Real-world expertise retrieval: The information seeking behaviour of recruitment professionals

2 authors:

Tony Russell-Rose
UXLabs

91 PUBLICATIONS   1,682 CITATIONS

Jon Chamberlain
University of Essex

20 PUBLICATIONS   144 CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   DALI - Disagreements in Language Interpretation View project

Real-world expertise retrieval:

The information seeking behaviour of recruitment professionals

Tony Russell-Rose, Jon Chamberlain

UXLabs, London, UK

tgr@uxlabs.co.uk, jchamb@essex.ac.uk

Abstract. Recruitment professionals perform complex search tasks in order to find candidates that match client job briefs. In completing these tasks, they have to contend with many core Information Retrieval (IR) challenges such as query formulation and refinement and results evaluation. However, despite these and other similarities with more established information professions such as patent lawyers and healthcare librarians, this community has been largely overlooked in IR research. This paper presents results of a survey of recruitment professionals, investigating their information seeking behaviour and needs regarding IR systems and applications.

1   Introduction

Research into how people find and share expertise can be traced back to the 1960s, with early studies focusing on knowledge workers such as engineers and scientists and the information sources they consult [1]. Since then, the process of finding human experts (or *expertise retrieval*) has been studied in a variety of contexts and become the subject of a number of evaluation campaigns (e.g. the TREC Enterprise track and Entity Track [2, 3]). This has facilitated the development of numerous research systems and prototypes, and led to significant advances in performance, particularly against a range of system-oriented metrics [4].

However, in recent years there has been a growing recognition that the effectiveness of expertise retrieval systems is highly dependent on a number of contextual factors [5]. This has led to a more human-centred approach, where the emphasis is on how people search for expertise in the context of a specific task. These studies have typically been performed in an enterprise context, where the aim is to utilize human knowledge within an organization as efficiently as possible (e.g. [5, 6]).

However, there is a more ubiquitous form of expertise retrieval that embodies expert finding in its purest, most elemental form: the work of the professional recruiter. The job of a recruiter is to find people that are the best match for a client brief and return a list of qualified candidates. This involves the creation and execution of complex Boolean expressions, including nested, composite structures such as the following:

```
Java AND (Design OR develop OR code OR Program) AND ("*
Engineer" OR MTS OR "* Develop*" OR Scientist OR technol-
ogist) AND (J2EE OR Struts OR Spring) AND (Algorithm OR
"Data Structure" OR PS OR Problem Solving)
```

Over time, many recruiters create their own collection of queries and draw on these as a source of intellectual property and competitive advantage. Moreover, the creation of such expressions is the subject of many community forums (such as Boolean Strings and Undercover Recruiter) and the discussions that ensue involve topics that many IR researchers would recognise as wholly within their field of expertise (such as query expansion and optimisation, results evaluation, etc.).

However, despite these shared interests, the recruitment profession has been largely overlooked by the IR community. Even recent systematic reviews of professional search behaviour make no reference to this profession [7], and their information seeking behaviours remain relatively unstudied. This paper seeks to address that omission. We summarise the results of a survey of 64 recruitment professionals, examining their search tasks and behaviours, and the types of functionality that they value.

2   Background

We are aware of no prior work investigating the recruitment profession from an information seeking perspective. However, there are studies of other professions with related characteristics, such as Joho et al.'s [8] survey of patent searchers and Geschwandtner et al.'s [9] survey of medical professionals.

Unfilled vacancies have high impact on the economy, costing the UK £18bn annually [10]. Recruitment or *sourcing* is the process of finding capable applicants for those vacancies. It is a skill that is to some extent emulated by expert finding systems [4], although recruiters also must take into account contextual variables such as availability, previous experience, remuneration, etc.

Sourcing is also similar to people search on the web where the goal is to analyse large volumes of unstructured and noisy data to return a list of individuals who fit specific criteria [11]. The professional recruiter must normalise and disambiguate the returned results [2], and then apply additional factors to select a smaller group of qualified candidates. The gold standard for evaluation in this case is recommending one or more candidates that successfully fulfil a client brief.

3   Method

The survey instrument consisted of an online questionnaire of 40 questions divided into five sections[1]. It was designed to align with the survey instruments of Joho et al. [8] and wherever possible also with Geschwandtner et al. [9], to facilitate comparisons between the different professions. The five sections were as follows:

---

[1]   Available from https://isquared.wordpress.com/

1. **Demographics**: The background and professional experience of the respondents.
2. **Search tasks**: The types of search task that respondents perform in their work.
3. **Query formulation**: The approaches and techniques used to construct queries.
4. **Evaluation**: How they assess and evaluate the results of their search tasks.
5. **Ideal search engine**: Any other features additional to those described above.

