Towards Explainability in Professional Search

Tony Russell-Rose
Department of Computing
Goldsmiths, University of London
London, UK
T.Russell-Rose@gold.ac.uk

Andrew MacFarlane
Centre for HCI Design
City University of London
London, UK
A.Macfarlane-1@city.ac.uk

ABSTRACT

Professions such as healthcare, law, recruitment and patent search all share an interest in the resolution of complex information needs. This typically involves the formulation of structured search strategies that are expressed as Boolean strings. However, creating effective Boolean queries remains an ongoing challenge, often compromised by a lack of transparency and reproducibility. In this paper we explore some of the shortcomings of current approaches, examine alternative solutions and make recommendations towards improved explainability in professional search.

CCS CONCEPTS

• Information Systems -> Information Retrieval -> Users and interactive retrieval

KEYWORDS

Explainable Search, Professional Search

ACM Reference format:

Tony Russell-Rose and Andrew Macfarlane. 2020. Towards Explainability in Professional Search. In Proceedings of EARS 2020, July, 2020, Xi'an, China, 5 pages. https://doi.org/10.1145/1234567890

1 Introduction

The field of information retrieval (IR) has given much attention to generic users who issue short queries in response to simple information needs [1]. Ranking models such as BM25 [2] have been successfully applied to address the needs of such users. However, there has been relatively little attention given to professional users who rely on logic models to resolve complex information needs [3]. In contrast to non-professional searchers, these users are characterized by the development of complex search strategies that attempt to identify conceptual structure within composite information needs.

978-1-4503-0000-0/18/06...$15.00
https://doi.org/10.1145/1234567890

Professional search (PS) is often focused on high-recall use cases. In professions such as healthcare, legal research, recruitment and patent search, failure to retrieve all the relevant literature could result in unwanted outcomes such as a systematic review becoming invalid, potential loss of a court case, overlooking a candidate with the right expertise, or a potential patent infringement.

The methods applied in PS should be transparent and explainable so that they can be verified and audited by third parties. For example, search strategies may need to be presented in a court of law to justify the inclusion or exclusion of a particular piece of prior art. Likewise, they may be presented as evidence of due diligence in the publication of a systematic literature review. This is particularly important in highly regulated professions such as healthcare [4], since clinical guidance may require regular updates. This requires a further iteration of the search process which is often performed by other individuals long after the original search was published. The aim of this paper is to outline the shortcomings of current approaches, explore alternatives and make recommendations towards improved explainability in professional search.

The paper is structured as follows. In Section 2 we provide examples of complex information needs that highlight the challenges faced by professional searchers. We then examine the explainability problem in section 3, and explore alternative approaches and solutions in section 4. We then draw these threads together to identify recommendations and design principles for explainable search in section 5. We conclude in section 6 with a summary and future work.

2 Professional Search Strategies

One of the key characteristics of professional search is a requirement to analyze and identify the conceptual structure within a composite information need (which is often determined by a client brief or protocol). This is achieved by subdividing the information need into discrete facets which are represented by sets of related terms or phrases. In a Boolean search strategy, the OR operator is usually applied within each facet, and a composite result is formed by applying the AND operator across facets. The

Boolean AND NOT operator may be used to eliminate unwanted concepts or terms. There are many techniques for constructing Boolean queries including Building Blocks and Successive Fractions [5]. The outcome is typically a multi-line search strategy as shown in Figure 1. This example is from healthcare, but strategies from other professions can be equally complex.

```
1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. clinical trials as topic.sh.
6. randomly.ab.
7. trial.ti.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. (animals not (humans and animals)).sh.
10. 8 not 9
11. exp Child/
12. ADOLESCENT/
13. exp infant/
14. child hospitalized/
15. adolescent hospitalized/
16. (child$ or infant$ or toddler$ or adolescen$ or teenage$).tw.
17. or/11-16
18. Child Nutrition Sciences/
19. exp Dietary Proteins/
20. Dietary Supplements/
21. Dietetics/
22. or/18-21
23. exp Infant, Newborn/
24. exp Overweight/
25. exp Eating Disorders/
26. Athletes/
27. exp Sports/
28. exp Pregnancy/
29. exp Viruses/
30. (newborn$ or obes$ or "eating disorder$" or pregnan$ or childbirth or virus$ or influenza).tw.
31. or/23-30
32. 10 and 17 and 22
33. 32 not 31
```

Figure 1: An example professional search strategy

A brief analysis of Figure 1 allows a number of observations to be made [6]. First, it is difficult to identify the overall structure of the strategy; in particular how concepts are related and how terms are combined to create those concepts. Second, such strategies scale poorly: an information need may often require the use of many concepts comprising hundreds of terms spread over a number of physical pages. Third, Boolean search strategies are hard to debug and maintain: even to the trained eye it is difficult to identify and correct syntactic or semantic errors. These issues all contribute to a lack of transparency and repeatability, and hence compromise explainability in professional search. In the next section we examine evidence from the literature that explores these issues in further detail.

