

First International Workshop on Professional Search

Suzan Verberne

Leiden University

s.verberne@liacs.leidenuniv.nl

Jiyin He

Signal Media

jiyin.he@signalmedia.co

Udo Kruschwitz

University of Essex

udo@essex.ac.uk

Gineke Wiggers

Leiden University

g.wiggers@law.leidenuniv.nl

Birger Larsen

Aalborg University

birger@hum.aau.dk

Tony Russell-Rose

UXLabs

tgr@uxlabs.co.uk

Arjen P. de Vries

Radboud University

arjen@acm.org

November 2, 2018

Abstract

In this report we describe the outcome of the First International Workshop on Professional Search, held in co-location with SIGIR 2018. The workshop addressed the specific requirements and challenges of professional search, as opposed to web search. The workshop included a survey held among 113 professional searchers, two keynote talks, six short paper talks, and two break-out sessions. From all parts of the workshop, we concluded that (1) we can define professional search with a set of common characteristics; but (2) there are definitely differences between domains; (3) we are in need of more knowledge about the search process in different domains, and (4) we need to collect data, not only Cranfield-style test collections, but also user observation data that help us understand the challenges and requirements of professional search better. Currently, there are many open problems that potentially have an impact on the IR community beyond the professional domains. We therefore call the scientific IR community to address professional search more actively.

1 Introduction

Professional search is a problem area challenging many facets of information retrieval, related to system aspects (e.g. distributed search), user aspects (e.g. complex information needs), as well as the interface between user and system (e.g. supporting exploratory search tasks). Professional search tasks have specific requirements, different from the requirements of generic web search.

The Professional Search workshop at SIGIR¹ addressed these requirements from multiple angles. We actively involved the target group of our research in the preparation of the workshop by distributing a survey in the months before the workshop. We defined the target group as “everyone who regularly performs complex search tasks at work in environments other than general web search”, including information specialists in various domains as well as librarians, scientists, lawyers, and other knowledge worker professions. The aim of the survey was to gather the experiences of these professional searchers: which systems do they use, what do they like about them, what functionalities do they miss, and what aspects/challenges do they think should be addressed by the academic IR community. The survey had 113 responses, and is discussed below in more detail (Section 3.1).

2 Overview of the talks

The workshop program consisted of a mix of keynotes, short oral presentations, and a break-out session during which participants focused on discussing the topics identified during the day. The workshop started with a presentation of the survey results (prepared by the organizers). Keynote speakers provided two different perspectives on professional search: Paul Bennett (Microsoft Research) on “Search and Recommendation in the Enterprise” and David D. Lewis (Brainspace) on “Research Challenges in IR for Legal Discovery and Investigations” We accepted six papers for short oral presentations.

2.1 Keynotes

Abstracts of the keynotes can be found in the proceedings.²

Paul Bennett: Search and Recommendation in the Enterprise

This keynote focused on information assistance and the challenge of understanding personal tasks and how they relate to information finding and task planning. Information assistance, like a great human assistant, should provide contextual intelligence by presenting the information you need before you ask for it, i.e. at the right time, place, and context; and addressing your information need in a way that relates to the task you are doing.

David D. Lewis: Research Challenges in IR for Legal Discovery and Investigations

This keynote focused on the application of information retrieval technology to documents and data relevant to legal procedures (e-discovery in litigation, open government requests, antitrust reviews, etc.). David Lewis gave many examples of IR research problems in e-discovery and investigations that are every bit as intellectually interesting as those in web-scale IR, even if they have been studied much less prominently.

2.2 Short papers

Automatic Shortlisting of Candidates in Recruitment

By Girish Keshav Palshikar, Rajiv Srivastava, Mahek Shah and Sachin Pawar [16]. Assessing

¹<https://jiyinhe.github.io/ProfS2018/>

²<http://ceur-ws.org/Vol1-2127/>

resumes is an important and effort-intensive task within human resources departments. This presentation described a system (TEAC) that creates standardized profiles based on resumés and the job description, and identifies a shortlist of candidates, along with an interpretable score for each. The resume scoring function uses domain-knowledge from recruitment experts as well as feedback about whether a given candidate was selected or rejected.