The survey was designed to be completed in approximately 15 minutes. To obtain a large and representative sample we sent it out to interest groups via social media and also engaged with SurveyMonkey's panel of HR professionals based in North America. The survey ran from 09-Jun-2015 to 01-Aug-2015. We received 416 responses, of which 69 passed the qualifying question "*Is your primary job function to recruit and hire professionals for your organization or for clients*?" A further five were eliminated due to contradictory or nonsensical answers, which left 64 complete responses.

4   Results

4.1   Demographics

Of the 64 respondents, 69% were female and 31% male, with 54% of respondents aged between 25 and 45 years - a profile that is more female-oriented and younger than the patent and medical search survey respondents. Most respondents worked full time (91%), and the clients they worked for were predominantly external (48%) rather than internal (34%). This contrasts sharply with patent searchers, whose clients were predominantly internal (88%). Most respondents had several years' experience as a recruiter, with a median of 10 years, which aligns with that of the patent searchers.

4.2   Search tasks

We then examined the broader query lifecycle. In total, the majority of respondents (80%) used examples or templates at least sometimes; suggesting that the value embodied in such expressions is recognised and re-used wherever possible. In addition, most respondents (57%) were prepared to share queries with colleagues in their workgroup and a further 22% would share more broadly within their organisation. However, very few (5%) were prepared to share publicly, underlining the competitive nature of the industry. Job boards such as Monster, CareerBuilder and Indeed were the most commonly used databases (77%), although a similar proportion (73%) also targeted social networks such as LinkedIn, Twitter and Facebook.

Table 1 shows the amount of time that recruiters spend in completing their most frequently performed search task, the time spent formulating individual queries, and the number of queries they use. On average, it takes around 3 hours to complete a search task which consists of roughly 5 queries, with each query taking around 5 minutes to formulate. This suggests that recruitment follows a largely iterative paradigm, consisting of successive phases of candidate search followed by other activities such as candidate selection and evaluation. Compared to patent search the task completion time is less (3 hours vs. 12 hours) but is longer than typical web search tasks

[12]. Also, the number of queries is fewer (5 compared to 15) but the average query formulation time is the same (5 minutes).

Table 1. Search effort of recruitment professionals

	Min	Median	Max
Search task completion time (hrs)	0.06	3	30
Query formulation time (mins)	0.1	5	90
Number of queries submitted	1	5	50
Ideal number of results	1	33	1000
Number of results examined	1	30	100000
Time to assess relevance (mins)	1	5	50

4.3 Query formulation

In this section we examine the mechanics of the query formulation process, by asking respondents to indicate a level of agreement to various statements using a 5-point Likert scale ranging from strong disagreement (1) to strong agreement (5). The results are shown in Fig. 1 as a weighted average across all responses.
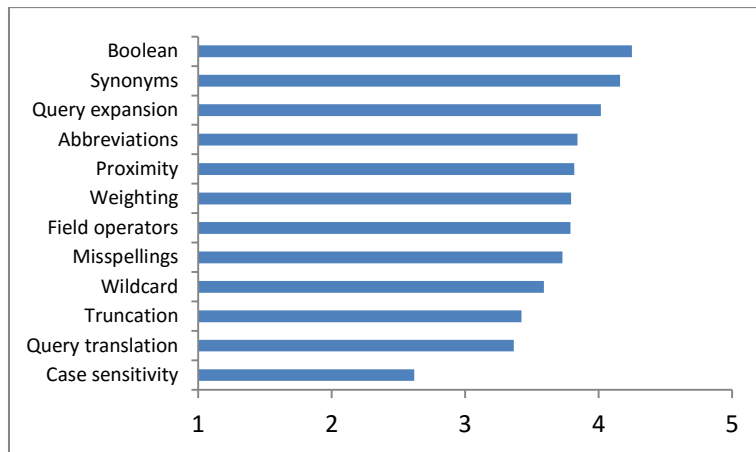

Fig. 1. Important query formulation features

The results suggest two observations in common with patent search. Firstly, the average of all but one of the features is above 3 (neutral), which suggests a willingness to adopt a wide range of search functionality to complete search tasks. Secondly, Boolean logic is shown to be the most important feature (4.25), closely followed by the use of synonyms (4.16) and query expansion (4.02). These scores indicate such functionality is desired by recruiters but the support offered by current search tools is highly variable. On the one hand, support for complex Boolean expressions is provided by many of the popular job boards. However, practical support for query formula-

tion and synonym generation is much more limited, with most current systems still relying instead on the expertise and judgement of the recruiter.

4.4 Results evaluation

In this section we examine the results evaluation process. Table 1 shows the ideal number of results returned, the number of results examined, and the time taken to assess relevance of a single result. Although the maximum probably represents outlier data, the median time to assess relevance of a single result is the same as that of the patent searchers (5 minutes). The number of results examined, however, is lower (30 vs. 100), suggesting that recruiters may adopt more of a satisficing strategy, evaluating only as many results as are required to create a shortlist of suitable candidates. This is supported by the median ideal number of returned results being similar (33).

