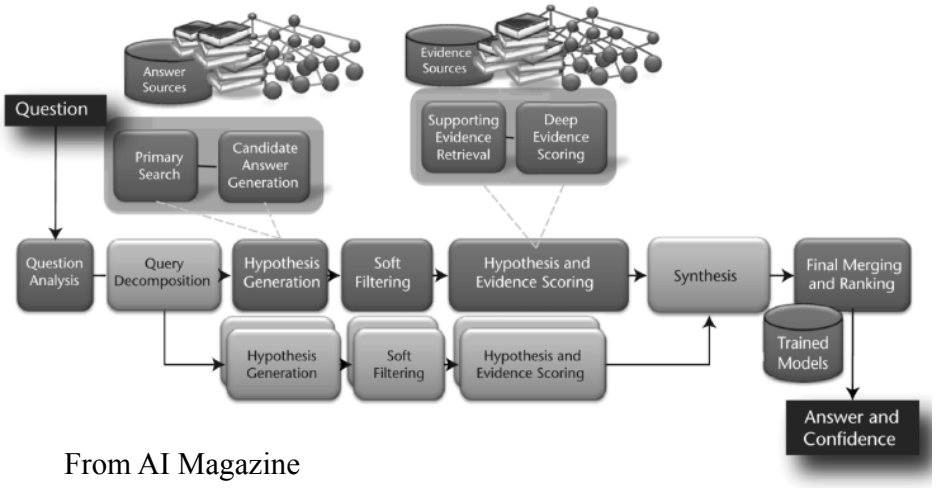
Question Answering

- **Overview and task definition**
- **History**
- **Open-domain question answering**
- **Basic system architecture**
  - Watson's architecture
- **Techniques**
  - Predictive indexing methods
  - Pattern-matching methods
  - Advanced techniques

Statistics

Development Team	25 people
Project Duration	4 years
Software	1,000,000+ SLOC
	700K Java, 300K C++, plus other bi
	~ 130 components
Hardware	90 IBM Power 750 servers
	2880 Power7 cores @ 80+ TFLOPS
	20 TB memory
	10 Gbps network

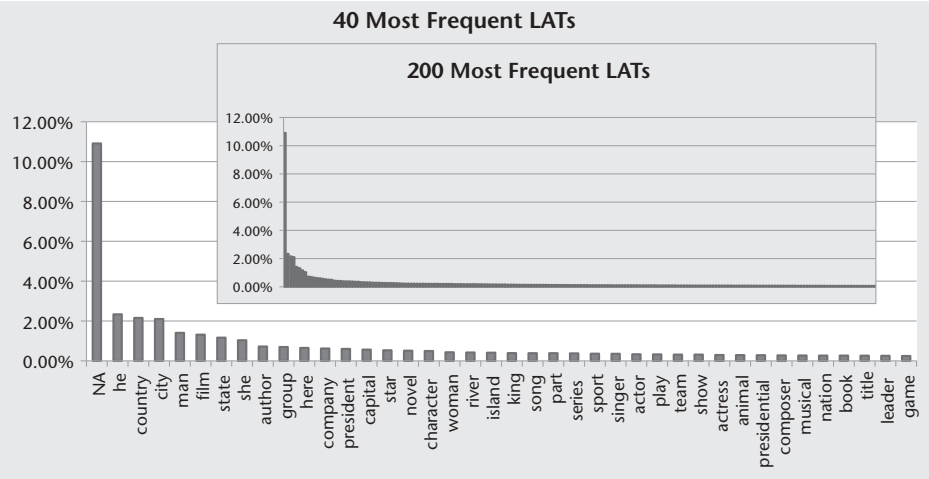
Watson's architecture



From AI Magazine

Question Analysis

- **Identify question type**
- **Determine if decomposition is needed**
- **Determine the "lexical answer type"**
- **Determine "focus" of question**
  - "When hit by electrons, a phosphor gives off elec-tromagnetic energy in **this form**"
  - "Secretary Chase just submitted **this** to me for the third time; guess what, pal. This time I'm accepting it."
- **Relation detection**
  - "They're the two states you could be reentering if you're crossing Florida's northern border"
    - borders(Florida,?x,north)



Decomposition

Category: "Rap" Sheet
Clue: This archaic term for a mischievous or annoying child can also mean a rogue or scamp.

