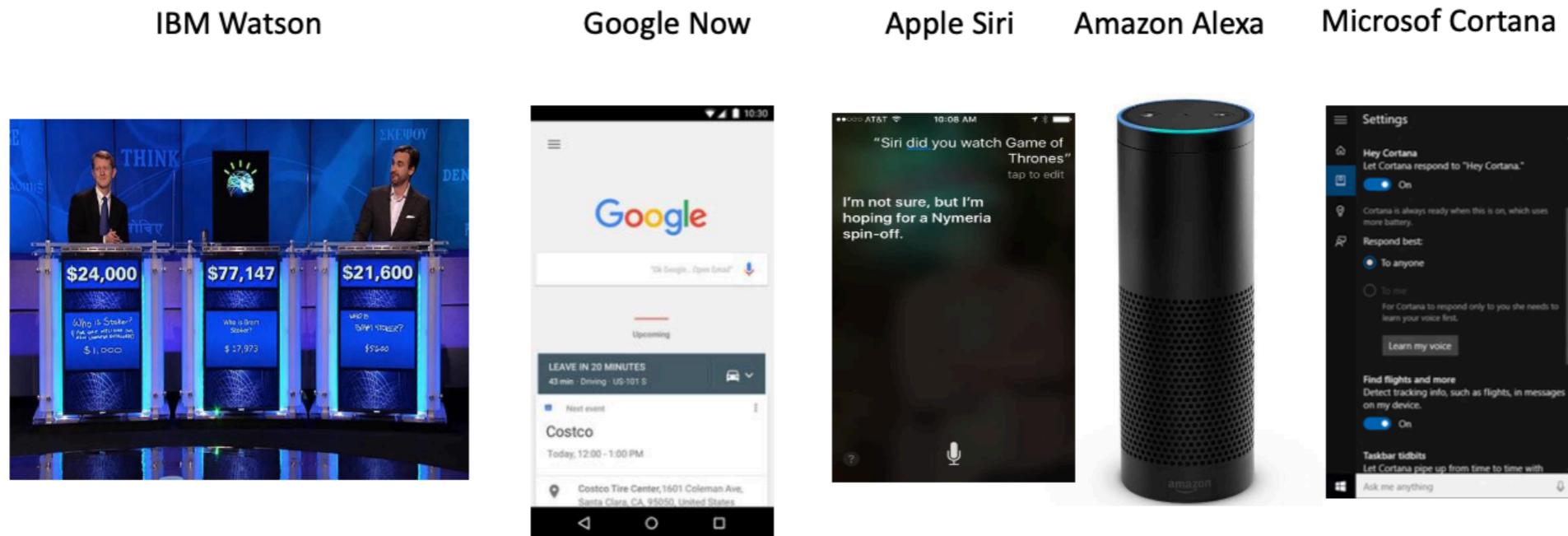
College of Arts, Media and Technology
Chiang Mai University, Chiangmai, Thailand

Agenda

- What is it?
- A storytelling about IBM Watson

Question-answering system

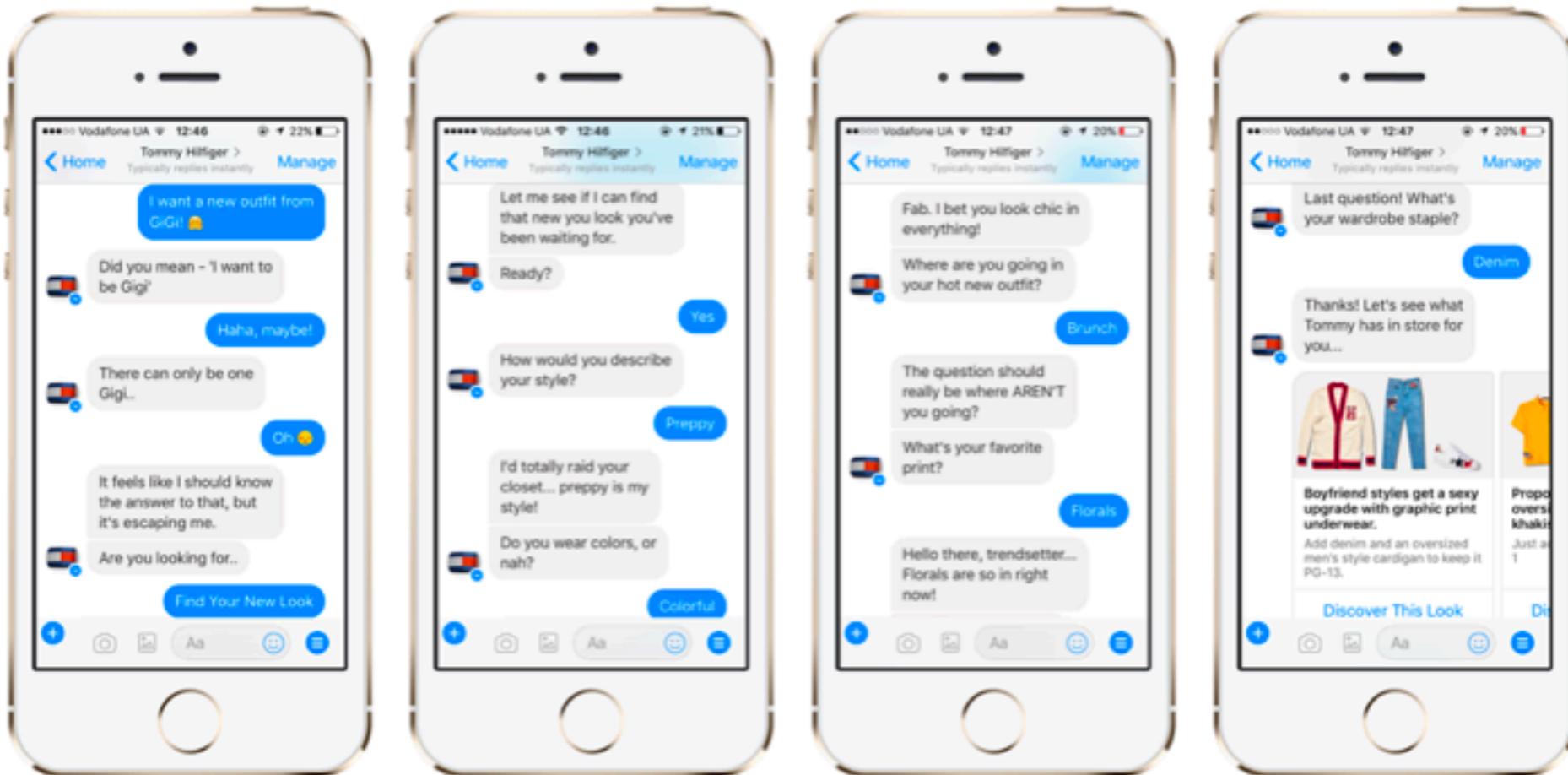
- Systems that answer natural language questions.



