Report on the SIGIR 2013 Workshop on Health Search and Discovery

Ryen W. White
Microsoft Research
ryenw@microsoft.com

Elad Yom-Tov
Microsoft Research
eladyt@microsoft.com

Eric Horvitz
Microsoft Research
horvitz@microsoft.com

Eugene Agichtein
Emory University
eugene@mathcs.emory.edu

William Hersh
Oregon Health & Science University
hersh@ohsu.edu

For more information about the symposium series, please visit
$http://research.microsoft.com/hsd2013$.

Abstract

The workshop brought together 40 researchers and practitioners from academia and industry to discuss search and discovery in the medical domain. Presentations and discussions spanned several challenging and important topics, including directions improving the accessibility of medical and health information for lay people (with associated enhancements to result ranking algorithms and search interfaces), and methods for discovering biomedical phenomena from the information that people seek online, as evidenced in query streams and other sources such as social and news media. A thread throughout the workshop was the opportunity for new methods and applications to enhance the quality of life of people suffering from medical disorders, carry out surveillance of emerging diseases and other threats, and, more generally, to improve the health and well-being of people via tools to support their health-related information behavior.

1   Introduction and Background

Health information is one of the most widely searched topics on the Web. A recent survey by the Pew Internet and American Life Project found that 59% of U.S. adults have looked online for health information in the past year, with 35% of respondents seeking to diagnose a medical condition online [4]. Health information search is conducted using both search systems such as Internet search engines and via social media (e.g., advice from friends and others). However, studies have found that cyberspace can be a costly place for those with no medical training. For example, the common task of performing self-diagnosis can lead to unwarranted anxiety and healthcare utilization associated with interactions between biases of indexing and retrieval in search engines and biases of user judgment [7].

Such challenges frame current and future research on enhancing information retrieval (IR) methods at both the infrastructure and user-facing levels. Directions include translating both consumer queries and technical terminology, personalizing retrieved content by considering searchers' domain knowledge, and identifying reliable information during ranking and crawling [6]. This domain also presents unprecedented opportunities to develop applications capable of monitoring and improving the quality of life of people affected by a variety of medical conditions via tools to support their health-oriented information behavior. Mining aspects of that behavior (e.g., queries and social media interactions) with consent in the aggregate across many users can assist in medical discoveries [8] and enhance public health monitoring [1, 2, 5] . The monitoring and use of health-related online behavior also highlights important tradeoffs between user/patient privacy and benefit to those people and wider society, which we believe need to be explored and understood in greater detail.

This workshop, held in conjunction with the annual ACM Special Interest Group on Information Retrieval (SIGIR) conference in Dublin, Ireland in August 2013, brought together researchers from a variety of disciplines with interest in health search and discovery. The goal in organizing the workshop in conjunction with the annual SIGIR meeting was to stimulate interest within SIGIR on health search and discovery and to capitalize on the expertise of members of the SIGIR community on such challenges as retrieval, personalization, expertise modeling, search interface design, data mining, and privacy preservation, all of which are critical to enabling advances this health search and discovery.

2   The Workshop

The one-day workshop comprised a range of activities including keynotes, oral presentations, posters, and an introductory ice breaker, and an associated workshop-wide discussion. In the remainder of this section we summarize each of the workshop activities in turn.

2.1   Keynotes

We had the privilege of hearing from two outstanding keynote speakers: Prof. John Brownstein from Harvard University and HealthMap.org, and Prof. Justin Zobel from the University of Melbourne.

John Brownstein, Harvard University and HealthMap.org: In his keynote, entitled "Digital Disease Detection," Prof. Brownstein presented research on how online activity from sources such as social media and Web searching could assist in public health monitoring. He argued that there is value in the fusion and visualization of diverse electronic resources (online epidemic intelligence, social networks, and mobile technology) for rapid identification and dissemination of epidemiological data. Prof. Brownstein described ongoing work on HealthMap.org, which employs data mining techniques to extract information from these sources and summarize them for consumption by the general public and partners such as the Centers for Disease Control and Prevention and the Food and Drug Administration. He discussed a number of case study applications–primarily involving infectious diseases such as influenza A (H1N1) [1]–demonstrating where the methods have performed well, and how mining big data offers outstanding opportunities for global infectious disease surveillance. Moving beyond the passive mining of social media and

