

Report on the SIGIR 2014 Workshop on Medical Information Retrieval (MedIR)

Lorraine Goeuriot
Dublin City University
Ireland
lgoeuriot@computing.dcu.ie

Liadh Kelly
Dublin City University
Ireland
lkelly@computing.dcu.ie

Gareth J.F. Jones
Dublin City University
Ireland
gjones@computing.dcu.ie

Henning Müller
HES-SO Valais
Switzerland
henning.mueller@hevs.ch

Justin Zobel
University of Melbourne
Australia
jzobel@unimelb.edu.au

Abstract

The workshop on Medical Information Retrieval took place at SIGIR 2014 in Gold Coast, Australia on July 11. The workshop included eight oral presentations of referred papers and an invited keynote presentation. This allowed time for lively discussions among the participants. These showed the significant interest in the medical information retrieval domain and the many research challenges arising in this space which need to be addressed to give added value to the wide variety of users that can profit from medical information search, such as patients, general health professionals and specialist groups such as radiologists who mainly search for images and image related information.

1 Introduction

Medical information retrieval refers to methodologies and technologies that seek to improve access to medical information archives via a process of information retrieval (IR). Such information is now potentially accessible from many sources including the general web, social media, journal articles, and hospital records. Health-related content is one of the most searched-for topics on the Internet, and as such this is an important domain for research in information retrieval.

Medical information is of interest to a wide variety of users, including patients and their families, researchers, general practitioners and clinicians, and practitioners with specific expertise such as radiologists. There are several dedicated services that seek to make this information more easily accessible, such as *Health on the Net's* medical search systems for the general public and medical practitioners: <http://www.hon.ch/>. Despite the popularity of the medical domain for users of search engines, and current interest in this topic within the IR research community, development of search and access technologies remains particularly challenging and under explored.

One of the central issues in medical information search is the diversity of the users of these services with corresponding differences in types and scopes of their individual needs. Their information needs will be associated with varied categories and purposes, they will typically have widely varying levels of medical knowledge, and, important in some settings, they will have differing language skills.

These challenges can be summarized as follows:

1. Varying information needs: While a patient with a recently diagnosed condition will generally benefit most from simple or introductory information on the disease and its treatment, a patient living with or managing a condition over a longer term will generally be looking for more advanced information, or perhaps support groups and forums. In a similar way, a general practitioner might require basic information quickly while advising a patient, but more detailed information if deciding on a course of treatment, while a specialist clinician might look for an exhaustive list of similar cases or research papers relating to the condition of a patient that they are currently seeking to advise. Understanding various types of users and their information needs is one of the cornerstones of medical information search, while adapting IR to best address these needs to develop effective, potentially personalized systems is one of its greatest challenges.
2. Varying medical knowledge: The different categories of users of medical information search systems will have widely varying levels of medical knowledge, and indeed the medical knowledge of different individuals within a user category can also vary greatly. This affects the way in which individuals pose search queries to systems and also the level of complexity of information which should be returned to them or the type of support in understanding / disambiguating returned material which will be required.
3. Varying language skills: Given that much medical content is written only in the English language, research to date in medical information search has predominantly focused on monolingual English retrieval. However, given the large number of non-English speakers on the Internet and the lack of content in their native language, effective support for them to search English language sources is highly desirable. The Internet in particular has affected the patient-physician relationship, and providing relevant, reliable information to patients in their own language is a key to alleviate such challenging situations and reduce instances of phenomenon such as cyberchondria.

In addition, the format, reliability, and quality of biomedical and medical information varies greatly. A single health record can contain clinical notes, technical pathology data, images, and patient-contributed histories, and may be linked by a physician to research papers. The importance of health and medical topics and their impact on people's everyday lives makes the need for retrieval of accurate and reliable information especially important. Determining the likely reliability of available information is challenging. Finally, as with IR in general, the evaluation of medical search tools is vital and challenging. For example, there are no established or standardized baselines or evaluation metrics, and limited availability of test collections. Further discussion and progression on this topic would be beneficial to the community.

Theme and purpose of the workshop

The objective of the workshop was to provide a forum to enable the progression of research in medical IR seeking to provide enhanced search services for all users with interests in medical information search. The workshop aimed to bring together researchers interested in medical information search with the goal of identifying specific research challenges that need to be addressed to advance the state-of-the-art and to foster interdisciplinary collaborations towards the meeting of these challenges. To enable this, we encouraged participation from researchers in all fields related to

medical information search including mainstream IR, but also natural language processing, multilingual text processing, and medical image analysis.

