

Systematic Review Automation Tools for End-to-End Query Formulation

Hang Li
University of Queensland
Brisbane, Australia
hang.li4@uq.net.au

Harrisen Scells
University of Queensland
Brisbane, Australia
h.scells@uq.net.au

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

ABSTRACT

Systematic reviews are used widely in the biomedical and health-care domains. Systematic reviews aim to provide a complete and exhaustive overview of the medical literature for a specific research question. Core to the construction of a systematic review is the search strategy. The main component of a search strategy is a complex Boolean query, typically developed by information specialists (e.g., librarians). The aim of the search strategy is to retrieve relevant studies that will contribute to the outcomes of the systematic review. One barrier information specialists face when developing a search strategy is the enormous amount of medical literature that exists in databases. This vast amount of literature means that search strategies often suffer from biases (e.g., lack of expertise, overconfidence, limited knowledge of the domain) and are incomplete, or retrieve far too many studies (possibly as a result of the biases, but also due to the tools used to develop search strategies). Retrieving too many studies impacts the time and financial costs of the review, and retrieving too few studies may impact the outcomes of the review. Therefore, it is vital to support expert searchers develop effective search strategies. In this paper, we present a novel end-to-end set of advanced tools for information specialists. These tools are tightly integrated into an existing Open Source search strategy refining package (searchrefiner). These tools aim to address the problems associated with search strategy development by providing a complete framework from query development, to refinement, to documentation. The implementation of these tools also offers a glimpse at the ease at which related tools may be implemented within the searchrefiner ecosystem. More information about the tools including installation, documentation, and screenshots is made available on the searchrefiner website: <https://ielab.io/searchrefiner>.

ACM Reference Format:

Hang Li, Harrisen Scells, and Guido Zuccon. 2020. Systematic Review Automation Tools for End-to-End Query Formulation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), July 25–30, 2020, Virtual Event, China*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3397271.3401402>

Permission to make digital or hard copies of all or part 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 components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SIGIR '20, July 25–30, 2020, Virtual Event, China

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-8016-4/20/07...\$15.00
<https://doi.org/10.1145/3397271.3401402>

1 DESCRIPTION OF SYSTEM

Systematic reviews are used as a highly trustworthy and comprehensive evidence source for informing clinical diagnosis and treatment, as well as institutional policy making. In order to be comprehensive, all relevant studies related to the research question of the review must be retrieved. However, there are many costly challenges in producing systematic reviews, e.g., developing a search strategy (Boolean query), and screening retrieved studies for relevance. These two processes in particular hinder the construction of systematic reviews significantly: for a typical systematic review, the screening process can take upwards of 47 days [3] and \$AUD 437,000 [8]. One solution to combat these costs is to use more effective search strategies [16]. The three main challenges of search strategy development are formulation [4, 6], refinement [16], and documentation [22]. These challenges are complex and time-consuming, especially for constructing and refining the search strategies. Currently, most search strategies are manually developed, where the presence of an expert is essential [13]. However, search strategy development is filled with biases [18] and subjectivity [10], and often search strategies retrieve far too many studies [11].

PubMed¹ is widely used by millions of users to search for biomedical information [11]. The tools afforded by PubMed, however, currently do not support information specialists to the extent that they require. To this end, we implemented several tools inside an existing Open Source framework, searchrefiner [17], as an end-to-end system for the **formulation**, **refinement**, and **documentation** of Boolean queries for systematic review literature search. The following is a summary of the three categories of search strategy development for which we have implemented tools for:

Query Formulation: Formulating a preliminary query is the first step in the development of a search strategy. Query formulation is typically ad-hoc and requires great expertise.

Query Refinement: After formulating a preliminary query, information specialists further refine the query. Refinement aims to reduce the total number of studies retrieved by the initial query, while maintaining recall. The refinement process, much like formulation, is ad-hoc, and arguably requires even greater expertise.

Query Documentation: In the final stage of the search strategy development process, the query must be sufficiently documented. It is common for search strategies to be under-reported, leading to irreproducible systematic reviews [22].

The following sections describe each of the problem categories, as well as the tool(s) we have implemented to address these problems. An overview of the names of the tools and the specific problem each of the tool addresses is presented in Table 1

¹The largest medical literature database, maintained by National Library of Medicine.