other data sources for evidence of disease outbreaks, he discussed "participatory epidemiology" where people can communicate disease information directly at time of interest using mobile applications (e.g., via tweets or text messages sent to the service) or through regular health surveys pushed to consumers. He suggested that Internet-based collaborative systems (where the public is actively involved) can play an important complementary role (to passively mining other data sources) in gathering information quickly and improving coverage and accessibility. This may also result in higher quality data for surveillance purposes that is currently attainable via other means such as mining search queries [2]. Prof. Brownstein described other projects outside the realm of infectious disease monitoring including estimating abortion rates using search engine traffic (studying the distribution of queries for "abortion" at different geographic locations), examining how searchers react to government legislation (e.g., cigarette tax increases), understanding how obesity can be tracked via social media using interest in television as a proxy for obesity level, and monitoring rumors regarding vaccines. The attempts to tap the power of digital tools covered in the keynote demonstrate important steps in improving health systems as well as engaging the public as participants in the public health process.

Justin Zobel, University of Melbourne: In his keynote entitled "Transformative Technologies in Computational Health," Professor Zobel presented insights on the roles of computation, connectedness, and crowdsourcing in enabling transformative new technologies. He focused on two examples of situations where transformative technologies are augmenting or complementing current practices: the diagnosis of infection, and creation of the systematic reviews that are intended to represent authoritative medical knowledge. Based on these examples, Prof. Zobel argued that dramatic, disruptive changes can be an unanticipated outcome of research that had very different initial goals. He presented insights into how research in computational health and biomedicine can be shaped by such advances. Prof. Zobel highlighted several unfulfilled promises of biomedicine and computation, including the use of genetic sequencing as a primary tool in healthcare, and pointed to issues with current data collection and analysis methods, including lack of clear confirmation that the data will help to answer interesting questions, insufficient exploration of parameter settings, and uninformed use of tools designed for other problems. He highlighted some technology-driven changes to clinical practice (e.g., outbreaks mapped as they occur and not in retrospect, medical decisions supported by automated knowledge) some of which are being realized, and asked critical questions such as how to reconcile the insights of the biomedical and computational communities. Prof. Zobel discussed several important challenges associated with making progress in this area, including communication difficulties between biomedical researchers and computer scientists arising in differences in goals and practices, as well as differences in expectations regarding data, technology, and definitions of quality results in the two communities.

2.2 Talks

Five of the submitted papers were selected for oral presentation at the workshop. Each paper was given a 20-minute speaking slot, including time for questions. Papers were grouped together in a single paper session, immediately preceding lunch. The following five papers were selected for presentation:

- *Learning Attribution Labels for Disorder Mentions in Online Health Forums.* Rimma Pivovarov presented research by her colleagues at Columbia University on developing a supervised

method to learn whether a disorder mentioned in a health forum post can be attributed to its author or whether it is in reference to another person. This has the potential to enhance tasks such as adverse drug event detection and more generally association rule discovery over health forums. The authors first classified disorder mentions (including side effects, symptoms, and findings) in two subtasks: general vs. specific, and personal vs. someone else. The authors found that support vector machine (SVM) classifiers augmented with word clusters from unlabeled data and with lexico-syntactic features significantly outperformed a bag-of-words SVM baseline. The system's performance indicates attribution can be uncovered from health forum data and is sufficiently accurate to be leveraged for further analysis.

- *HealthTrust: A PhD Dissertation on the Retrieval of Trustworthy Health Social Media.* Luis Fernandez-Luque presented research on finding trustworthy social media content. He described some of the challenges that need to be addressed to facilitate the retrieval of trustworthy health information from social media sources. For example, misleading and harmful content, such as promoting anorexia as lifestyle or macabre amputations, is often very popular hindering the visibility of high quality online resources. Luis also outlined some next steps, including the integration of trustworthiness features into ranking algorithms and the development of a Web portal for deployment and evaluation of these methods broadly.

- *Assisted Query Formulation for Multimodal Medical Case-based Retrieval.* André Mourão presented a medical retrieval system that supports semantic multimodal queries for medical case-based search (medical.novasearch.org). The system combines a powerful framework based on state-of-the-art image and text processing algorithms with a simple yet powerful multimodal search interface to provide a valuable tool to retrieve medical data. The authors also propose a new data-fusion technique, which they called inverted squared rank, to better deal with the combination of ranked lists from various heterogeneous systems. Studies of system performance with healthcare professionals are pending.