Topics of interest included but are were not limited to:

- Users and information needs
- Semantics and natural language processing (NLP) for medical IR
- Reliability and trust in medical IR
- Personalised search
- Evaluation of medical IR
- Multilingual issues in medical IR
- Multimedia technologies in medical IR
- The role of social media in medical IR

2 Keynote – Karin Verspoor

The keynote talk was given by Dr Karin Verspoor (University of Melbourne, Australia), on "Practice-based Evidence in Medicine: Where Information Retrieval Meets Data Mining" [1].

A new approach in medical practice is emerging thanks to the increasing availability of large-scale clinical data in electronic form. In practice-based evidence, the clinical record is mined to identify patterns of health characteristics, such as diseases that co-occur, side-effects of treatments, or more subtle combinations of patient attributes that might explain a particular health outcome. This approach contrasts with what has been the standard of care in medicine, evidence-based practice, in which treatment decisions are based on (quantitative) evidence derived from targeted research studies, specifically, randomised controlled trials. Advantages of consulting the clinical record for evidence rather than relying solely on structured research include avoiding the selection bias of the inclusion criteria for a clinical trial and monitoring of longer-term outcomes and effects. The two approaches are, of course, complementary - a hypothesis derived from large-scale data mining could in turn form the starting point for the design of a clinical trial to rigorously investigate that hypothesis. Information retrieval can play an important role in both approaches to collecting medical evidence. However, the use of information retrieval methods in collecting practice-based evidence requires moving away from traditional document-oriented retrieval as the end goal in itself, to viewing that retrieval as an intermediate step towards knowledge discovery and population-scale data mining. Furthermore, it may require the development of more context-specific retrieval strategies, designed to identify specific characteristics of interest and support particular tasks in the medical context.

3 Presented Papers

Of the twenty papers submitted to the workshop, eight were selected for inclusion in the workshop proceedings and for presentation at the workshop:

- Patrick Cheong-Iao Pang, Karin Verspoor, Shanton Chang and Jon Pearce. Designing for Health Exploratory Seeking Behaviour [3]
- Miji Choi, Karin Verspoor and Justin Zobel. Evaluation of Coreference Resolution for Biomedical Text [4]
- Yihan Deng, Matthaeus Stoehr and Kerstin Denecke. Retrieving Attitudes: Sentiment Analysis from Clinical Narratives [5]
- Bevan Koopman and Guido Zuccon. Why Assessing Relevance in Medical IR is Demanding [6]
- Dimitrios Markonis, Roger Schaer and Henning Müller. Multi-modal relevance feedback for medical image retrieval [7]

-
- Liqiang Nie, Mohammad Akbari, Tao Li and Tat-Seng Chua. A Joint Local-Global Approach for Medical Terminology Assignment [8]
 - Rajendra Prasath and Philip O'Reilly. Exploring Clustering Based Knowledge Discovery towards Improved Medical Diagnosis [9]
 - Guido Zuccon and Bevan Koopman. Integrating Understandability in the Evaluation of Consumer Health Search Engines [10]

4 Discussion Session

The discussion sessions started with a brainstorming activity to identify the key challenges in medical IR. The two main areas identified were the lack of available data sets and the need for better evaluation. Two groups were formed to discuss these two topics.

The first group discussed the lack of data sets. One of the reasons for this is the limited amount of publicly available data (i.e. clinical data, query logs, etc.). Aside from the patient related issues of confidentiality and privacy, medical data being very varied and changing, getting representative and up-to-date data sets is very challenging. These variations can be found at different levels. The level of specialization and targeted readers is the first one: consumer information varies greatly from clinical practice information. Then, linguistic variations such as shifting vocabulary are impacting information extraction (IE) and IR results. In order to deal with these changing characteristics, what could the value of abstraction into controlled vocabularies be? Moreover, controlled vocabulary would help in alleviating ambiguity. But how can it be efficiently incorporated into a retrieval approach? Concept-based representation of data and indexing are investigated but their efficiency in IR is still to be proven. Finally, some modalities are very specific to the medical domain, such as temporality, negativity, and patients' characteristics in clinical data such as age, gender, co-morbidities, etc. To understand and automatically process these, training data is necessary (raw data and gold standard annotations), but is difficult and expensive to obtain.