Ref: <http://knowledgeworks.cn/>

Other applications

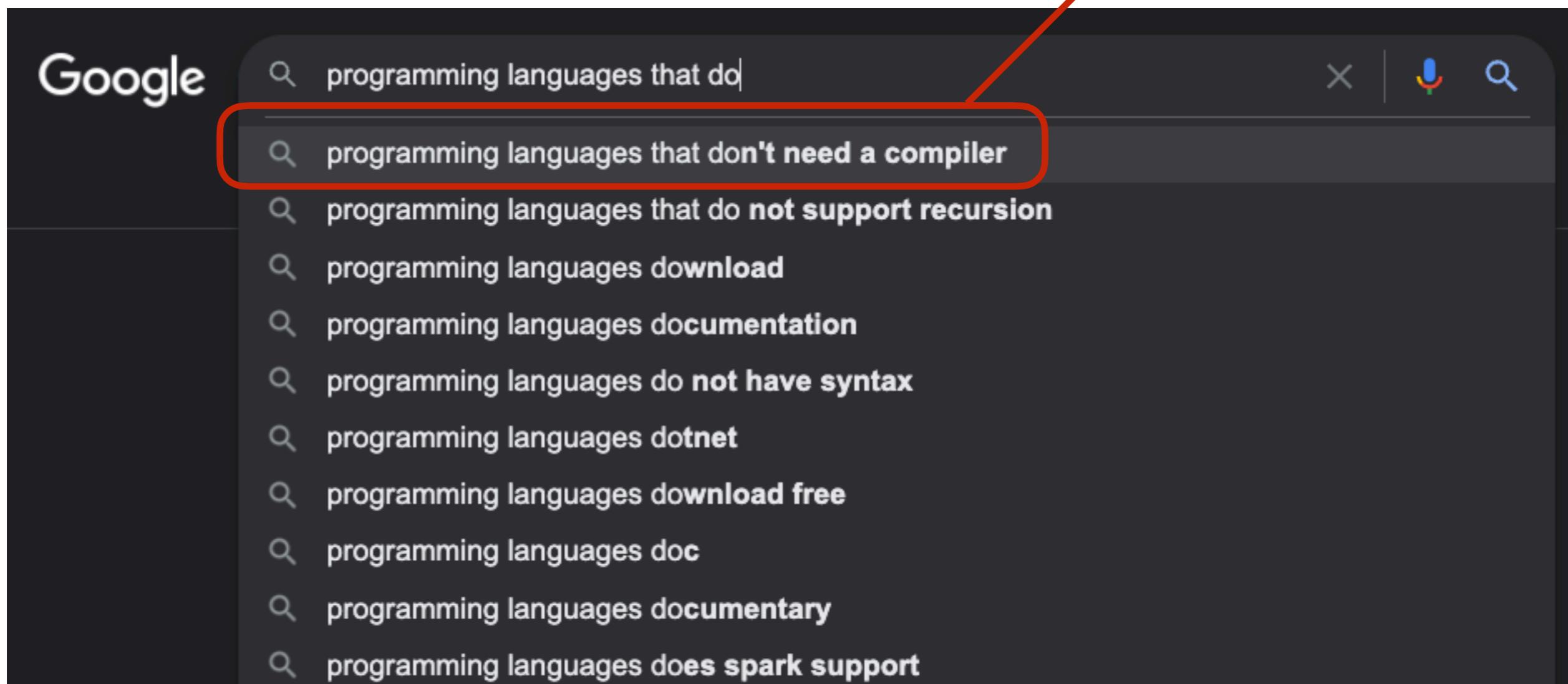
- Natural human-computer interaction
 - Chatbot