3 The Explainability Problem

Explainability in professional search can be thought of as entailing two distinct (but complementary) criteria:

- The degree to which an articulated information need produces the results that the user expected and intended;
- The degree to which an unarticulated information need can be parsimoniously articulated by the user and represented with maximal fidelity to their intent with minimal opportunity for error.

Ostensibly, the conventional formalism of Boolean logic offers a satisfactory solution to the first criterion since its semantics are deterministic and in principle the output of any given expression should be predictable and repeatable. This contrasts sharply with the default ranking mechanisms of most proprietary web and enterprise search engines, which are rarely open to public scrutiny and often subject to continual change. However, a closer examination reveals a number of issues that compromise this criterion and hence undermine explainability in PS.

First, professional searchers routinely formulate expressions that in addition to Boolean logic often include proximity, truncation field codes and other operators. Support for these operators can vary radically across different databases, not only in their availability but also in the way they are applied and in their relative precedence. This inconsistency adds complexity and further undermines explainability. As a consequence, repeatability in professional search is typically confined to a more narrow interpretation predicated on specific platforms and syntaxes.

Second, developing professional search skills involves a steep learning curve [7]. Searchers will not always know the optimal approach to adopt to address a given information need, and may require multiple iterations to develop an effective solution [8]. A lack of transparency and repeatability can make this undertaking unduly time-consuming and resource-intensive.

Third, the current practice of using document-centric media such as PDF or MS Word for search strategy development introduces further errors and inefficiencies [19]. For example: auto-correction can undermine truncation and corrupt truncated formats; spell checking can obfuscate differences between British and US English and can create unwanted duplicates; and copying and pasting text fragments between word-processing tools can lead to loss of non-print characters. In sum, the practice of manipulating search strategies as text strings compromises their ability to function as transparent, reproducible, explainable artifacts.

Let us turn now to the second criterion. Arguably, it is this requirement that is most poorly served by current formalisms and in particular the ubiquitous Boolean string. The issues of obfuscation, poor scalability and propensity for error raise significant questions regarding its fitness for purpose. So why has it remained ubiquitous among the professional search community?

Part of the reason may be attributed to a lack of incentive to innovate among the database vendors whose monopoly over data sources affords their platforms a privileged degree of exclusivity and profitability. Ironically, it transpires their users may unwittingly collude in this status quo, by finding ingenious ways to work around the above shortcomings, in many cases actively cultivating 'black art' search skills to mitigate the risk of potential disintermediation by automated solutions. In addition, many professional searchers will have developed a substantial body of knowledge and competence gained through years of experience, and they understandably may be reluctant to forfeit this knowledge (however imperfect) without the prospect of demonstrably greater reward. And on a more prosaic note, the introduction of novel methods would inevitably disrupt established workflows such as the production of systematic literature reviews [9,10].

A further, more pragmatic issue lies with the tooling itself, namely the command-line query builder and its underlying paradigm of capturing information needs as a set of expressions connected by line number. If this approach is intended to provide a principled mechanism for representing structure in composite information needs, why does it rely on something as arbitrary as a line number? This is the conceptual equivalent of the GOTO statement in first generation BASIC, which is an approach that was discredited many decades ago [11].

Finally, we should consider the degree to which existing formalisms are found to be error prone. This issue is highlighted by McGowan and Sampson [12] and Salvador-Olivan et al [13], who found that as many as 90% of the search strategies in one sample they reviewed contained at least one error, and that 80% of those errors that had a direct effect on recall. This finding underlines the shortcomings outlined above, and further motivates the pursuit of alternative solutions. It is to these that we turn to in the next section.

4 Alternative Approaches to Explainability

Professional search has a long history, much of which pre-dates the search experiences that we now consider mainstream [5]. Arguably, explainability has been a key criterion from the outset, in the sense that professional searchers operate in environments that require a level of governance and auditability. However, the translation of that requirement into effective and scalable solutions would appear to have been less than wholly successful. In this section we review some alternatives to the conventional approach of command-line query builders and Boolean strings.

Anick et al. [14] is an early example of an alternative approach. They developed a system that could parse natural language queries and represent them as movable tiles on a visual canvas. The user could rearrange the tiles to reformulate the expression and to activate or deactivate alternative elements to modify the query. These innovations helped mitigate many of the syntactic errors associated with query string manipulation.

In subsequent work, Fishkin and Stone [15] investigated applying direct manipulation techniques to database query formulation using a system of "lenses" to refine and filter the data. Users could combine lenses by stacking them and applying a suitable operator or combine them to create compound lenses, supporting the encapsulation of complex queries. Jones [16] proposed an influential approach in which concepts are expressed using a Venn diagram notation combined with integrated query result previews. Users could formulate queries by overlapping objects within the workspace to create intersections and disjunctions, and they could select subsets to achieve a further refined set of results.