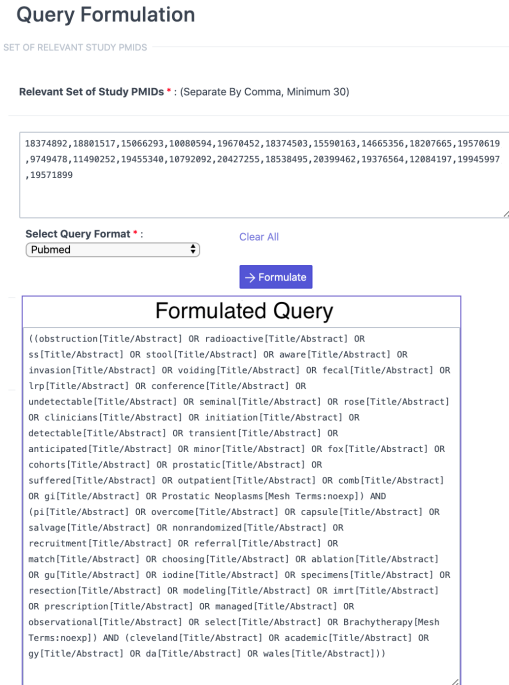


Figure 1: Formulating a PubMed query using the AutoFormulate tool. Both a query with MeSH headings, and a query without MeSH headings is generated.

1.1 Query Formulation

There has been much research into search strategy development methodology [1, 2, 5]. Recently, Russell-Rose et al. [14] proposed a visual approach *2DSearch* that consists of a graph editor in which Boolean queries are expressed as objects on a two-dimensional canvas. Their method eases the query formulation process by allowing users to visually construct their query, possibly reducing common errors. However, this tool still requires the same cognitive effort to develop a search strategy as existing methods and tools.

Among all previous works, another common method of developing a search strategy is the *conceptual* method [4, 9], which formulates the query by dividing the proposed research question into several high-level concepts, and then identifying keywords to match these concepts. This process is performed manually, in an ad-hoc manner, often and preferably by trained information specialists, commonly without much rigour in terms of evaluation [18]. To avoid the problems associated with the conceptual approach, an *objective* method of formulating search strategies for systematic reviews has been proposed [6, 12]. Using this method, a small collection of ‘gold standard’ studies is first identified to seed a semi-automatic keyword selection process, and then used for validation. This method, however, is still ad-hoc as it requires the information specialists to manually test and choose the keywords to add to the query. This is a time-consuming process as well, as it still requires human effort and input. Moreover, it is possible that this method will introduce extra errors and bias into the final queries.

AutoFormulate (Figure 1) is an interface for supporting the automatic formulation of search strategies using the objective method. To reduce the human workloads, unnecessary errors, and bias in formulating search strategies, Scells et al.[18] have proposed a

Tool	Problem Category
AutoFormulate	Query Formulation
KeywordSuggest	Query Refinement
QueryVis	Query Refinement
QueryLens	Query Refinement
AutoDoc	Query Documentation

Table 1: The specific problem each of the tools addresses.

computational adaptation to the objective method which aims to approximate human intuition in constructing a search query. It does not require manual human involvement, nor trial-and-error procedures, and is capable of generating a query automatically. AutoFormulate provides a direct interface to this computational adaptation of the objective method. To use this tool, information specialists first input a set of seed studies. AutoFormulate divides these studies into three different categories: development, validation, and unseen. The development set is used to identify suitable terms for the query. The validation set is used to tune which combination of terms results in the most effective best query (queries can be tuned for different evaluation measures, e.g., precision, recall, F_β , etc.). The unseen set is used to evaluate the resulting query.

1.2 Query Refinement

Query refinement is integral to search strategy development. It is also the lengthiest and most challenging part of the search strategy development process. To this end, we have focused our efforts on three tools to assist information specialists refine their queries. These tools assist information specialists by suggesting keywords to add to their query, visualising their query, and allowing them to explore the space of query reformulations. The three tools are described in further details below.

KeywordSuggest (Figure 2) is a tool that provides keyword suggestions given an input keyword. KeywordSuggest utilises clinical concept embeddings [21], and PubMed term frequency statistics. Keywords are ranked by similarity and merged into a single list by normalising the similarity scores from each source.