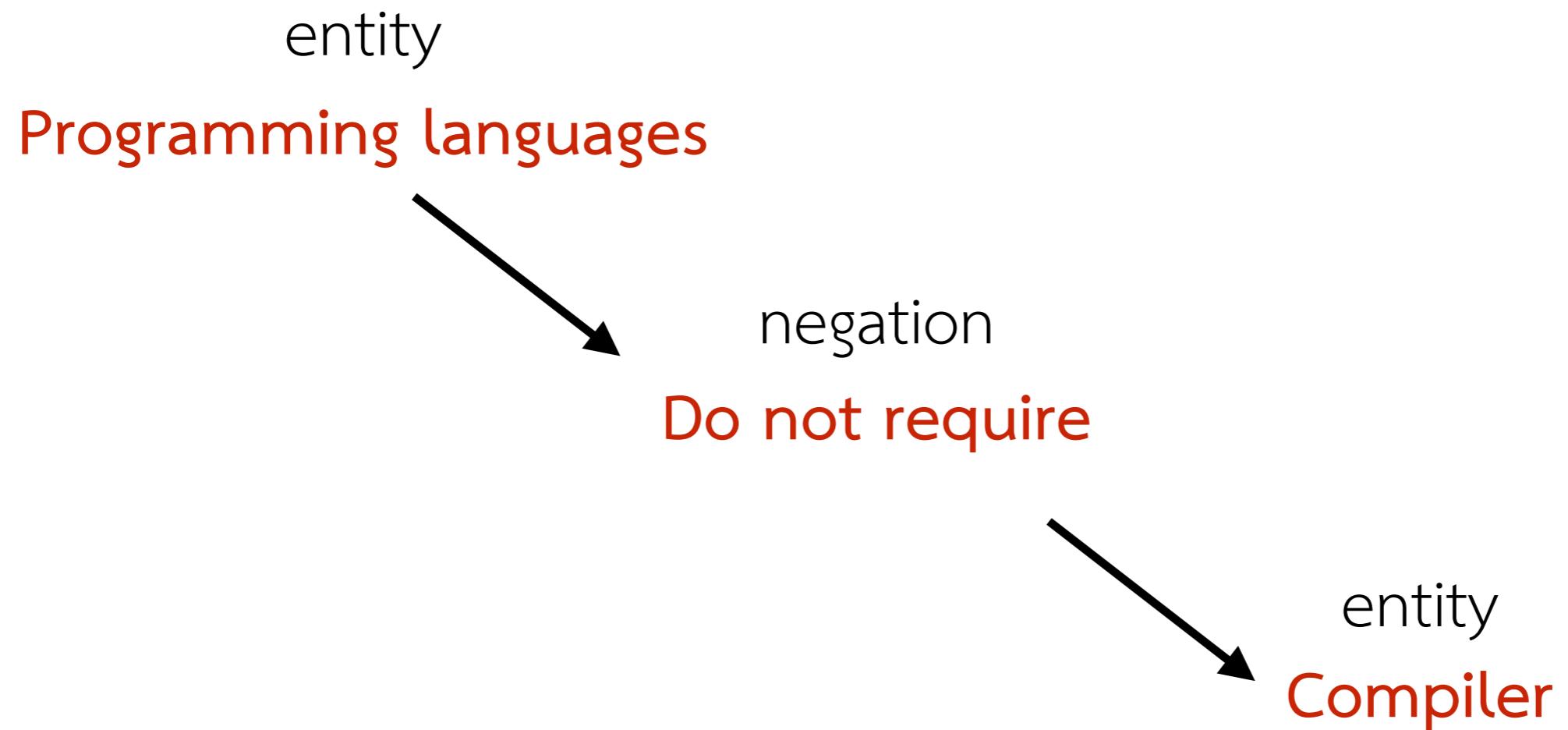
Ref: <https://rubygarage.org/blog/why-retail-should-use-chatbots>

Example processes

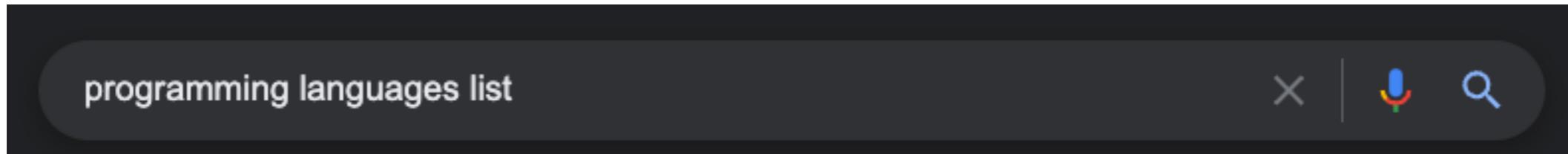
Started from frequent queries



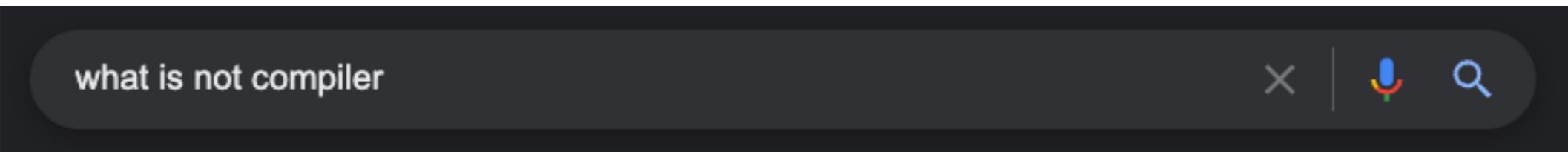
Example processes



Example processes



- e.g., C, python, java, ...



- Interpreter

Example processes

A screenshot of a search results page from a search engine. The search bar at the top contains the query "programming languages that don't need a compiler". Below the search bar are navigation links: All, Images, Videos, News, Shopping, More, and Tools. A horizontal line separates these from the search results. The results section starts with a summary: "About 17,100,000 results (0.61 seconds)". Below this, a snippet of text reads: "Scripting languages don't require to be compiled rather they are interpreted." followed by an ellipsis "...". Another snippet follows: "Examples of Scripting Language (These languages are traditionally used without explicit compilation process. But today some of them may vary.)". A bulleted list of languages is provided: • JavaScript. • Lua. • Perl. • Python. • PHP. • VBScript etc.

Approaches

- Text-based QA
- Knowledge-based QA
- Hybrid QA

Text-based QA

- Question processing
 - Detect question type, answer type, focus, relations
 - Then, formulate queries to send to a search engine
- Passage retrieval
 - Retrieve ranked documents
 - Break into suitable passages and rerank
- Answer processing
 - Extract and rank candidates
 - Format to natural language

Text-based QA

Word Alignment for Question Answering QA (1999-2005)

TREC

What is the fastest car in the world?