Yi et al. [17] developed a system based around a "dust and magnet" (p 239) metaphor, in which users could represent dimensions of interest within the data as magnets on a visual canvas. The effect of the "magnetic forces" on individual "data particles" reflected the relationships between points in the data, using interaction and animation to communicate cause and effect. Nitsche and Nürnberger [18] developed a system based around a radial interface in which users could integrate and manipulate queries and results. The concept used a pseudo-desktop metaphor in which objects of interest clustered toward the centre. Query objects could be entered directly onto this canvas, and their proximity to the center and to other objects was a relevance cue, facilitating real time feedback and exploration. More recently, Scells & Zuccon [20] developed a tool assist in formulating, visualising, and understanding Boolean queries. Their searchrefiner interface allows researchers to edit Boolean queries by dragging and dropping clauses in a structured editor. In addition, the tools provided by searchrefiner allow researchers to visualise why the queries they formulate retrieve citations, and ways to understand how to refine queries into more effective ones.

A further example is 2Dsearch [19], in which queries are formulated by manipulating objects on a two-dimensional canvas. Search results update in real-time and individual blocks with hit counts can be enabled/disabled on demand. Query suggestions are provided via an NLP services API, and support is offered for optimising and translating search strategies for different databases. Queries are analysed and validated, with common errors detected and corrections offered, and then stored as executable objects.

There have also been notable non-academic or practitioner-focused attempts to develop tools to support structured searching. Boolify[1] was one of the earliest examples, which allowed users to generate simple Boolean expressions by dragging terms and operators onto a 2D canvas. Boolio[2] offers a further variant, focusing on recruitment use cases and using a grid of rows and columns to allow to express disjunctions and conjunctions. Search Whiteboard[3] follows a similar approach, using the tabular structure of Excel spreadsheets to encode nested expressions as a series of rows and columns.

[1] https://www.kidzsearch.com/boolify/
[2] https://www.scoperac.com/boolio/
[3] https://exeterhealth.libguides.com/searching/Resources

Each of the above systems offers an alternative way to articulate complex information needs. In the next section, we review the collective insights they offer and explore ways in which future systems might better support explainability in PS.

5 Recommendations for Explainability

Let us now return to the criteria outlined earlier which define explainability in terms of the degree to which:

- an articulated information need produces the results that the user expected and intended
- an unarticulated information need may be parsimoniously articulated by the user with maximal fidelity and minimal error

Based on the issues identified in Section 3 and the insights provided by the solutions in Section 4, we propose the following initial design principles for explainabiity in PS. We also indicate (to the best of our knowledge) which of the systems described in Section 4 provide support for each principle (shown in parentheses).

1. Support transparency in the mapping between logical structure and physical structure:
   a. Allow users to express concepts and relationships using direct manipulation [6, 14, 15, 16, 18, 20, Boolify]
   b. Use visual cues to communicate conceptual structure [6, 14, 16, 20, Boolio, Search Whiteboard]
   c. Use interaction and animation to communicate cause and effect [6, 14, 15, 16, 18, 20, Boolify]
   d. Encourage exploration and query optimisation through real-time feedback [6, 14, 15, 18, 20]
2. Adopt scalable formalisms that accommodate complexity:
   a. Facilitate abstraction by allowing users to switch between overview and detail views, and to expand & collapse elements on demand [6, 20]
   b. Facilitate encapsulation by allowing users to independently manipulate and test sub-components of a composite search [6, 15, 17, 20, Search Whiteboard]
3. Delegate lower-level syntactic operations to system functions:
   a. Replace error-prone string manipulation with controlled object manipulation [6, 14, 15, 16, 17, 20]
   b. Provide automated translation of search syntax across databases, and semi-automated where appropriate, e.g. mapping of controlled vocabulary terms [6, 20]
4. Provide real-time feedback on query effectiveness
   a. Allow users to evaluate the contribution of individual query elements [6, 14, 15, 17, 20]
   b. Provide insights to help users understand how to make queries more effective [20]
5. Provide support for collaboration and team working:
   a. Facilitate versioning, sharing and peer-review
   b. Support repositories of best practice examples and templates [6]
   c. Provide automated support for search strategy reporting [20]

It is evident that none of the solutions in Section 4 addresses all of these principles. Moreover, it is our hope that additional principles will emerge following review and discussion of this initial set, and in that spirit we welcome feedback, refinement and suggestions.

