

Aspects of Medical Information Search

Henning MÜLLER^a, Allan HANBURY^b, Celia BOYER^c

^a*University of Applied Sciences Western Switzerland, Sierre, Switzerland*

^b*Technical University of Vienna, Austria*

^c*Health On the Net Foundation (HON), Geneva, Switzerland*

Abstract. The Internet is increasingly used to find health information worldwide. Online health information search can be beneficial for novice users but due to the overwhelming quantity and uneven quality of online health information it may also be a threat. Although medical professionals are qualified to evaluate information quality, they can be overwhelmed with its quantity and it is impossible to be up to date even in a specialized field. However, the main threat is information quality. The main initiatives in online health information quality are explored, and the HONcode certification is described. The workshop will also detail other aspects of medical information search, text retrieval using ontologies for semantic text analysis and visual retrieval to include images and videos. The EU project KHRESMOI will be presented – it aims at using semantic, textual and visual information retrieval targeted at three main groups: the general population, GPs and radiologists. Time is foreseen for discussion among participants.

Introduction to Medical Information Search

Nothing has changed clinical practice more fundamentally than the Internet. The main reason is that the Internet is available to patients and physicians and this is redefining the roles [1]. For professionals the Internet is primarily an information source, useful for keeping up with the increasing number of medical papers being published [2]. It is also used in difficult cases, for example to “Google for a diagnosis” [3]. For patients, in addition to an information source, the Internet can also be a source for finding people who experience the same health problems. Examples provide anecdotal evidence of positive influence of information found on the Internet, for example in diagnosing rare diseases in children [4]. However, placing such a tool in the hands of patients can potentially be dangerous: entering symptoms into a search engine is perilous for someone prone to hypochondria [1], leading to Cyberchondria [5]. It is thus important that physicians can counsel patients on the use of the Internet for health search.

The two main risks with online health information are: (i) the misinterpretation of information that can be caused by the reader’s health literacy differing from that of the target audience of a site, and (ii) the quality of online health information being often questionable [6]. To combat the latter risk, there several initiatives, including: code of conduct, self-evaluation by users, user-guidance and web site certification [7], such as the HONcode [8]. In medical information search, there are many further challenges to be met in order to improve the access to health information on the Internet, such as: extremely large datasets, inclusion of images and videos in the retrieval [9] and

multilingual access to information [10]. The KHRESMOI project is a 4-year EU funded project that is working towards some of these challenges. Workshop goals are:

- to inform the participants about current tools for medical information search;
- to present tendencies for determining health web site quality and labeling web sites, including a description of the HONcode;
- to present new challenges such image search and access to very large data;
- to present the current possibilities of using semantic data in retrieval;
- to foster the discussion among the participants on topics in information search.

1. Workshop speakers

Célia Boyer is the Executive Director of the Health On the Net (HON) Foundation, Geneva, Switzerland. She is a recognized expert in quality assessment of medical information on the Internet and has taken part in several research projects (PIPS, WRAPIN) and conferences, both on a European and an International level.

Allan Hanbury is Senior Researcher at the Technical University of Vienna, Austria. His research interests include information retrieval, multi-modal information access, and the evaluation of information retrieval results. Currently he is scientific coordinator of the EU project KHRESMOI.

Stefan Darmoni is an advanced technology unit manager in the Computing & Networks Department of Rouen University Hospital and also an associate professor of Medical Informatics, Rouen Medical School, France. His interests include information systems, internet/intranet, decision support, documentary information system, and computer-aided clinical trials management, both from practical and theoretical aspects.

Patrick Ruch is a professor in information sciences at the University of Applied Sciences Geneva, Switzerland. Over the past years he has had several positions in industry and academia working for the medical informatics service at the Geneva University hospitals, IBM research in Zürich and the National Library of Medicine, Bethesda, MD, USA. Patrick Ruch is a member of the EFMI WG NLU.