QueryVis [20] (Figure 3) is a tool for visualising search strategies. QueryVis presents a query as a tree with information about the number of studies retrieved as well as number of seed studies retrieved by each leaf node. These layers of information on top of the hierarchical representation of a query assist information specialists by allowing them to gain a deep understanding of what a query retrieves and exactly how it retrieves it. Furthermore, the above KeywordSuggest tool is integrated into this interface to allow information specialists to explore suggestions.

QueryLens [19] (Figure 4) is a tool that automatically generates variations for a query (by, e.g., adding or removing keywords, and fields, rewriting Boolean operators, or exploding MeSH headings). The middle panel provides statistics and evaluation results. The right panel plots each variation in precision-recall space. A point in this plot can be clicked to see information about the referring query in the middle panel. The variations produced by QueryLens are ranked as suggestions using a learning to rank model optimised for evaluation measures (e.g., precision, recall, F_β , etc.) [15, 16].

Word Suggestion

Term: (Required)
neoplasm

Return size: (Optional, Default 10, Max 30)
10

Number of docs for search in ES index: (Optional, Default 20, Max 100)
50

Merge Results (Optional, Default is true)

Data Sources: (Optional, Default Using cui,es)

CUI2Vec

Elastic Search Index

Clear All

→ Search

10 Alternatives Found
In 31.262 seconds

1. neoplasm recurrence Score: 1 Source: CUI	2. atherosclerosis Score: 1 Source: ES	3. ischemic Score: 0.985 Source: ES
4. secondary malignancies Score: 0.8 Source: CUI	5. no metastases (tumor staging) Score: 0.8 Source: CUI	6. health Score: 0.8 Source: ES
7. heart Score: 0.769 Source: ES	8. ihd Score: 0.523 Source: ES	9. congenital Score: 0.231 Source: ES
10. reasoning Score: 0.108 Source: ES		

Figure 2: Suggesting keywords using the KeywordSuggest tool. Given an input term, suggestions from multiple sources are presented for consideration.

Text Editor Visual Editor

22626 citations retrieved
8 citations relevant
8 citations relevant retrieved

Query Language List
PubMed

Alternative Words Found

1.) edge to vector cancer abct 07 Score: 11 Source: CUI	6.) International union against cancer Score: 11 Source: CUI
2.) ar@QdR Score: 11 Source: CUI	7.)))) Score: 11 Source: CUI
3.)))) Score: 11 Source: CUI	8.)))) Score: 11 Source: CUI
4.)))) Score: 11 Source: CUI	9.)))) Score: 11 Source: CUI
5.)))) Score: 11 Source: CUI	10.)))) Score: 11 Source: CUI

brain[All Fields] OR cancer[All Fields]

Figure 3: Manually refining a query using QueryVis and KeywordSuggestor tools. The top-left panel is where queries can be input to the tool. The bottom-left panel presents alternative terms found by the KeywordSuggestor tool. When keywords in the right panel are clicked, keyword suggestions fill the bottom-left panel.

1.3 Query Documentation

Search strategies are included in systematic review for reproducibility purposes. According to the Cochrane Handbook, the report of the search strategy used should consist of seven elements [7]: (1)

Continue Editing QueryVis > Results >

- Number of variations: 14
- Type of transformation of this variation: Original
- PS: 0.000869705020697
- Precision: 0.00086756011758908
- Recall: 1
- # Retrieved: 7021

0.0010
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 4: Automatically refining the query with QueryLens[19]. The left panel is where queries can be input to the tool. The middle panel displays query statistics. The right panel plots all variations in terms of recall and precision. The sliding bar controls which variation is inspected. Clicking points in the plot also selects variations. The bottom plot shows how the query changes over time.

Auto Doc

GENERAL INFORMATION

Years Covered By Search * : 1990 - 2000

Query Author (Optional) * : Harrison Scells

Search Strategy Summary * : Use the query to search for relevant studies.

Language Restrictions * : English only.

DATABASE INFORMATION

Date Search Was Run * : November 2019

Databases Searched * : PubMed

Name of Host (Optional) * : Host That Hosts the Database Above

Complete Search Strategy * : No Types Found.

(((schistosomiasis OR cca OR used OR found OR hematuria OR either) AND (strips OR urinalysis OR dipstick OR dipsticks OR ng OR Parasite Egg Count) AND (village OR villages OR kg))

Use the + and - button below to add or remove information if multiple databases are used.

