

Improving Systematic Review Creation With Information Retrieval

Harrisen Scells

Queensland University of Technology
Brisbane, Queensland, Australia
harrisen.scells@hdr.qut.edu.au

ABSTRACT

Systematic reviews, in particular medical systematic reviews, are time consuming and costly to produce but are of value for clinical decision making, policy, and regulations. The largest contributing factors to the time and monetary costs are the searching (including the formulation of queries) and screening processes. These initial processes involve researchers reading the abstracts of thousands and sometimes hundreds of thousands of research articles to determine if the retrieved articles should be included or excluded from the systematic review. This research explores automatic methodologies to reduce the workload relating to the searching and initial screening processes. The objective of this research is to use Information Retrieval techniques to improve the retrieval of literature for medical systematic reviews. The following three research questions elaborate on this:

RQ1: *How to determine the effectiveness of systematic review queries a priori, and generate better queries?*

When performing a systematic review, the information need of medical researchers is to answer the research question of that review. Often, the Boolean queries used to identify potentially relevant literature in the screening phase are highly complex, and may be a poor representation of the information need of the researchers. A good query in the context of systematic review search is one that not only satisfies the information need, but does so in an effective manner (i.e. by restricting the number of non-relevant citations). Prior research has investigated frameworks for formulating Boolean queries to meet these needs [2, 5], however there has been no work in developing these methods for predicting the effectiveness of Boolean queries prior to the screening phase.

This question relates to determining if a Boolean query in a systematic review is well suited to the information need (the research question posed by the review), and will investigate *Query Performance Predictors* (QPPs) as one approach for predicting the effectiveness of a Boolean query. Additionally, in order to generate queries, this research will investigate query transformation methods and apply them to Boolean queries. The focus of this question will be the use of QPPs to predict the effectiveness of Boolean queries within the context of systematic reviews, and to investigate query generation methods (i.e. query expansion, reduction, transformation) to produce better, more effective Boolean queries.

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).
SIGIR '18, July 8–12, 2018, Ann Arbor, MI, USA
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-5657-2/18/07.
<https://doi.org/10.1145/3209978.3210226>

It is envisioned that this solution can be integrated into the query formulation stage of systematic review creation, and will assist researchers in formulating more effective queries. Answering this research question is the primary focus of this work.

RQ2: *How to improve systematic review search with new retrieval models?*

Reducing the workload for researchers in the screening phase reduces the cost in time and money involved with producing a systematic review. This research question relates to investigating methods for reducing the workload of the screening phase via retrieval, and methods for measuring the workload reduction. Specifically this research question will investigate query and document structure for improving retrieval, e.g. Scells et al. [4].

RQ3: *How can ranking assist reviewers prioritise citations in the screening phase?*

A ranking function orders a retrieved set of documents based on some notion of *relevance*. The Boolean retrieval model defines a simple form of ranking where all documents retrieved are equivalent in terms of relevance [1]. This means that Boolean queries retrieve documents *unordered* (i.e. a set).¹ Within the context of systematic reviews, ranking can be considered equivalent to screening prioritisation; i.e. the citations resulting from the search strategy is ordered by some notion of relevance. A review of the literature has identified that this is an area of research that has not been explored in enough detail for this domain. This research question will investigate learning to rank [3] as a potential solution for screening prioritisation. The hypothesis is that with domain-specific and discriminative features for training, learning to rank can be used to effectively rank citations retrieved with Boolean queries.

REFERENCES

- [1] W Bruce Croft, Donald Metzler, and Trevor Strohman. 2010. *Search engines: Information retrieval in practice*. Vol. 283. Addison-Wesley Reading.
- [2] Gary Marchionini. 1996. *Information Seeking in Electronic Environment*. Vol. 9. Cambridge University Press.
- [3] Harrisen Scells, Guido Zuccon, Anthony Deacon, and Bevan Koopman. 2017. QUT ielab at CLEF eHealth 2017 technology assisted reviews track: Initial experiments with learning to rank. In *CEUR Workshop Proceedings: Working Notes of CLEF 2017: Conference and Labs of the Evaluation Forum*, Vol. 1866. CEUR Workshop Proceedings, Paper–98.
- [4] Harrisen Scells, Guido Zuccon, Bevan Koopman, Anthony Deacon, Leif Azzopardi, and Shlomo Geva. 2017. Integrating the framing of clinical questions via PICO into the retrieval of medical literature for systematic reviews. In *CIKM 2017 International Conference on Information and Knowledge Management*. ACM, Singapore.
- [5] Eero Sormunen. 2000. A novel method for the evaluation of Boolean query effectiveness across a wide operational range. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 25–32.

¹Although some type of weak ordering could be enforced, e.g. coordination level matching.