

Trustworthiness and relevance in Web-based clinical question answering

Sarah Cruchet, Lonneke van der Plas, Célia Boyer
Health On the Net Foundation

Presented by Allan Hanbury for the KHRESMOI project
MIE, Pisa, August 29, 2012

- Why a question answering system (QA)?
- Why in the Health domain?
- What is a restricted QA?
- Research question
- Structure of the QA
- Method
- Results and Conclusion

Why a question answering system?

- Full-fledged question in natural language and not just keywords
- Precise answer to the user's questions and not a list of relevant documents
- Adapted to devices with limited size such as smart phones
- Reduced time for the search

Most famous QA system which won the Jeopardy game in 2011.

IBM's Watson Computer Beats the Superstars of Jeopardy! But What Does it Mean?



Michio Kaku on February 17, 2011, 12:43 PM

A milestone of sorts was passed this week when IBM's Watson supercomputer bested the two human superstars of the Jeopardy! TV program, answering questions that would stump an ordinary person. Tensions were high during this contest, since the stakes are potentially very great. This contest harkens back to the famous test mentioned by the AI pioneer Alan Turing, who said that, sooner

Why in the health domain?

- Many health documents on the Internet. The answer exists but no system proposes direct answers to a given query
- The health domain is very important for the citizen. 1 to 5 patients use the Internet to search for health information
- Relevance and trustworthiness of the answers is crucial for the health domain

Restricted QA system

- Intended for the General public
- Only trustworthy websites which have been HONcode certified
- Use the semantic resources UMLS (Unified Medical Language system)
- Based on the redundancy found in large document collections
- Comparative evaluation where the same QA is given access to either the whole World Wide Web or only certified websites

- Hypothesis: HONcode certified websites should be of better quality than the average websites found on the web, and therefore give more relevant answers.
- Does limiting the document collection based on trustworthiness hurt the performance of a clinical QA system in term of relevance?

Method: presentation of the QA

- Question analysis module (uses semantic resources such as UMLS)
- Query generation module:
 - Translate the question formulated in natural language into a query that can be processed by the search engine
 - Expands query with synonyms
- Document retrieval module:
 - Collects documents
 - Analyses documents semantically
 - Extracts medical terms and semantic types to be used by the answer extraction module
- Automatic Knowledge extraction module:
 - Passage extraction module
 - Answer extraction module
 - Ranking of the best answers

Question:

What is hypertension?

Question analysis
module

Query generation
module

Restricted to

Standard information
retrieval module



Knowledge
extraction module

Best Answers

- Study based on the methodology developed for a study conducted at the U.S. National Library of Medicine (NLM) on information retrieval systems
- Comparison of two systems:
 - Certified_only: restricted to only the database of 6 800 HONcode certified health and medical websites
 - Unrestricted: access the World wide Web via Google results
- Only the first 10 answers for each question are taken into account
- A medical expert graded the answers:
 - A+ (very relevant and reliable), A (relevant and reliable), A- (not the whole answer)
 - B+ (leading to response), B (may lead to the answer), B- (unclear)
 - C (not relevant)

Results

Metric	Certified only	Unrestricted
MAP	0.59	0.36
Bpref	0.50	0.34
R-prec	0.59	0.38
MRR	0.76	0.86
P@5	0.54	0.36
P@10	0.32	0.22

- MAP is higher for the restricted system
- Out of the first 5 answers found by the restricted system, 54% are relevant against only 36% for the unrestricted system

100 %

Answers:	Language:	Expected question type:	Expected medical type:	Rate answers: Is it appropriate or not ?	
141	English	definition	disease		
<p>1: Answer from www.ehealthmd.com Diabetes mellitus, commonly known as diabetes, is a disorder in the way the body gets energy from food.</p>				✓	✗
<p>2: Answer from www.medicinenet.com Diabetes mellitus, commonly referred to as diabetes (as it will be in this article) was first identified as a disease associated with "sweet urine," and excessive muscle loss in the ancient world.</p>				✓	✗
<p>3: Answer from www.medicinenet.com Diabetes mellitus is a group of metabolic diseases characterized by high blood sugar (glucose) levels, that result from defects in insulin secretion, or action, or both.</p>				✓	✗
<p>4: Answer from diabetes.niddk.nih.gov Diabetes is a disorder of metabolism—the way the body uses digested food for growth and energy.</p>				✓	✗
<p>5: Answer from www.lifeclinic.com For reasons that are not yet clear, diabetes is increasing in our population to the point where public health authorities are calling diabetes an "epidemic" that requires urgent attention.</p>				✓	✗
<p>6: Answer from diabetes.niddk.nih.gov Diabetes is a disease in which levels of blood glucose, also called blood sugar, are above normal.</p>				✓	✗
<p>7: Answer from www.ehealthmd.com Diabetes affects the way the body uses food.</p>				✓	✗
<p>8: Answer from www.womenshealthmatters.ca Diabetes Mellitus is the name given to a common, chronic health condition, characterized by above-normal levels of glucose (a type of sugar) in the blood.</p>				✓	✗

Conclusion and next steps

- Trustworthiness of answers returned to the user is an important issue in clinical QA
- Limiting the document collection to HONcode-certified health websites does not alter the performance of our clinical QA system
- Integration of the restricted QA in a health search engine in order to combine a keyword based and the NL based queries:
KHRESMOI project



Thank you!

CELIA.BOYER@HEALTHONNET.ORG

WWW.KHRESMOI.EU

WWW.HEALTHONNET.ORG