The Jaguar XJ220 is the dearest, fastest and most sought after car on the planet.

[Harabagiu & Moldovan, 2001]

- Assume that there is an underlying alignment
 - Describes which words in and can be associated

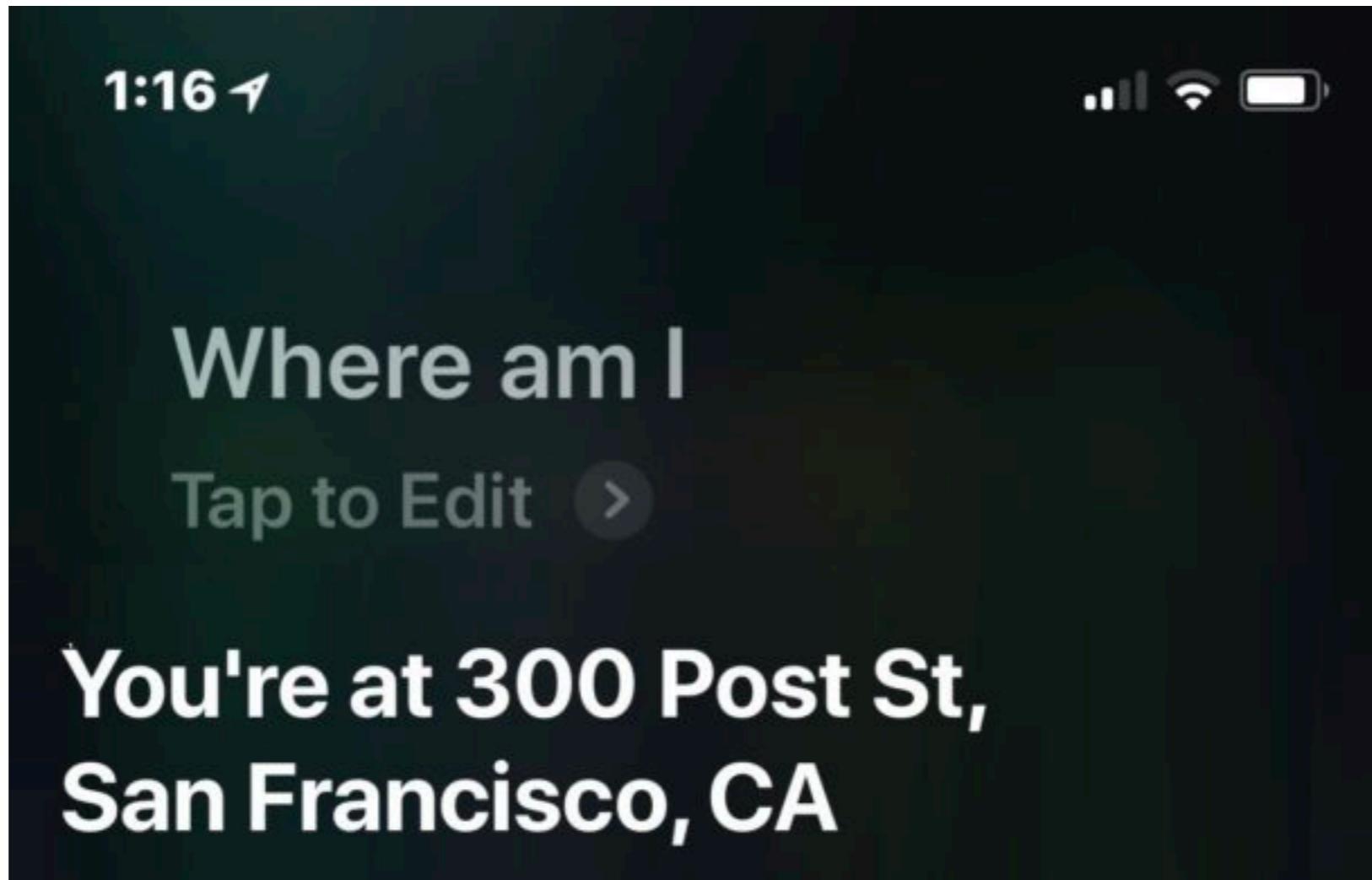
See if the (syntactic/semantic) relations support the answer

Ref: <https://slidetodoc.com/introduction-to-information-retrieval-14-questionanswering-qa-most/>

Knowledge-based QA

- Build a semantic representation of the query
 - e.g., times, dates, locations, entities, numeric quantities
- Map from this semantics to query structured data or resources
 - e.g., Lat-long databases
 - Ontologies — Wikipedia
 - Hotel/restaurant review places
 - Scientific databases
 - Wolfram alpha
 - etc

Knowledge-based QA



Ref: <https://osxdaily.com/2018/10/23/get-current-location-siri-where-am-i/>

Hybrid QA

- Build a semantic representation of the query
 - e.g., times, dates, locations, entities, numeric quantities
- Generate answer candidates using IR methods
 - augmented with ontologies and semi-structured data
- Score each candidate using richer knowledge sources
 - Lat-long databases
 - Temporal reasoning
 - Taxonomical classification

IBM Watson



Ref: <https://medium.com/@giacaglia/how-ibm-watson-works-40d8d5185ac8>

A storytelling about IBM Watson

- The Jeopardy!



