

Watson

Elizabeth Hines

Kevin Riden

Introduction

- * Machine to compete on Jeopardy
- * Simple description
 - * Question-answer system with natural language processing
- * IBM's description
 - * An application of advanced Natural Language Processing, Information Retrieval, Knowledge Representation and Reasoning, and Machine Learning technologies to the field of open domain question answering

Watson & AI

- * Searching
 - * Watson does searching, but IBM doesn't share exactly how
 - * DeepQA is the overall term for their search
- * Formalism used to present knowledge
 - * IBM hides this level of detail
- * Why is this an AI project?
 - * Natural language processing coupled with computing confidence levels and finally deciding on an answer
- * Pitfall
 - * “IBM invented an ingenious program—not a computer that can think” –John Searle

System Details

* Hardware

- * Cluster of 90 IBM Power 750 servers in 10 racks
- * Totals of 2880 POWER7 processor cores and 16 Terabytes of RAM
- * Each Power 750 server uses a 3.5 GHz POWER7 eight core processor
- * Process 500 gigabytes, the equivalent of a million books, per second
- * Hardware cost about \$3 million

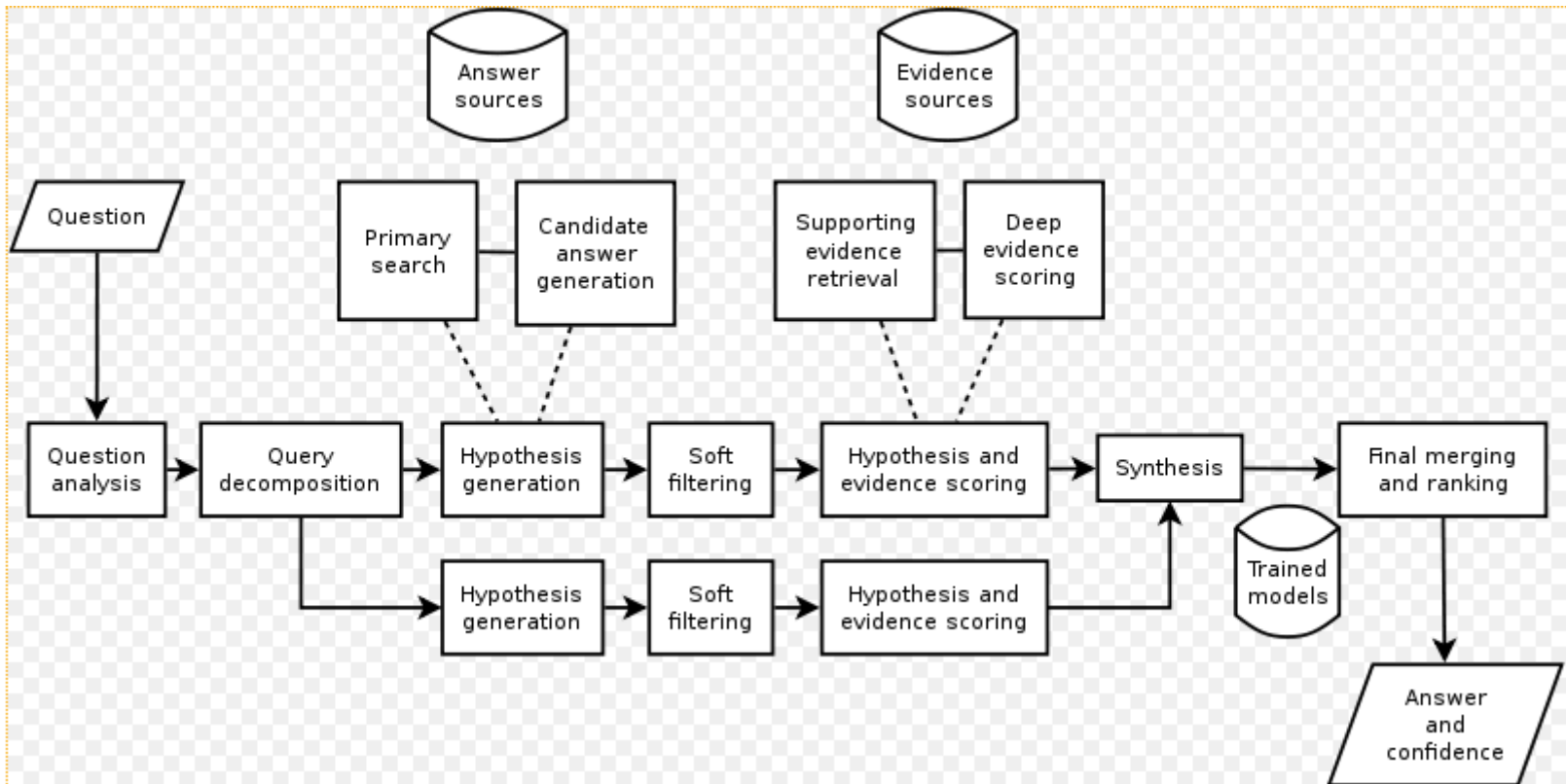
* Software

- * At least Java, C++, and Prolog
- * Uses Apache Hadoop framework for distributed computing
- * Apache UIMA (Unstructured Information Management Architecture) framework
- * IBM's DeepQA
- * SUSE Linux Enterprise Server 11 operating system

* Data

- * 200 million pages of structured and unstructured content
- * 4 terabytes of disk storage

System Details



Demonstration

- * Overall System Architecture
 - * <http://www-03.ibm.com/innovation/us/watson/building-watson/how-watson-works.html>
- * Answering a Question
 - * <http://www-03.ibm.com/innovation/us/watson/what-is-watson/science-behind-an-answer.html>
- * Playing Jeopardy
 - * http://www.youtube.com/watch?v=WFR3lOm_xhE

Contributions

Healthcare

Industry

Education

Customer Service