Subclue 1: This archaic term for a mischievous or annoying child.

Subclue 2: This term can also mean a rogue or scamp. Answer: Rapscallion

Decomposition

Category: Diplomatic Relations
Clue: Of the four countries in the world that the United States does not have diplomatic relations with, the one that's farthest north.

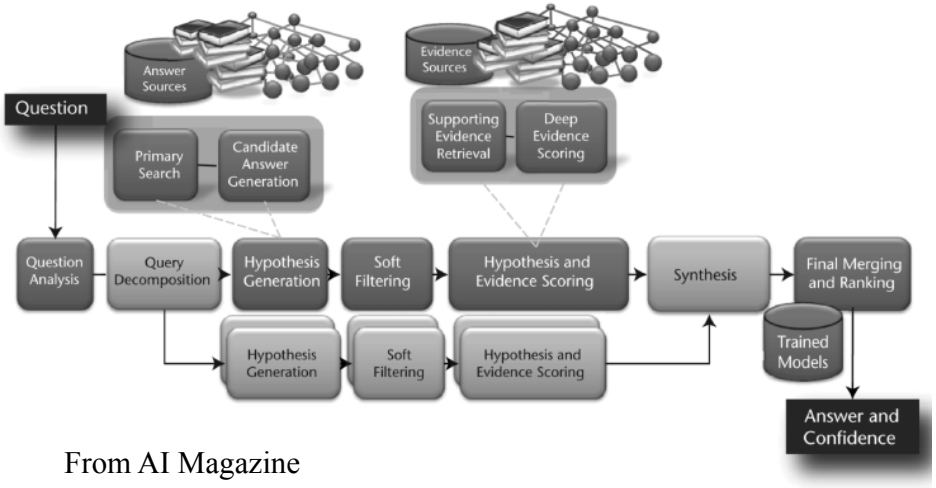
Inner subclue: The four countries in the world that the United States does not have diplomatic relations with (Bhutan, Cuba, Iran, North Korea).

Outer subclue: Of Bhutan, Cuba, Iran, and North Korea, the one that's farthest north. Answer: North Korea

Content acquisition

- **Depends on answer types**
  - LATs
- **Wide range of**
  - encyclopedias
  - Dictionaries
  - Thesauri
  - Newswire articles
  - Literary works
  - Taxonomies, ontologies, WordNet
- **Automatic corpus expansion**

Watson's architecture

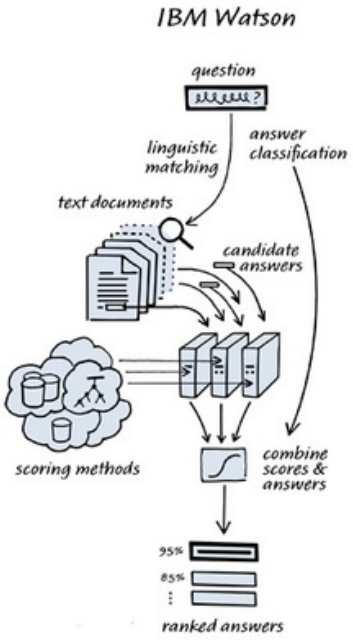


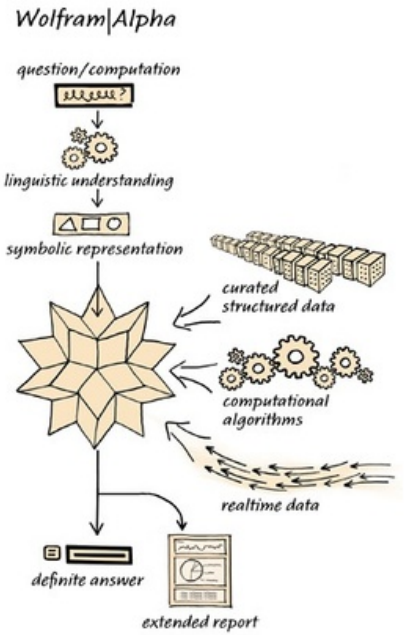
From AI Magazine

The Rest

- **Primary search**
  - Top 250 candidates
- **Candidate answer generation**
  - Extracts the answer from the text/passage/db entry
- **Soft filtering**
  - A bit mysterious…whittle down to top 100
- **Hypothesis and evidence scoring**
  - "rigorous evaluation process"
- **Final merging and ranking**
  - Uses many scoring models
  - Many are question-type-specific

