

TOWARDS AN OPEN, TRUSTWORTHY AND MULTILINGUAL SEARCH ENGINE FOR MEDICAL PRACTITIONERS: NEWS FROM THE EUROPEAN KHRESMOI PROJECT

Samwald M^{1,2,3}, Kritz M⁴, Gschwandtner M⁴, Stefanov V²,
Hanbury A²

Abstract

Medical practitioners often have unmet information needs that impact patient care. However, currently available web-based search engines are not suitable for routine use. Finding relevant information takes too long, assessing the trustworthiness of found information is difficult, and support for the heterogeneity of languages and nomenclature across European countries is lacking. In this paper, we analyze the current barriers to web-based searching by medical practitioners and introduce the European Khresmoi project, which aims to dismantle these barriers.

Keywords – *Medical Practice, Information Retrieval, Semantic Technologies, Decision Support, Multilinguality*

1. Background

Physicians often have unmet information needs. These have been reported as occurring for up to 2 of every 3 patients seen [1], or more recently for 41% of the questions they pursued [2]. Although these medical professionals have many tools for information search available (such as PubMed), studies have revealed that they do not use them to their full capabilities.

Most questions arise during consultations and have a direct impact on the medical decision process [3]. There is evidence suggesting that physicians primarily respond to their information needs when they perceive the question to be urgent and believe that definitive answers can be found [4]. However, physicians are often restricted in their search by time constraints [5]. Physicians search on average for less than 5 minutes to answer questions [6]. A so-called “90 second rule” has been described in the literature — meaning physicians do not even attempt to find information unless they think they can do it in a minute and a half [7]. Hence, it is important that the pertinent information is found during this time. However, the time taken to answer questions using PubMed averages 30 minutes [1] and the information found is often scattered over multiple articles, making

1 Section for Medical Expert and Knowledge-Based Systems, Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Austria

2 Information Retrieval Facility, Vienna, Austria

3 Digital Enterprise Research Institute (DERI), National University of Ireland Galway, Ireland

4 Society of Physicians in Vienna (Billrothhaus), Vienna, Austria

PubMed searching impractical for routine clinical use [6]. Furthermore, physicians that are not native English speakers using systems in English language are prone to use erroneous search terms, resulting in poorer returned results [8].

The World Wide Web has a lot to offer in terms of both quantity and quality of medical information [9]. There is no consensus within the literature as to what extent doctors currently rely on web-based searching as compared to other information sources. One line of research suggests that physicians often find it quicker and easier to look up answers in a pocket reference book or ask a colleague for advice [10] rather than searching on the Internet. In addition current web-based solutions fail to provide psychological support, guidance, affirmation, sympathy, judgement, and feedback, which colleagues can provide within the daily decision making process of a physician. A review by Davies [3] compared relevant research between 2000 and 2005, and found that text books (39%) and colleagues (25%) were the information sources physicians accessed most frequently, while computer resources were used only by 13%. However, an upward trend of Internet use is visible as the highest percentage of use examined in the study, 53%, occurred in the latest published research from 2005. A Spanish study published in 2007 [11] found that the majority of physicians still relied on colleagues, drug compendiums and textbooks rather than on web-based resources. A possible explanation is language as a potential barrier to web-based searching, and it appears to be inadequately addressed by current web-based solutions within the medical domain. In contrast to these findings is research claiming a clear preference of the Internet as a primary informational resource amongst physicians [12],[13]. Both studies provide support for the notion that the Internet has become an important information source amongst physicians. Possible explanations of conflicting study outcomes could be the variance among medical specialities, different geographical locations and potential biases introduced by different methodologies used in the studies (e.g., user observations versus self reports). Furthermore, the rapid changes and advances in the field of information technology make comparisons over larger time spans difficult.

There is also conflicting data about which web sites and tools physicians use to look for medical information on the web. Some publications suggest that general-purpose search services such as Google can play a useful role in the medical decision making process [14],[15]. In contrast, Leo et. al. [16] reported that physicians mistrust the quality of results from such search engines and prefer to directly access specialized medical websites. A study by Yu and Kaufman [14] suggests that Google is preferred for finding medical definitions, as it is easy to use and provides good answers to simple questions. However, for more complex information needs more advanced search systems may be required. A recent study funded by Google [13] is in strong contrast with prior findings. It postulates that the majority of physicians use Google or a similar search engine as their primary information sources in the clinical decision making process. However, it is unclear as to what extent the study was biased in terms of sample selection. Another study showed that general practitioners use Google as their first information source, primarily to lead them to higher quality websites [17]. Thus, it appears that physicians are currently willing to use a search engine for simple questions and as an initial source to help them find their way to higher quality websites.