All tooltips are referenced from Analysis of the reporting of search strategies in Cochrane systematic reviews in J Med Libr Assoc 97(1) January 2009 Authors are: Adriana Yoshii, MLS, AHP; Daphne A. Plaut, MLS, AHP; Kathleen A. McGraw, MA, HLS; Margaret J. Anderson, MS; Kay E. Wellik, MLS, AHP DOI: 10.3163/1536-5050.97.1.004

Clear All

→ Generate Report

Search Strategy Report

We are using 1 database in our search process.
The authors for the search strategies:
Harrison Scells
The database is:
PubMed database, the search was conducted in [2019-11] with years covered from 1990 to 2000.
The complete search strategy we used for this database is as follows:
(((schistosomiasis OR cca OR used OR found OR hematuria OR either) AND (strips OR urinalysis OR dipstick OR dipsticks OR ng OR Parasite Egg Count) AND (village OR villages OR kg))
To summarise our search strategy:
Use the query to search for relevant studies.
There are some language restrictions applied to our search strategy:
English only.

Figure 5: Documenting the query using AutoDoc. This tool validates both queries and each of the seven elements that should be reported. Once validated, a report can be generated and can be sent to the authors of the systematic review.

databases searched (e.g., PubMed), (2) name of host (e.g., Ovid), (3) date search was run, (4) years covered by search, (5) complete search strategy, (6) summary of the search strategy, and (7) language restrictions. When reporting search strategies, authors rarely include all seven elements, e.g., Yoshii et al. [22] analysed 65 systematic reviews and found that none included all seven elements.

AutoDoc (Figure 5) is the first tool of its kind for supporting the documentation of search strategies. AutoDoc validates both search strategies and the seven elements that should be reported. Firstly, the tool checks for spelling mistakes and logical errors in the query. Next, it validates the forms that information specialists are required to fill that cover each of the seven elements that should be reported. After the query and forms have been validated, a report is generated. The generated search strategy report can finally be copied into the relevant section of a systematic review.

2 CASE STUDY

Next, a case study to demonstrate how an information specialist may use this end-to-end set of tools to complete the task of search strategy development is presented. Search strategies are used to identify as many relevant studies as possible, and this usually requires a large amount of time and effort. Without an effective search strategy, the search results may be biased or incomplete. Hence, we present the workflow that one may go through to formulate, refine, and document a complete, and effective search strategy.

First, the information specialist acquires a set of ‘seed’ studies. These studies are known to be relevant, typically provided by the authors of the systematic review. Seed studies can be entered into searchrefiner through a settings interface (not shown). Many of the tools presented here access seed studies for a variety of functions. The information specialist then proceeds to the **AutoFormulate** tool. A query is automatically derived from the seed studies, and once satisfied, the information specialist proceeds to the next step.

Once a query has been formulated, the information specialist moves into the tools for query refinement. The information specialist may first visualise the query using **QueryVis** to gain an understanding of what studies it retrieves, which clause retrieves the most relevant studies, and how it is retrieving them. The information specialist manually removes clauses they think are not contributing to the search based on the query visualisation provided by QueryVis, but also wishes to add terms to broaden the search. Rather using their intuition, they could use the **KeywordSuggest** component that is embedded into the QueryVis, providing them with keyword suggestions that can be replaced with the original keywords or added into the search query. Once satisfied with their manual refinement, the information specialist then explores any other modifications they might be able to make that they could have missed using **QueryLens**. The information specialist is presented with a series of refinements they were not aware of from manual refinement that improve the effectiveness of the query. The information specialist accepts the refinements and once more returns to QueryVis to see the impact of this. Once satisfied, they are ready to document their search strategy.

The information specialist then sends their query into the **AutoDoc** tool, and proceeds to fill out the seven elements that should be included in a search strategy report. The information specialist gets to the end of the forms and notices an alert that says there is a spelling mistake in their query. Once the error has been corrected, the information specialist generates the report and sends it to the authors of the review, who will execute the search strategy and begin screening studies for inclusion in the systematic review.

Acknowledgements. Harrison Scells is the recipient of a CSIRO PhD Top Up Scholarship. Dr Guido Zuccon is the recipient of an Australian Research Council DECRA Research Fellowship (DE180101579).