A storytelling about IBM Watson

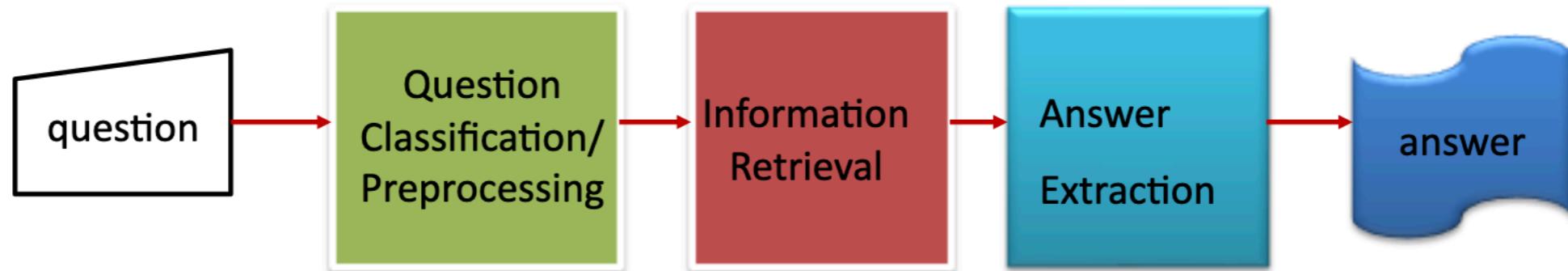
- The Jeopardy!
- Jeopardy! is a quiz competition where a clue in the form of an answer and the contestants have to form a question.
- For example, a clue could be: “He likes his martinis shaken, not stirred.”
- The correct response, phrased in the form of a question, would be, “Who is James Bond?”

A storytelling about IBM Watson

- To educate Watson,
 - Hadoop was utilized to process various data sources such as encyclopedias, dictionaries, news feeds, literature, and the entire contents of Wikipedia.

A storytelling about IBM Watson

- Automatic question-answering system: The basic



e.g.

*What is
Calvados?*

/Q is /A
where:/Q="(Calvados)"

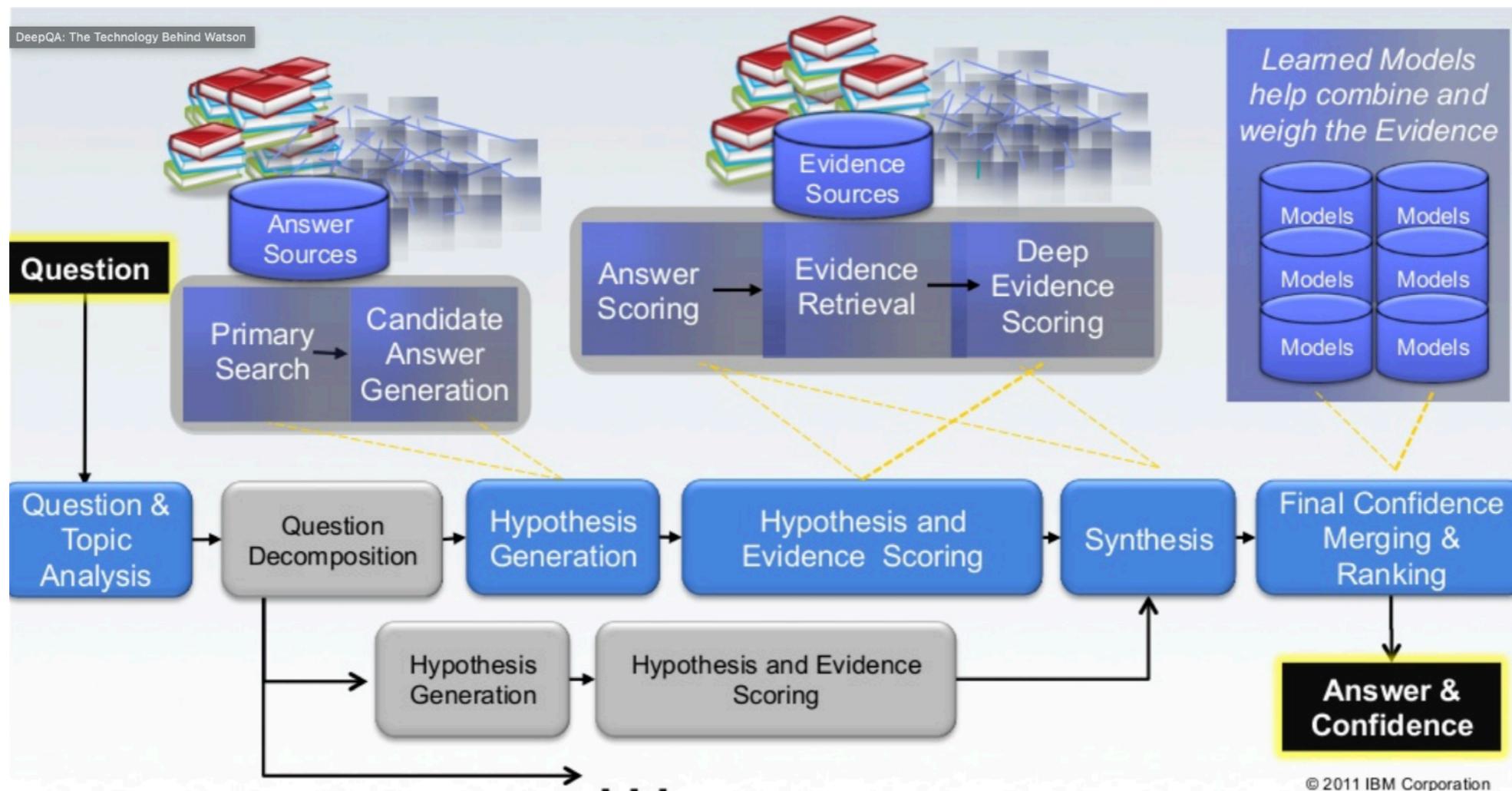
Query="Calvados is"
Text retrieved
I="...Calvados is often used in cooking...Calvados is a dry apple brandy made in..."