6 Conclusion

Professional search is predicated on the resolution of complex information needs in a context of due diligence and accountability. This requires that search strategies be transparent, reproducible and explainable. We define explainability in terms of a mapping between an information need and its representation and a mapping between a representation and its effect. Current methods exhibit significant shortcomings regarding both of these mappings. In this paper we explore several alternative approaches and identify an initial set of design principles that could mitigate these shortcomings.

Explainability in PS is a complex and relatively under-investigated problem. We believe it deserves greater attention from the IR community, and we hope this paper serves to initiate a dialogue toward the development of more explainable approaches in professional search.

REFERENCES

[1] Jansen, B.J., Spink, A., Bateman, J. and Saracevic, T., 1998, April. Real life information retrieval: A study of user queries on the web. In ACM Sigir Forum (Vol. 32, No. 1, pp. 5-17). New York, NY, USA: ACM.
[2] Robertson, S.E., Walker, S., Jones, S., Hancock-Beaulieu, M.M. and Gatford, M., 1995. Okapi at TREC-3. Nist Special Publication Sp, p.109.
[3] Russell-Rose, T., Chamberlain, J. and Azzopardi, L., 2018. Information retrieval in the workplace: A comparison of professional search practices. Information Processing & Management, 54(6), pp.1042-1057
[4] Shokraneh, F., 2019. Reproducibility and replicability of systematic reviews. World Journal of Meta-Analysis, 7(3).
[5] MacFarlane, A. and Russell-Rose, T.G. 2016. Search Strategy Formulation: A Framework for Learning. Proceedings of 4th Spanish Conference in Information Retrieval, Grenada, 14-16 June 2016.
[6] Russell-Rose, T. and Shokraneh, F., 2020. Designing the Structured Search Experience: Rethinking the Query-Builder Paradigm. Weave: Journal of Library User Experience, 3(1).
[7] Yoo I, Mosa ASM. 2015. Analysis of PubMed User Sessions Using a Full-Day PubMed Query Log: A Comparison of Experienced and Nonexperienced PubMed Users. JMIR Med Inform3(3):e25.
[8] Hoang, L. and Schneider, J., 2018, Opportunities for computer support for systematic reviewing-a gap analysis. In International Conference on Information (pp. 367-377). Springer, Cham.
[9] van Altena, A.J., Spijker, R. and Olabarriaga, S.D., 2019. Usage of automation tools in systematic reviews. Research synthesis methods, 10(1), pp.72-82.
[10] O'Connor, A.M., Tsafnat, G., Thomas, J., Glasziou, P., Gilbert, S.B. and Hutton, B., 2019. A question of trust: can we build an evidence base to gain trust in systematic review automation technologies?. Systematic reviews, 8(1), p.143.
[11] Dijkstra, E.W., 1968. Letters to the editor: go to statement considered harmful. Communications of the ACM, 11(3), pp.147-148.

[12] Sampson, M. and McGowan, J., 2006. Errors in search strategies were identified by type and frequency. Journal of clinical epidemiology, 59(10), pp.1057-e1.
[13] Salvador-Oliván, J.A., Marco-Cuenca, G. and Arquero-Avilés, R., 2019. Errors in search strategies used in systematic reviews and their effects on information retrieval. Journal of the Medical Library Association: JMLA, 107(2), p.210.
[14] Anick, P. G., Brennan, J. D., Flynn, R. A., Hanssen, D. R., Alvey, B. and Robbins, J. M. 1990. A Direct Manipulation Interface for Boolean Information Retrieval via Natural Language Query. In Proceedings of the 13th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '90) . ACM, New York, NY, USA, 135–150.
[15] Fishkin, K and Stone, M. C. 1995. Enhanced Dynamic Queries via Movable Filters. ACM Press, 415–420.
[16] Jones, S. 1998. Graphical Query Specification and Dynamic Result Previews for a Digital Library. In Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology (UIST '98). ACM, New York, NY, USA, 143–151.
[17] Yi, J.S., Melton, R., Stasko, J. and Jacko, J. A. 2005. Dust & Magnet: Multivariate Information Visualization Using a Magnet Metaphor. Information Visualization 4, 4 (Oct. 2005), 239–256.
[18] Nitsche, M., and Nürnberger, A. 2006. QUEST: Querying Complex Information by Direct Manipulation. In: Yamamoto S. (eds) Human Interface and the Management of Information. Information and Interaction Design. HIMI 2013. Lecture Notes in Computer Science 8016 (2006).
[19] Russell-Rose, T., Chamberlain, J. and Shokraneh, F., 2019, A Visual Approach to Query Formulation for Systematic Search. In Proceedings of the 2019 Conference on Human Information Interaction and Retrieval (pp. 379-383).
[20] Scells, H., & Zuccon, G. (2018). searchrefiner: A Query Visualisation and Understanding Tool for Systematic Reviews. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management - CIKM '18 (pp. 1939–1942). New York, New York, USA: ACM Press. https://doi.org/10.1145/3269206.3269215