REFERENCES

- [1] Robert G Badgett, Daniel P Dylla, Susan D Megison, and E Glynn Harmon. 2015. An experimental search strategy retrieves more precise results than PubMed and Google for questions about medical interventions. *PeerJ* 3 (2015), e913.
- [2] Wichor M Bramer, Gardien B de Jonge, Melissa L Rethlefsen, Frans Mast, and Jos Kleijnen. 2018. A systematic approach to searching: an efficient and complete method to develop literature searches. *Journal of the Medical Library Association: JMLA* 106, 4 (2018), 531.
- [3] Krystal Bullers, Allison M Howard, Ardis Hanson, William D Kearns, John J Orriola, Randall L Polo, and Kristen A Sakmar. 2018. It takes longer than you think: librarian time spent on systematic review tasks. *Journal of the Medical Library Association: JMLA* 106, 2 (2018), 198.
- [4] Justin Clark. 2013. Systematic Reviewing. In *Methods of Clinical Epidemiology*, Gail M. Williams Suhail A. R. Doi (Ed.). Springer.
- [5] Mette Brandt Eriksen and Tove Faber Frandsen. 2018. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *Journal of the Medical Library Association: JMLA* 106, 4 (2018), 420.
- [6] Elke Hausner, Siv Waffenschmidt, Thomas Kaiser, and Michael Simon. 2012. Routine development of objectively derived search strategies. *Systematic reviews* 1, 1 (2012), 19.
- [7] Julian PT Higgins and Sally Green. 2011. *Cochrane handbook for systematic reviews of interventions*. Vol. 4. John Wiley & Sons.
- [8] Joseph Lau. 2019. Editorial: Systematic review automation thematic series. *Systematic Reviews* 8, 1 (2019), 70.
- [9] Carol Lefebvre, Eric Manheimer, and Julie Glanville. 2008. Searching for studies. *Cochrane handbook for systematic reviews of interventions: Cochrane book series* (2008), 95–150.
- [10] Ying-Hsang Liu. 2009. *The impact of MeSH (Medical Subject Headings) terms on information seeking effectiveness*. Ph.D. Dissertation. Rutgers University-Graduate School-New Brunswick.
- [11] Zhiyong Lu, W John Wilbur, Johanna R McEntyre, Alexey Iskhakov, and Lee Szilagyi. 2009. Finding query suggestions for PubMed. In *AMIA Annual Symposium Proceedings*, Vol. 2009. American Medical Informatics Association, 396.
- [12] Alison O’Mara-Eves, James Thomas, John McNaught, Makoto Miwa, and Sophia Ananiadou. 2015. Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic reviews* 4, 1 (2015), 5.
- [13] Melissa L Rethlefsen, Ann M Farrell, Leah C Osterhaus Trzasko, and Tara J Brigham. 2015. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *Journal of clinical epidemiology* 68, 6 (2015), 617–626.
- [14] Tony Russell-Rose and Jon Chamberlain. 2019. An open-access platform for transparent and reproducible structured searching. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1293–1296.
- [15] Harrison Scells, Leif Azzopardi, Guido Zuccon, and Bevan Koopman. 2018. Query variation performance prediction for systematic reviews. In *SIGIR’18*.
- [16] Harrison Scells and Guido Zuccon. 2018. Generating better queries for systematic reviews. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 475–484.
- [17] Harrison Scells and Guido Zuccon. 2018. searchrefiner: A query visualisation and understanding tool for systematic reviews. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 1939–1942.
- [18] Harrison Scells, Guido Zuccon, and Bevan Koopman. 2020. A computational approach for objectively derived systematic review search strategies. In *Proceedings of the 42nd European Conference on Information Retrieval*.
- [19] Harrison Scells, Guido Zuccon, Bevan Koopman, and Justin Clark. 2019. Automatic search strategy reformulation interface for systematic reviews. In *Proceedings of the 2019 Cochrane Colloquium*.
- [20] Harrison Scells, Guido Zuccon, Bevan Koopman, and Justin Clark. 2019. Visualising systematic review search strategies to assist information specialists. In *Proceedings of the 2019 Cochrane Colloquium*.
- [21] Anton H. van der Vegt, Guido Zuccon, and Bevan Koopman. 2019. Learning Inter-Sentence, Disorder-Centric, Biomedical Relationships from Medical Literature. In *Proceedings of AMIA 2019*.
- [22] Adriana Yoshii, Daphne A Plaut, Kathleen A McGraw, Margaret J Anderson, and Kay E Wellik. 2009. Analysis of the reporting of search strategies in Cochrane systematic reviews. *Journal of the Medical Library Association: JMLA* 97, 1 (2009), 21.