2. The Khresmoi approach

Summarizing the issues described above, medical practitioners could benefit from a widely deployed search engine that:

- provides relevant, summarised output, yielding answers to complex questions in just a few minutes

- is openly available on the web without barriers (e.g., requirement of subscription, registration fees)
- is multilingual and supports terminologies (e.g., drug brand names in various countries)
- is fast, easy to use and can be viewed on mobile devices
- is developed in close cooperation with medical practitioners
- is able to deal with short and underspecified queries
- is continuously kept up-to-date
- responds to the need for psychological support and affirmation

It is the aim of the Khresmoi project⁵ to design and build a search engine that meets these criteria. The project is executed by a consortium of 12 academic and industrial partners and has a budget of approximately € 10 million. It started in September 2010 and will run for four years. Incremental versions will be made available during the project. Results of interviews and questionnaires about search engine use and information needs of European medical practitioners will be available by mid-2011. A first prototype will be released after one year, the final stable version of the infrastructure will be released in 2014.

Khresmoi aims to meet the needs of a variety of users (Fig. 1): Members of the general public that do not have any medical expertise, as well as medical practitioners within their specialities, and radiologists (which form a separate user group because of their distinct information needs).



Figure 1: The Khresmoi infostructure

The prototype system is based on existing software tools developed by the consortium partners. The GATE framework [18] is used for natural language processing, information extraction and the automated annotation of biomedical entities mentioned in websites and articles. The crawling, processing and querying of several terabytes of medical literature from web sources is done by the Vienna-based Information Retrieval Facility⁶ (IRF). To query both unstructured and semi-structured, annotated text, the *Mimir* information retrieval engine⁷ is employed. Mimir is able to recognize and normalize linguistic variants, different units of measurement, annotations and document sections. It provides a sophisticated query language based on these text features. Mimir is interlinked with and complemented by the BigOWLIM semantic repository⁸, a highly scalable storage and reasoning system for data in RDF/OWL9 format. BigOWLIM is used to house a large medical knowledge base composed of biomedical ontologies (such as those from the Unified

5 <http://khresmoi.eu>

6 <http://www.ir-facility.org/>

7 <http://gate.ac.uk/family/mimir.html>

8 <http://www.ontotext.com/owlim/big/>

9 <http://www.w3.org/TR/owl2-primer/>

Medical Language System, UMLS¹⁰) and relevant medical and pharmacological datasets harvested from the web of data .

The quality of medical websites used by the Khresmoi system is ensured by Health on the Net¹¹ (HON). HON is a non-governmental organization that gives a certificate (the 'HONcode') to trustworthy medical websites, based on a rigorous check of the website content. The HONcode is among the most widely acknowledged quality certificates for medical information on the web.

The project has a dedicated budget of around 120,000 Euros for the manual creation of annotations on professional medical literature, which will add rich contextual information and raise the data quality.

The ability to query over an integrated database of annotated text from medical websites and literature, biomedical ontologies and well-structured biomedical datasets is expected to improve the quality of search results, and to reduce the time that is needed to find relevant, valid answers to medical questions.

Of course, such a system needs to be designed in close collaboration with the targeted group of end users. The Society of Physicians in Vienna¹² is a member of the Khresmoi consortium and responsible for guiding project developments according to the needs of European medical practitioners. The society has over 2400 members and around 200 new members join each year. The website of the society is one of the leading Austrian websites for continuing medical education. It offers an online library, courses, webcasts, podcasts, as well as mobile services for physicians. In the Khresmoi project, the Society of Physicians conducts a large-scale user study to gather current data about the use of web-based information sources by medical practitioners in Austria and other European countries, employing both qualitative and quantitative methodologies (interviews and questionnaires). The study will serve to gather user requirements for the development of the Khresmoi system. Furthermore, the society will evaluate and test the initial and final versions of the Khresmoi search engine amongst its members.