The second group focused on the evaluation of medical information retrieval. They identified as the main issues the lack of evaluation campaigns and benchmarks for medical IR, and the lack of information on the few existing campaigns. Based on the experience of the group members, a few key challenges were focused on, in order to get more benchmarks, and improve their quality. Firstly, it is crucial to design realistic tasks, which involve a deep understanding of the users and their needs. Only once that has been done can the dataset be built, with realistic data. Along with the task and use case scenario, the evaluation scheme and the definition of relevance needs to be very carefully planned, in order to maximize the outcome of the task. In [2], relevance is modelled according to several relevance dimension factors: understandability, topicality, novelty, scope and reliability. For instance, a task focusing on IR for patients or laypeople would define relevance as mainly based on the topicality, the reliability and the understandability. These factors need to be taken into account during the relevance judgement and results evaluation stages [1]. This would allow personalization of the search, characterizing the users and their information needs. Lastly, benchmark creators should be incited to make their datasets available to the public, for specific tasks or any related research work.

5 Conclusions

This was the first SIGIR workshop on 'Medical Information Retrieval', and followed on nicely from the SIGIR 2013 medical workshop on 'Health Search and Discovery: Helping Users and Advancing Medicine'. The volume of interest in the workshop, both through the number of paper submissions and large number of workshop participants, highlight both the activity and interest in the medical information retrieval space within the community. The workshop provided greater insights into the

active areas of research within this space and helped in progression of the many challenges facing the space. Special attention was paid to evaluation within this space and possibilities for progression within the data set creation and benchmarking initiatives discussed.

6 Acknowledgements

We would like to thank SIGIR 2014 for hosting the workshop. Thanks also go to the program committee (Eiji Aramaki, Kyoto University, Japan; Celia Boyer, Health on the Net, Switzerland; Ben Carterette, University of Delaware, USA; Allan Hanbury, Vienna University of Technology, Austria; William Hersh, Oregon Health and Science University, USA; Jung-Jae Kim, Nanyang Technological University, Singapore; Gang Luo, University of Utah, USA; Iadh Ounis, University of Glasgow, UK; Patrick Ruch, HES-SO, Switzerland; Stefan Schulz, Medical University Graz, Austria; Karin Verspoor, NICTA, Australia; Ellen Voorhees, NIST, USA; Ryen White, Microsoft Research, USA; Elad Yom-Tov, Microsoft Research, USA), paper authors and workshop attendees, without whom the workshop would not have been the success it was.

7 References

- [1] Verspoor, K. 2014. Practice-based Evidence in Medicine: Where Information Retrieval Meets Data Mining. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*, keynote.
- [2] Zhang, Y., Zhang, J., Lease, M. and Gwizdka J. 2014. Multidimensional Relevance Modeling via Psychometrics and Crowdsourcing. In *Proceedings of SIGIR 2014*.
- [3] Cheong-lao Pang, P., Verspoor, K., Chang, S. and Pearce, J. 2014. Designing for Health Exploratory Seeking Behaviour. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.
- [4] Choi, M., Verspoor, K. and Zobel, J. 2014. Evaluation of Coreference Resolution for Biomedical Text. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.
- [5] Deng, Y., Stoehr, M. and Denecke, K. 2014. Retrieving Attitudes: Sentiment Analysis from Clinical Narratives. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.
- [6] Koopman, B. and Zuccon, G. 2014. Why Assessing Relevance in Medical IR is Demanding. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.
- [7] Markonis, D., Schaer, R. and Müller, H. 2014. Multi-modal relevance feedback for medical image retrieval. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.
- [8] Nie, L., Akbari, M., Li, T. and Chua, T.S. 2014. A Joint Local-Global Approach for Medical Terminology Assignment. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.
- [9] Prasath, R. and O'Reilly, P. 2014. Exploring Clustering Based Knowledge Discovery towards Improved Medical Diagnosis. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.
- [10] Zuccon, G. and Koopman, B. 2014. Integrating Understandability in the Evaluation of Consumer Health Search Engines. In *Proceedings of the SIGIR workshop on Medical Information Retrieval (MEDIR 2014)*.