- The Research Team
- The Algorithms Team





Question answering

- **Overview and task definition**
- **History**
- **Open-domain question answering**
- **Basic system architecture**
  - Watson's architecture
- **Techniques**
  - Predictive indexing methods
  - Pattern-matching methods
  - Advanced techniques

Indexing with predictive annotation

- **Some answers belong to well-defined semantic classes**
  - People, places, monetary amounts, telephone numbers, addresses, organizations
- **Predictive annotation: index a document with "concepts" or "features" that are expected to be useful in (many) queries**
  - E.g. people names, location names, addresses, etc.

Predictive annotation

In the early part of this century, the only means of transportation for travelers and mail between <LOCATION> Europe </LOCATION> and <LOCATION> North America </LOCATION> was by passenger steamship. By <DATE> 1907 </DATE>, the <COMPANY> Cunard Steamship Company </COMPANY> introduced the largest and fastest steamers in the <LOCATION> North Atlantic </LOCATION> service: the <NAME> Lusitania </NAME> and the <NAME> Mauritania </NAME>. Each had a gross tonnage of <WEIGHT> 31,000 tons </WEIGHT> and a maximum speed of <SPEED> 26 knots </SPEED>.

– From K. Felkins, H.P. Leighly, Jr., and A. Jankovic. "The Royal Mail Ship Titanic: Did a Metallurgical Failure Cause a Night to Remember?" *JOM*, 50 (1), 1998, pp. 12-18.

Advantages and disadvantages

+ **Most of the computational cost occurs during indexing**
  - Allows use of more sophisticated methods
+ **Annotator has access to complete text of document**
  - Important for recognizing some types of features
– **Must know ahead of time which types of concepts are likely to be important**
– **Increases size of index considerably**
  - E.g. by an order of magnitude if many features

Used (in varying amounts) by almost all open-domain Q/A systems

Simple pattern-based QA

- **Observation: there are many questions… but fewer types of questions**
- **Each type of question can be associated with**
  - Expectations about answer string characteristics
  - Strategies for retrieving documents that might have answers
  - Rules for identifying answer strings in documents

Question answering

- **Overview and task definition**
- **History**
- **Open-domain question answering**
- **Basic system architecture**
  - Watson's architecture
- **Techniques**
  - Predictive indexing methods
  - Pattern-matching methods
  - Advanced techniques

Example

- **Who is the President of Cornell?**
  - Expectation: answer string contains a person name
    - Named entity identification
  - Search query: "president Cornell *PersonName"
  - Rule: "*PersonName, President of Cornell"
    - Matches "…David Skorton, President of Cornell"
    - Answer = "David Skorton"

Question analysis

- **Input: the question**
- **Output**
  - Search query
  - Answer expectations
  - Extraction strategy
- **Requires**
  - Identifying named entities
  - Categorizing the question
  - Matching question parts to templates
- **Method: pattern-matching**
  - Analysis patterns initially created manually…

Question analysis example

- **"Who is Elvis?"**
  - Question type: "who"
  - Named-entity tagging: "Who is <person-name>Elvis</person-name>"
  - Analysis pattern: if question type = "who" and question contains <person-name> then
    - Search query doesn't need to contain a *PersonName operator
    - Desired answer probably is a description
    - Likely answer extraction patterns
      - "Elvis, the X"
        - » "…Elvis, the king of rock and roll…"
      - "the X Elvis"
        - » "the legendary entertainer Elvis"