- *Towards Discovery-Oriented Patient Similarity Search.* Haggai Roitman presented research on the patient similarity task, where the goal is to find patients similar along some dimensions with a given patient. This has utility in a range of applications, including helping professionals diagnose a condition by finding others who may have been afflicted with it. The authors propose a solution implemented using a novel query language for semi-structured entity-relationship data using the IBM MedICS system. The system allows the similarity challenge to be reduced to a patient discovery problem that can be efficiently implemented. Extensions to their approach include the use of combinations of data visualization, data summarization, and provenance techniques.

- *Why Is It Difficult to Detect Outbreaks in Twitter?* Nattiya Kanhabua presented research on the challenges of gathering epidemic intelligence (EI) through social media and provided guidance for related research. Event-based EI frameworks leverage social media data to provide public health officials with the tools to identify salient information, namely, disease outbreak events. The challenges of building systems that are capable of doing this include: (1) dynamic classification to enable message filtering, (2) signal generation producing reliable warnings based on observed term frequency changes in the filtered messages, and (3)

providing search and recommendation functionalities to domain experts, for better assessment of the potential outbreak threats associated with the generated signals. The authors outline possible approaches to solve these important challenges as well as discussing areas where further research is required.

The papers that accompany the presentations are in the proceedings, available on the workshop website.

2.3 Posters

Nine of the submitted papers were selected for presentation during an interactive poster session that was held during the afternoon of the workshop. Authors had the opportunity to present their work and engage directly in discussions with other participants. Each of the posters is briefly summarized below.

- *A Tool for Monitoring and Analyzing HealthCare Tweets* describes a platform to help governments or healthcare providers monitor health tweets on social media in different regions.

- *Test Collections for Medical Information Retrieval Evaluation* describes the development of two benchmarks for medical IR evaluation, one targeting medical professionals and another to evaluate patient search in context.

- *Khresmoi Professional: Multilingual Semantic Search for Medical Professionals* describes two systems to support medical search by general practitioners and consultant radiologists using semantic, multilingual text, and image-based (including 2D and 3D radiology images) search.

- *Query Expansion using open Web-based SKOS Vocabularies* describes a query expansion method for improving medical search that exploits term labels and semantic relationships in vocabularies such as the simple knowledge organization system (SKOS).

- *Systems for Improving Electronic Health Record Note Comprehension* describes a biomedical natural language processing system called NoteAid (clinicalnotesaid.org) which translates medical jargon into consumer-oriented language, assisting patients in understanding their health records.

- *Towards a Gold Standard for the Evaluation of Health Recommender Systems* describes methods to evaluate recommender systems for the medical domain in a controlled study, including the construction of a test collection for case-related recommendations.

- *Towards Intelligent and Socially Oriented Query Recommendation for Electronic Health Records Retrieval* describes an electronic health records (EHR) search engine, EHR-SE, comprising a query recommendation service to improve search results and a social search feature enabling the sharing of EHR search knowledge among medical professionals.

- *Clinical Information Retrieval with Split-layer Language Models* describes a split-layer language model that embeds linguistic layers from existing natural language processing systems in retrieving medical documents.

- *Extracting Adverse Drug Reactions from Forum Posts and Linking them to Drugs* describes methods for extracting adverse drug reactions (ADRs) from forum posts and linking extracted ADRs to the drugs that users claim are responsible for them. Short papers on research presented in the proceedings available on the workshop website.

2.4   Ice Breaker and Discussion

Following the first keynote, we conducted a group ice-breaker activity to encourage people to reflect and to engage one another around key opportunities and challenges in health search and discovery that were important to them or topics they believed should be important for the community. People were encouraged to target opportunities where IR technology or expertise could be particularly useful. Once they had collected their thoughts, they were encouraged to write them on sticky notes and attach them to the wall of the workshop room. Participants were encouraged to read each other's notes, cluster them if appropriate, and to introduce themselves to the person standing next to them. Following completion of the ice breaker, the workshop organizers reviewed the sticky notes and identified the following themes that emerged as important to workshop participants:

- Improving and understanding the reliability of online medical content (including rumors, validating medical claims made in social and the news media, understanding and mitigating the effects of bad and harmful content);

