WSDM 2019 Tutorial on Health Search (HS2019)
A Full-day from Consumers to Clinicians

Bevan Koopman
Australian e-Health Research Centre, CSIRO
Brisbane, Australia
bevan.koopman@csiro.au

Guido Zuccon
University of Queensland
Brisbane, Australia
g.zuccon@uq.edu.au

ABSTRACT
The HS2019 tutorial will cover topics from an area of information retrieval (IR) with significant societal impact — health search. Whether it is searching patient records, helping medical professionals find best-practice evidence, or helping the public locate reliable and readable health information online, health search is a challenging area for IR research with an actively growing community and many open problems. This tutorial will provide attendees with a full stack of knowledge on health search, from understanding users and their problems to practical, hands-on information on current tools and techniques, evaluation resources, as well as important open questions and future directions. Tutorial material is available at https://ielab.io/health-search-tutorial/.

ACM Reference Format:

1 MOTIVATION AND OVERVIEW
With modern medicine increasingly reliant on information technology, the demand for IR systems that search medical content has grown significantly. The increasing need to retrieve medical advice (by both consumers and clinicians), and the adoption of electronic medical records are two factors driving the demand for health search. IR research has much to offer here by developing new tools and techniques specific to this domain [9].

The range of health information available (primary research sources, secondary research sources, patient records, web pages and popular publications, etc.), plus the range of end users (health consumers, different clinicians — general practitioners, specialists, researchers, etc.), and the range of tasks (searching evidence-based-medicine literature [22], searching patient records and cohort selection [30], searching for medical advice on the Web [34], searching the literature for drug-drug interactions and co-morbidities, searching for clinical trials [13], searching literature to form systematic reviews [10], etc.) all leads to complex requirements that often require novel solutions to these different problems.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WSDM ’19, February 11–15, 2019, Melbourne, VIC, Australia
© 2019 Copyright held by the owner/author(s).
https://doi.org/10.1145/3289600.3291379

1 https://www.nlm.nih.gov/research/umls/
2 http://www.snomed.org/snomed-ct
• provide an understanding of the users, their information needs, tasks and challenges that exist in this domain. This is critical as many of these differ from other domains.

• provide attendees with hands-on experience with health search techniques, tools and problems.

• present an analysis of open questions in the domain.

2 OBJECTIVES

The main aims of HS2019 will be to: (1) Summarise the basics of search in the health domain; (2) Present the different end user requirements for multiple user groups interested in health search, including tasks; (3) Provide an overview of the current use of IR techniques in the health domain; (4) Provide a hands-on introduction to domain-specific tools which can be exploited in health search; (5) Present resources and campaigns for evaluation in health search, including novel evaluation approaches; (6) Present challenges and opportunities for further research in the health domain and discuss how these could be met. This will allow IR researchers to identify promising ways of applying their work to health problems, allowing them to contribute to a domain of rapidly growing importance.

3 TUTORIAL TOPICS AND STRUCTURE

Session 1: Types of health information, end users and tasks. This section covers the characteristics of different types of health information sources important for health search, e.g. patient related (e.g., electronic health records [13, 30]), knowledge related (e.g., scientific papers [10, 22]), consumer related (e.g., patient forums [35]). We also discuss sources of domain knowledge such as medical ontologies, terminologies and classification systems. In addition, an analysis of the end user (from consumers [6, 31, 34] to clinicians [20, 28]) characteristics and tasks in health search is presented.

Session 2 & 3: Techniques, methods and tools. This section covers the state-of-the-art in health search, summarising the most important research methods and results in this area with respect to tasks in health search, highlighting common trends across tasks. This will cover techniques such as query expansion and reformulation, (e.g. [23, 27, 29]), use of domain knowledge and inference mechanisms (e.g. [7, 8, 14, 17, 24, 33]), learning to rank and other learning methods (e.g. [2, 4, 18, 19, 23, 26]), task-based information (e.g. [11, 16]), and specifically handling clinical text (e.g. [5, 12, 15, 25]). This part of the tutorial will also present an overview of tools for extracting clinical and biomedical information, providing a hands-on demonstration of how these tools work and an outlook of how they have been used to enhance information representation and the whole IR process.

Session 4: Evaluation and open challenges. Tasks and challenges in evaluating health search are covered in this section, including evaluation techniques specific to health search (e.g. [11, 32]), and datasets for evaluating health search (e.g. [10, 13, 22, 30, 35]).

REFERENCES