3. Conclusions

The Khresmoi project is at an early stage. In order to increase the uptake of web-based search systems by medical practitioners and to have a meaningful impact on medical decision making, the development of the system needs to be done in close collaboration with medical practitioners. The community of physicians in Vienna could spearhead the uptake of web-based decision support in the daily medical routine.

4. Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 257528 (KHRESMOI).

10 <http://www.nlm.nih.gov/research/umls/>

11 <http://healthonnet.org>

12 <http://www.billrothhaus.at>

5. References

- [1] HERSH, W.R, HICKAM, D.H., How Well Do Physicians Use Electronic Information Retrieval Systems? A Framework for Investigation and Systematic Review, *JAMA: The Journal of the American Medical Association*, vol. 280, Okt. 1998, S. 1347 -1352.
- [2] ELY, J.W., OSHEROFF, J.A., MAVIGLIA, S.M., ROSENBAUM, M.E., Patient-Care Questions that Physicians Are Unable to Answer, *Journal of the American Medical Informatics Association : JAMIA*, vol. 14, 2007, S. 407-414.
- [3] DAVIES, K., The information-seeking behaviour of doctors: a review of the evidence, *Health Information and Libraries Journal*, vol. 24, 2007, S. 78-94.
- [4] GORMAN, P.N., HELFAND, M., Information Seeking in Primary Care, *Medical Decision Making*, vol. 15, Juni. 1995, S. 113 -119.
- [5] SACKETT, D.L., STRAUS, S.E., Finding and Applying Evidence During Clinical Rounds, *JAMA: The Journal of the American Medical Association*, vol. 280, Okt. 1998, S. 1336 -1338.
- [6] HOOGENDAM, A., STALENHOF, A.F., DE VRIES ROBBÉ, P.F., OVERBEKE, A.J.P., Answers to Questions Posed During Daily Patient Care Are More Likely to Be Answered by UpToDate Than PubMed, *Journal of Medical Internet Research*, vol. 10.
- [7] SANDERS, S., Clever searching for evidence, *BMJ*, vol. 330, 2005, S. 1162-1163.
- [8] HOOGENDAM, A., STALENHOF, A.F., DE VRIES ROBBÉ, P.F., OVERBEKE, A.J.P., Analysis of queries sent to PubMed at the point of care: Observation of search behaviour in a medical teaching hospital, *BMC Medical Informatics and Decision Making*, vol. 8, S. 42-42.
- [9] PANDOLFINI, C. und BONATI, M., Follow up of quality of public oriented health information on the world wide web: systematic re-evaluation, *BMJ*, vol. 324, März. 2002, S. 582 -583.
- [10] RAMOS, K., LINSCHIED, R., SCHAFFER, S., Real-time information-seeking behavior of residency physicians, *Family Medicine*, vol. 35, Apr. 2003, S. 257-260.
- [11] GONZALES-GONZALES, A.I., DAWES, M., SANCHEZ-MATEOS, J., RIESGO-FUERTES, R., ESCORTELL-MAYOR, E., SANZ-CUESTA, T., HERNANDEZ-FERNANDEZ, T., Information Needs and Information-Seeking Behavior of Primary Care Physicians, *Ann Fam Med*, vol. 5, Juli. 2007, S. 345-352.
- [12] GRAVATT, Z.B.H., ARROLL, B., What resources do Auckland general practitioners use for answering immediate clinical questions and for lifelong learning?, *Journal of Primary Health Care*, vol. 2, Juni. 2010, S. 100-104.
- [13] HALL AND PARTNERS, Connecting with physicians online: searching for answers., 2009.
- [14] YU H., KAUFMAN D., A cognitive evaluation of four online search engines for answering definitional questions posed by physicians, *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 2007, S. 328-339.
- [15] TANG, H., NG, J.H.K., Googling for a diagnosis—use of Google as a diagnostic aid: internet based study, *BMJ*, vol. 333, Dez. 2006, S. 1143 -1145.
- [16] DE LEO, G., LEROUGE C., CERIANI, C., NIEDERMAN, F., Websites Most Frequently Used by Physician for Gathering Medical Information, vol. 2006, 2006, S. 902-902.
- [17] SIM, M.G., KHONG, E., JIWA, M., Does general practice Google?, *Australian Family Physician*, vol. 37, Juni. 2008, S. 471-474.
- [18] CUNNINGHAM, H., MAYNARD, D., BONTCHEVA, K., TABLAN, V., GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications, 2002.