- Predictive analytics and knowledge discovery (including enhanced epidemiology and drug safety, and discovering new associations between symptoms, conditions, and treatments);

- Combining multiple sources of evidence, including clinical data such as EHR and other reports, social media, news media, blogs, search log data, and medical literature;

- Handling the vocabulary mismatch between consumers and health professionals (including finding appropriate synonyms for medical/consumer terminology and offering assistance to patients in understanding medical terminology);

- Protecting user privacy, especially when sensitive data sources such as EHR are used;

- Developing tools to support self-diagnosis and comparisons of different treatment options by the public, with appropriate disclaimers and connections to professional medical care as needed;

- Building test collections for medical IR, with important real-world use cases and topics, large volumes of rich and diverse data, and adequate quantities of human judgments;

- Enhanced peer support (including tools to search social media, build scalable and self-sustaining social networks to support collaborative care, and find similar patients);

- Personalization of medical search and analysis services via user/patient profiling, and;

- Developing support to help patients find medical professionals, not just based on distance to the current location, but by using other criteria (e.g., finding the best physician to treat some disease, reputation system for physicians).

These highlight the important directions that our participants believed the IR community needs to take in this area. This also seeded a workshop-wide discussion later in the day, where participants talked about a number of these issues in greater detail, relating them to their personal experience and expanding on points where appropriate. As part of the discussion, the possibility of follow-up events such as other workshops at alternative venues such as the annual symposium of the American Medical Informatics Association (AMIA) was raised, as was the possibility of organizing a journal special issue on this topic. The discussion also touched on the possible value and means for organizing a standing working group on health search and discovery, perhaps hosted within the AMIA community. We will explore the available options and update the research community about our progress.

3   Concluding Remarks

Overall, the workshop was an enjoyable and engaging event and the feedback that we received from participants was positive. We thank our participants for investing time and energy in the workshop and for contributing so much to its success, as well as SIGIR (and the workshop chairs Arjen and Vincent) for giving us the opportunity to hold this event in conjunction with the conference this year. Health search and discovery emphasizes the important role that technology can and will play in areas such as public health and medicine. Since multiple challenges in this area rely on efficient access to reliable content the IR community should be at the forefront of such advances. We look forward to seeing more health-related research at SIGIR and other IR conferences, as well as more IR-related research at venues such as AMIA or other ACM venues such as SIGCHI, SIGKDD, and WSDM. We foresee numerous future developments in this area, including new approaches for detecting and monitoring diseases, supporting self-diagnosis and triage, and identifying adverse drug reactions. We envisage closer collaborations between researchers and other stakeholders in the medicine, medical informatics, and IR communities. We also hope that forthcoming research results will inform and support clinical practice, and realize benefits for health-conscious consumers, patients, and society.

References

[1] J. S. Brownstein, C. C. Freifeld, and L. C. Madoff. Influenza a (h1n1) virus, 2009 — online monitoring. *New England Journal of Medicine*, 360(21):2156–2156, 2009. PMID: 19423868.

[2] D. Butler. When google got flu wrong. *Nature*, 494(7436):155–156, Feb 2013.

[3] M. De Choudhury, S. Counts, and E. Horvitz. Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3267–3276, New York, NY, USA, 2013. ACM.

[4] S. Fox and M. Duggan. Health online 2013. pew internet and american life project. `http://pewinternet.org/Reports/2013/Health-online.aspx`, 2013.

[5] M. J. Paul and M. Dredze. You are what you tweet: Analyzing twitter for public health. In *Proceedings of the Fifth International Conference on Weblogs and Social Media.* AAAI Press, 2011.

[6] T. T. Tang, D. Hawking, N. Craswell, and K. Griffiths. Focused crawling for both topical relevance and quality of medical information. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, CIKM '05, pages 147–154, New York, NY, USA, 2005. ACM.

[7] R. W. White and E. Horvitz. Cyberchondria: Studies of the escalation of medical concerns in web search. *ACM Transactions on Information Systems*, 27(4):23:1–23:37, Nov. 2009.

[8] R. W. White, N. P. Tatonetti, N. H. Shah, R. B. Altman, and E. Horvitz. Web-scale pharmacovigilance: listening to signals from the crowd. *Journal of the American Medical Informatics Association*, 20(3):404–408, 2013.