/A is : a dry
apple
brandy

Answer:
/Q is /A:
"Calvados"
" is "a dry
apple
brandy"

A storytelling about IBM Watson

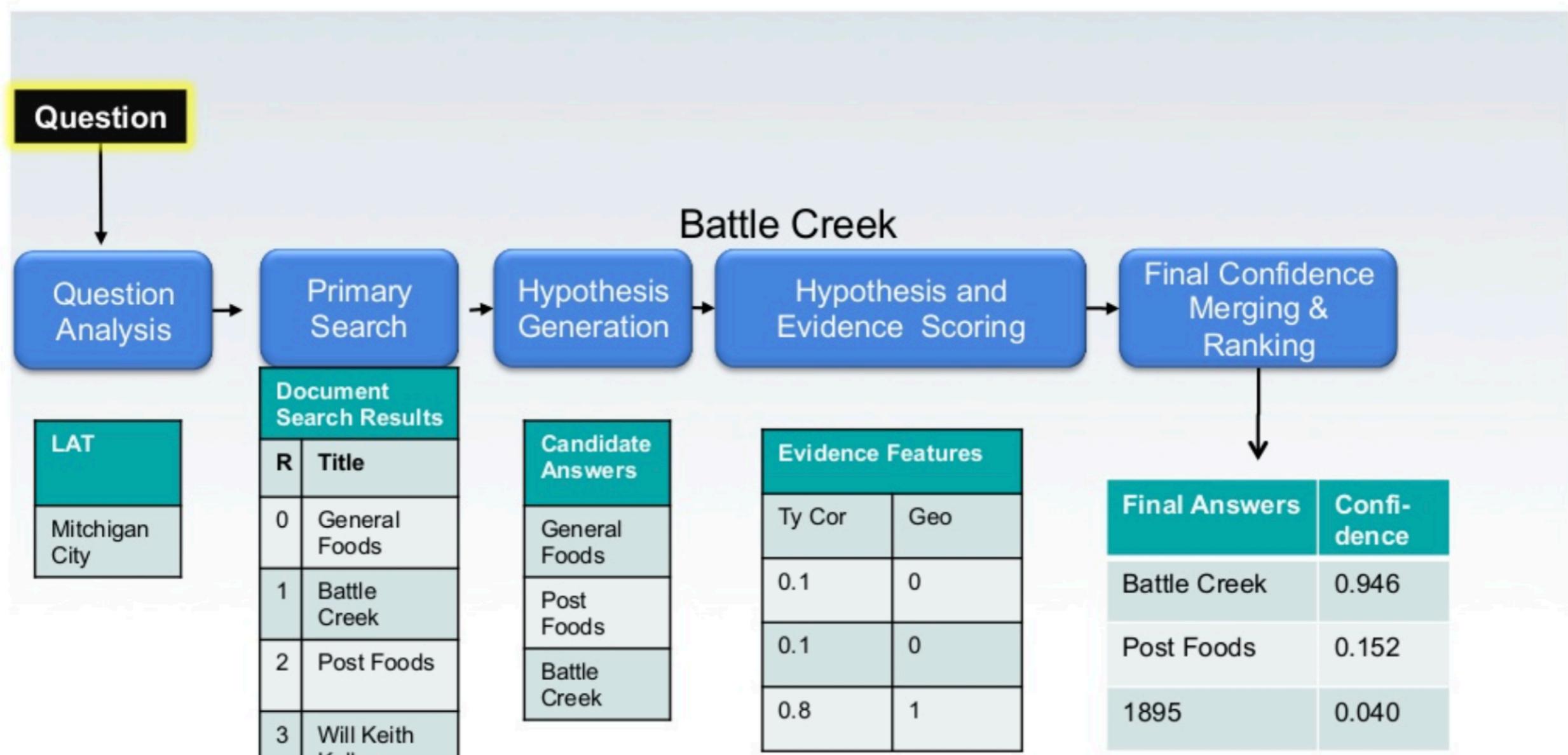
● Inside Watson



A storytelling about IBM Watson

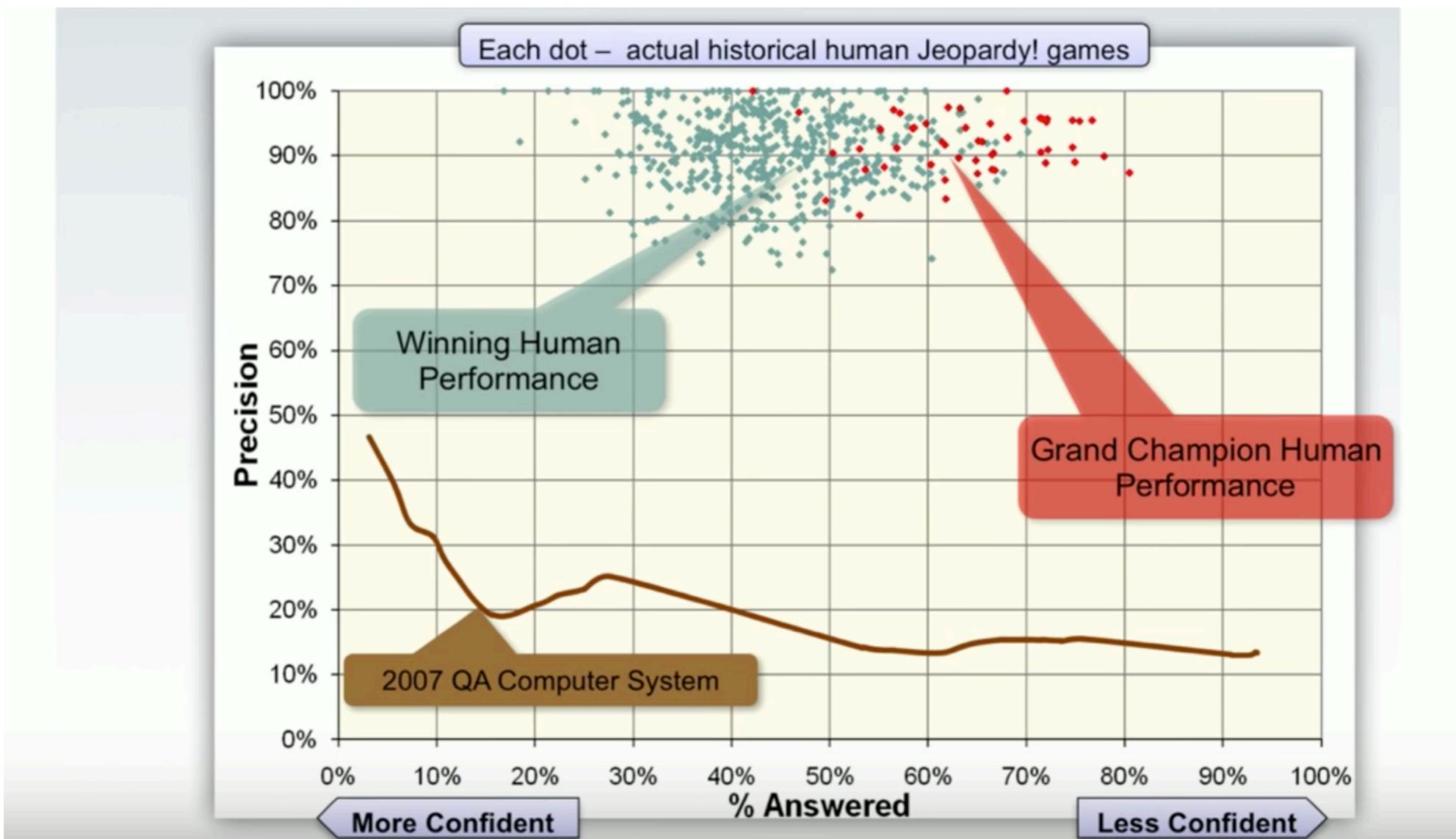