Henning Müller is a professor at the University of applied sciences Western Switzerland (HES-SO). His research interests include image retrieval, IR system benchmarking, multimodal information retrieval. He is participating in several EU and national projects and is the project coordinator of KHRESMOI.

2. Learning objectives of the workshop

Learning objectives focus on enabling participants to explore the main online health information sources and labels for measuring their quality. Techniques will be explained for health information access such as the extraction of semantics from free text and medical documents and the best ways of ranking documents for retrieval to deliver relevant content. The inclusion of images and videos into the retrieval process using image retrieval will be explained and also the combinations of media. Visual content can be read quicker than text but in terms of retrieval there are limitations. Discussions among workshop participants will be fostered to allow for highlighting the current challenges in making the massive amounts of medical knowledge accessible. In the KHRESMOI project, various user groups such as laypersons, GPs and specialists

are addressed separately as there might not be a single retrieval type that fits all users. Participants should leave the workshop knowing more about the state of the art in multimodal semantic-enhanced health information retrieval and of the challenges ahead.

3. Conclusions

Medical knowledge is becoming increasingly accessible with large repositories being available via the Internet. Besides peer reviewed articles and well controlled content from public bodies this information also contains many personal reports such as blogs and private sites or collaborative tools like Wikis. Information quality is a problem and needs to be addressed. The increased availability of digital content also creates new challenges such as finding the right information in the haystack of data, for example through the extraction of semantics and mapping of texts onto medical terminologies such as MeSH and UMLS. Other challenges are the multilingual nature of the web, meaning that content can be in a language that the user is not comfortable with and can be misunderstood. Multimedia content is another tendency that needs to be addressed as images and videos are not only produced massively but many topics can be understood much more easily through the use of images and videos.

References

- [1] Hartzband P, Groopman J, Untangling the Web — Patients, Doctors, and the Internet, *New England Journal of Medicine* volume 362, pages 1063-1066, 2010.
- [2] Smith R, Strategies for coping with information overload, *British Medical Journal* volume 341, 2010.
- [3] Tang H, Ng JHK, Googling for a diagnosis—use of Google as a diagnostic aid: internet based study, *British Medical Journal*, volume 333, page 1143, 2006.
- [4] Bouwman MG, Teunissen QGA, Wijburg FA, Linthorst GE, ‘Doctor Google’ ending the diagnostic odyssey in lysosomal storage disorders: parents using internet search engines as an efficient diagnostic strategy in rare diseases, *Archives of Disease in Childhood*, volume 95, pages 642-644, 2010.
- [5] White R, Horvitz E, Cyberchondria: Studies of the Escalation of Medical Concerns in Web Search, Microsoft Research Technical Report, MSR-TR-2008-177, 2008.
- [6] Jadad AR, Gagliardi A, Rating Health Information on the Internet: Navigating to Knowledge or to Babel?, *Journal of the American Medical Association*, volume 279(8), pages 611-614, 1998.
- [7] Risk A, Dzenowagis J, Review Of Internet Health Information Quality Initiatives, *Journal of Medical Internet Research*, volume 3(4), 2001.
- [8] Gaudinat A, Ruch P, Joubert M, Uziel P, Strauss A, Thonnet M, Baud R, Spahni S, Weber P, Bonal J, Boyer C, Fieschi M, Geissbuhler A, Health search engine with e-document analysis for reliable search results. *International Journal Medical Informatics*, volume 75(1), pages 73-85, 2006.
- [9] Müller H, Michoux N, Bandon D, Geissbuhler A, A review of content-based image retrieval systems in medicine – clinical benefits and future directions, *International Journal of Medical Informatics*, volume 73, pages 1-23, 2004.

[10] Hanbury A, Boyer C, Gschwandtner M, Müller H, KHRESMOI: Towards a Multi-Lingual Search and Access System for Biomedical Information, Med-e-Tel, Luxembourg, 2011.