Simple pattern-based Q/A: assessment

- **Extremely effective when**
  - Question patterns are predictable
    - Fairly "few" patterns cover the most likely questions
      - Could be several hundred
    - Not much variation in vocabulary
      - Simple word matching works
    - The corpus is huge (e.g., Web)
      - Odds of finding an answer document that matches the vocabulary and answer extraction rule improves
- **Somewhat labor intensive**
  - Patterns are created and tested manually

Common problem: improving answer extraction patterns

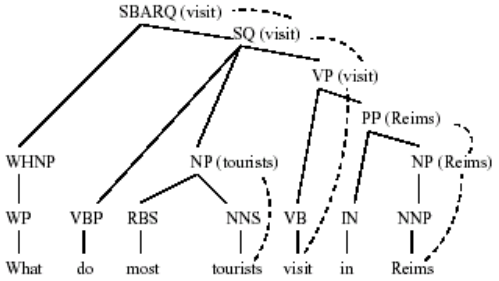
- **Word sequence patterns have limited power**
- **Solution: create patterns that use syntactic information**
  - Partial syntactic parsing of documents
    - Is this noun the subject or the object of the sentence?
  - Allows more complex patterns
    - Question: "Who shot Kennedy?"
    - "Who" implies a person that should be subject of answer sentence/clause
    - "Kennedy" should be direct object of answer
    - Pattern: <subject> shot Kennedy
    - Matching text: "Oswald shot Kennedy"

Question answering

- **Overview and task definition**
- **History**
- **Open-domain question answering**
- **Basic system architecture**
  - Watson's architecture
- **Techniques**
  - Predictive indexing methods
  - Pattern-matching methods
  - Advanced techniques

Question analysis

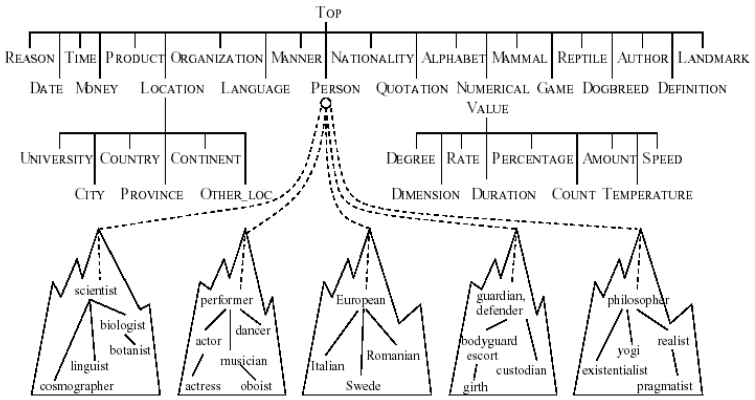
- **Parsing and named entity recognition**
- **Expected answer type determined by parsing**



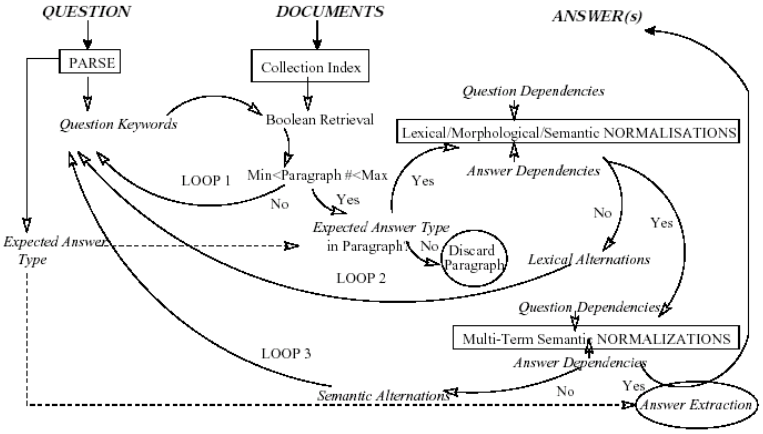
- **Exceptions for "special cases"**

(Q-P1): *What {is|are} <phrase_to_define>?*
(Q-P2): *What is the definition of <phrase_to_define>?*
(Q-P3): *Who {is|was|are|were} < person_name(s)>?*

Expected answer types



Feedback loops



Answer verification

- **Parse passages to create a dependency tree among words**
- **Attempt to unify logical forms of question and answer text**