Category: MICHIGAN MANIA

Clue: In 1894 C.W. Post created his warm cereal drink Postum in this Michigan city



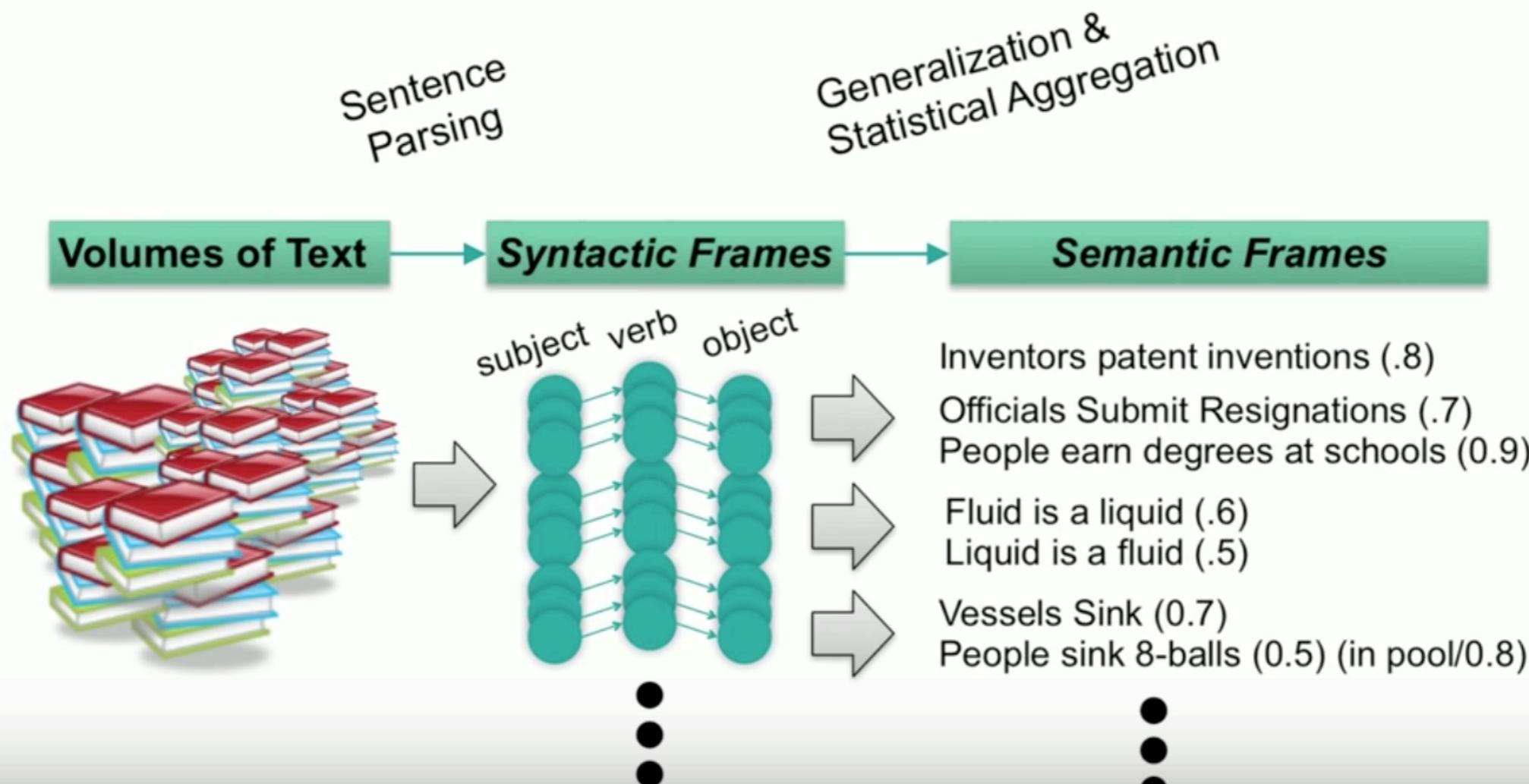
A storytelling about IBM Watson

- Back in 2007



A storytelling about IBM Watson

Automatic Learning From “Reading”



