FindZebra: A search engine for rare diseases

Background: The web has become a primary information resource about illnesses and treatments for both medical and non-medical users. Standard web search is by far the most common interface for such information. It is therefore of interest to find out how well web search engines work for diagnostic queries and what factors contribute to successes and failures. Among diseases, rare (or orphan) diseases represent an especially challenging and thus interesting class to diagnose as each is rare, diverse in symptoms and usually has scattered resources associated with it.

Methods: We use an evaluation approach for web search engines for rare disease diagnosis which includes 56 real life diagnostic cases, state-of-the-art evaluation measures, and curated information resources. In addition, we introduce FindZebra, a specialized (vertical) rare disease search engine. FindZebra is powered by open source search technology and uses curated freely available online medical information.

Results: FindZebra outperforms Google Search in both default setup and customised to the resources used by FindZebra. We extend FindZebra with specialized functionalities exploiting medical ontological information and UMLS medical concepts to demonstrate different ways of displaying the retrieved results to medical experts.

Conclusions: Our results indicate that a specialized search engine can improve the diagnostic quality without compromising the ease of use of the currently widely popular web search engines. The proposed evaluation approach can be valuable for future development and benchmarking. The FindZebra search engine is available at http://www.findzebra.com/.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Technical University of Denmark, Royal School of Library and Information Science, Copenhagen University Hospital
Pages: 528–538
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: International Journal of Medical Informatics
Volume: 82
Issue number: 6
ISSN (Print): 1386-5056
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 3.56 SJR 1.247 SNIP 1.85
Web of Science (2017): Impact factor 2.957
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 3.7 SJR 1.188 SNIP 2.01
Web of Science (2016): Impact factor 3.21
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 3.46 SJR 1.321 SNIP 2.157
Web of Science (2015): Impact factor 2.363
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 2.85 SJR 0.992 SNIP 1.862
Web of Science (2014): Impact factor 2.004
BFI (2013): BFI-level 2
Scopus rating (2013): CiteScore 3.81 SJR 1.365 SNIP 2.261
Web of Science (2013): Impact factor 2.716
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): CiteScore 3.22 SJR 1.115 SNIP 2.181