We then asked respondents to indicate on a Likert scale how frequently they use various criteria to narrow down results. Job function was the most important (4.34), followed by location (4.29). These choices mirror some of the factors found to influence expert selection [6]. However, it contrasts with those of the medical searchers, who favoured content-based criteria such as type of source, date range, language, etc.

We also examined recruiters' strategies for interacting with results sets. The most popular approach was to start with the result that looked most relevant (56%). The number of respondents who targeted the most trustworthy source was relatively low (9%), which contrasts with the medical professionals and the claim [6] that "*source quality is the most dominant factor in the selection of human information sources*".

5 Discussion and conclusions

This paper summarises the results of a survey of the information seeking behaviour of recruitment professionals, uncovering their search needs in a manner that allows comparison with other, better-studied professions. In this section we briefly discuss the findings with verbatim quotes from respondents shown in italics where applicable.

Sourcing is shown to be something of a hybrid search task. The goal is essentially a people search task, but, the objects being returned are invariably documents (e.g. CVs and resumes), so the practice also shares characteristics of document search. Recruiters' display a number of professional search characteristics that differentiate their behaviour from web search [13], such as lengthy search sessions, different notions of relevance, different sources searched separately, and the use of specific domain knowledge: "*The hardest part of creating a query is comprehending new information and developing a mental model of the ideal search result.*"

Recruitment professionals use complex search queries, and actively cultivate skills in the formulation of such expressions. The search tasks they perform are inherently interactive, requiring multiple iterations of query formulation and results evaluation "*it is the limitations of available technology that force them to downgrade their concept tree into a Boolean expression*". In contrast with patent searchers, recruiter search behaviour is characterised by satisficing strategies, in which the objective is to

identify a sufficient number of qualified candidates in the shortest possible time *"Generally speaking, it's a trade-off between time and quality of results".* The average time spent evaluating a typical result was 5 minutes, rather than the 7 seconds reported in previous eye tracking studies [14].

These findings also have important consequences for the IR community and the assumptions underlying many of its research priorities. For example, much academic research continues to assume that searches are formulated as natural language queries, but this study shows that many professions prefer to formulate their queries as Boolean expressions [15]. In closing, we would hope that these findings may inspire the creation of new test collections focused on recruitment tasks, and thus facilitate the translation of IR research into real-world impact on a growing information profession.

6   References

1. H. Menzel. Information needs and uses in science and technology, *Annual Review of Information Science and Technology*, vol. 1, 41–69, 1966.
2. K. Balog, I. Soboroff, P. Thomas, N. Craswell, A. P. de Vries, and P. Bailey. Overview of the TREC 2008 Enterprise Track, In: *The Seventeenth Text Retrieval Conference Proceedings (TREC'08),* NIST, 2009.
3. K. Balog, P. Serdyukov, and A. P. de Vries. Overview of the TREC 2011 Entity Track, In: *Proceedings of the Twentieth Text REtrieval Conference (TREC'11),* 2012.
4. K. Balog, Y. Fang, M. de Rijke, P. Serdyukov, and L. Si. Expertise Retrieval, *Foundations and Trends in Information Retrieval*, 6(2-3):127-256, 2012.
5. K. Hofmann, K. Balog, T. Bogers, and M. de Rijke. Contextual factors for finding similar experts. *Information Science and Technology* 61, 5, 994-1014, 2010.
6. L. S. E. Woudstra and B. J. Van den Hooff. Inside the source selection process: Selection criteria for human information sources, *Information Processing and Management*, vol. 44, 1267–1278, 2008.
7. E. Vassilakaki and V. Moniarou-Papaconstantinou. A systematic literature review informing library and information professionals' emerging roles, *New Library World*, Vol. 116 Issue 1/2, 37–66, 2015.
8. H. Joho, L. Azzopardi and W. Vanderbauwhede. A survey on patent users: An analysis of tasks, behavior, search functionality and system requirements. In: *Proceedings of the 3rd Symposium on Information Interaction in Context (IIiX 2010)*, 2010.
9. M. Geschwandtner, Marlene Kritz, and Celia Boyer. D8.1.2: Requirements of the health professional search. Technical report, Khresmoi Project, 2011.
10. J. Cann. IOR Recruitment Sector Report: Report No.1 (UK), Institute of Recruiters, 2015.
11. Z. Guan, G. Miao, R. McLoughlin, X. Yan and D. Cai. Co-occurrence-based diffusion for expert search on the web. *Knowledge and Data Engineering*, IEEE Transactions on 25.5: 1001-1014. 2013
12. A. Broder. A taxonomy of web search. *SIGIR Forum*, 36(2):3–10, 2002.
13. M. Lupu, M. Salampasis, and A. Hanbury. Domain specific search. *Professional search in the modern world*. Springer International Publishing 96-117. 2014.
14. W. Evans. Eye tracking online metacognition: Cognitive complexity and recruiter decision making, TheLadders, 2012.
15. J. Tait: An introduction to professional search. *Professional search in the modern world,* Springer International Publishing 1-5., 2014