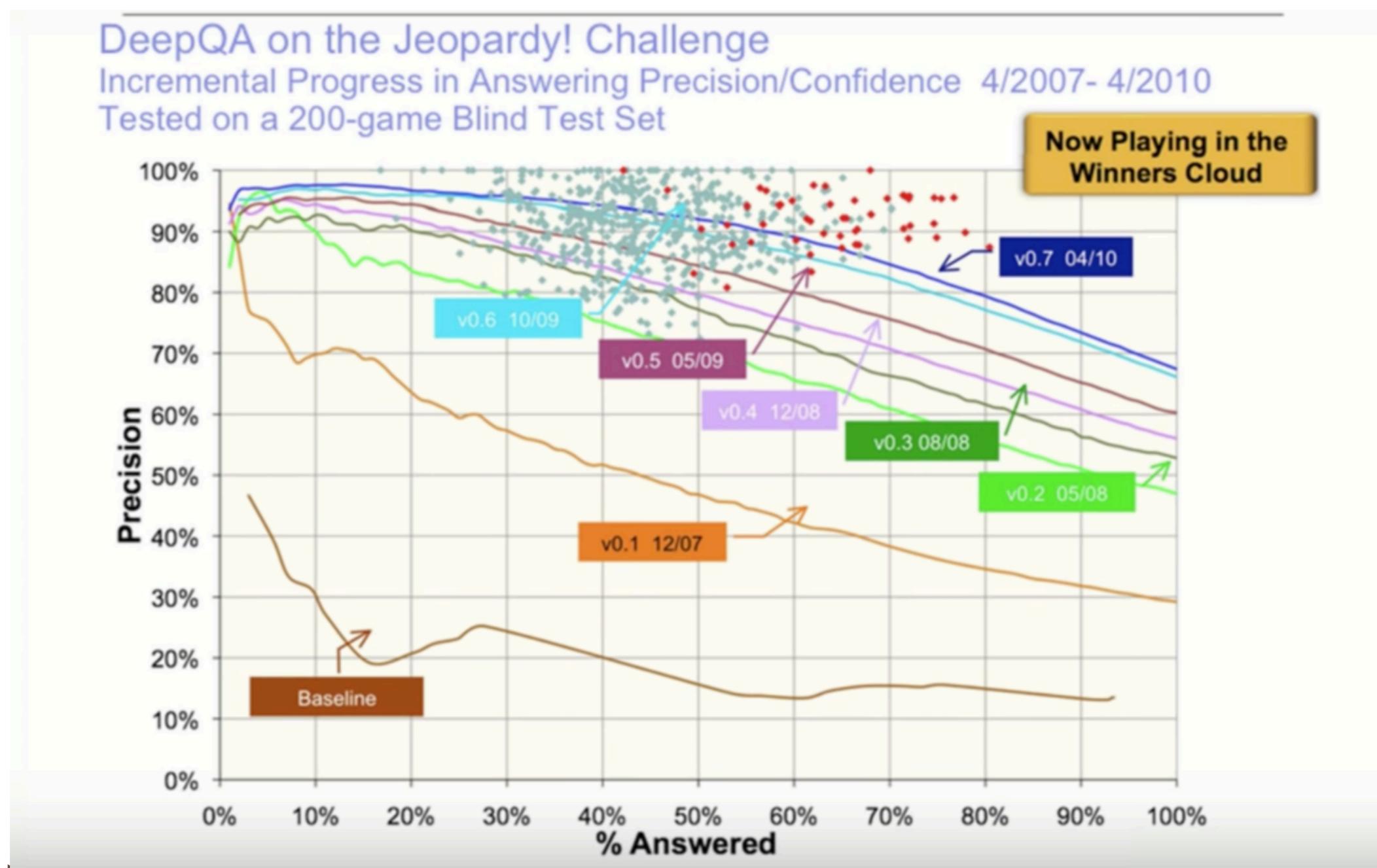
A storytelling about IBM Watson

Category: Celebrations of the Month

Clue	Type	Watson's Answer	Correct Answer	
D-Day anniversary & Magna Carta Day	day	Runnymede	June	✗
National Philanthropy Day & All Soul's Day	day	Day of the Dead	November	✗
National Teacher Day & Kentucky Derby Day	day/month(0.2)	Churchill Downs	May	✗
Administrative Professionals Day & National CPAs Goof-Off Day	day/month(0.6)	April	April	✓
National Magic Day & Nevada Admission Day	day/month(0.8)	October	October	✓

A storytelling about IBM Watson

- The continual improvement



A storytelling about IBM Watson

- In 2011, IBM Watson participated in the Jeopardy against two of the best Jeopardy champions in the show's history and Watson was able to defeat the two human champion.

A storytelling about IBM Watson

- For each clue provided during the game, Watson had to perform the following tasks in less than three seconds:
 - Deconstruct the provided clue into words and phrases.
 - Establish the grammatical relationship between the words and the phrases.
 - Create a set of similar terms to use in Watson's search for a response.
 - Use Hadoop to coordinate the search for a response across terabytes of data.
 - Determine possible responses and assign their likelihood of being correct.
 - Actuate the buzzer.
 - Provide a syntactically correct response in English

A storytelling about IBM Watsons

- In recent days
 - Among other applications, Watson is being used in the medical profession to diagnose patients and provide treatment recommendations.

Time for questions