Searching for Relevant Lessons Learned Using Hybrid Information Retrieval Classifiers: A Case Study in Software Engineering

By Tamer Mohamed Abdellatif, Luiz Fernando Capretz and Danny Ho [1]. Software organizations can have a lessons learned (LL) repository, which contains unstructured information regarding previous working solutions for historical software management problems, or former success stories. In this presentation a previously created automatic information retrieval based LL classifier solution was extended by examining the impact of the hybridization of multiple LL classifiers on performance.

Refresh Strategies in Continuous Active Learning

By Nimesh Ghelani, Gordon V. Cormack and Mark D. Smucker [6]. Continuous Active Learning (CAL) is a technique where the model presents a human reviewer a set of documents likely to be relevant and the reviewer provides relevance feedback. During a refresh, the machine learning model is trained with the relevance judgments and a new list of likely-to-be-relevant documents is produced for the reviewer to judge. The presentation showed that more frequent refreshes can significantly reduce the human effort required to achieve a desired level of recall.

Validating the Importance of Work Tasks for Professional Search

By Thomas Schoegje, Toine Pieters and Egon L. van den Broek [18]. In professional search many work tasks share some structure and are likely to recur. Retrieval in the context of such tasks can exploit prior knowledge about these tasks to return more relevant results. This position paper presented initial work into this direction, discussing a filtering mechanism that narrows down results to the document types most relevant to the work task.

Challenges in the Development of Effective Systems for Professional Legal Search

By Piyush Arora, Murhaf Hossari, Alfredo Maldonado, Clare Conran and Gareth J. F. Jones [3]. The development and evaluation of information retrieval systems relies on the availability of suitable datasets and/or test collections. These typically consist of a target document collection, example search queries, and relevance data. For professional search applications domain experts are required to gather this information. This presentation showed the vital need for close interaction between the professionals for which the application is being developed and the IR researchers throughout the development life cycle of the search system.

Explainable IR for personalizing professional search

By Suzan Verberne [20]. This position paper established the need for transparency in personalized professional search. The central idea of the paper is the professional personal knowledge graph. Graphs are a natural and transparent means of representing knowledge,

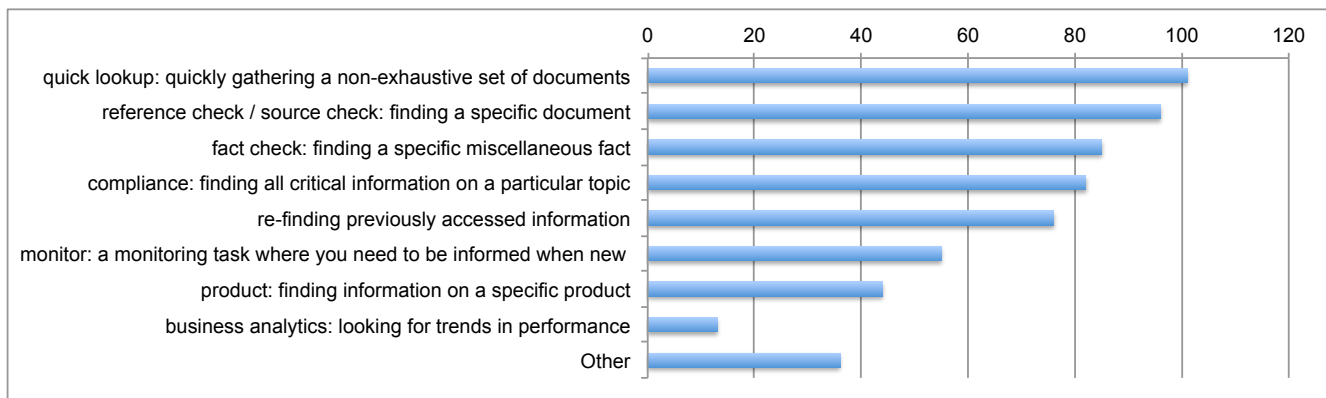


Figure 1: What type of search tasks do you undertake? (select all that apply)

and learning a personal graph from sparse user interaction data is a new research direction. A graph-based search paradigm would enable and stimulate the exploratory search behaviour for complex information needs that are inevitable in professional work environments.

3 Themes and challenges

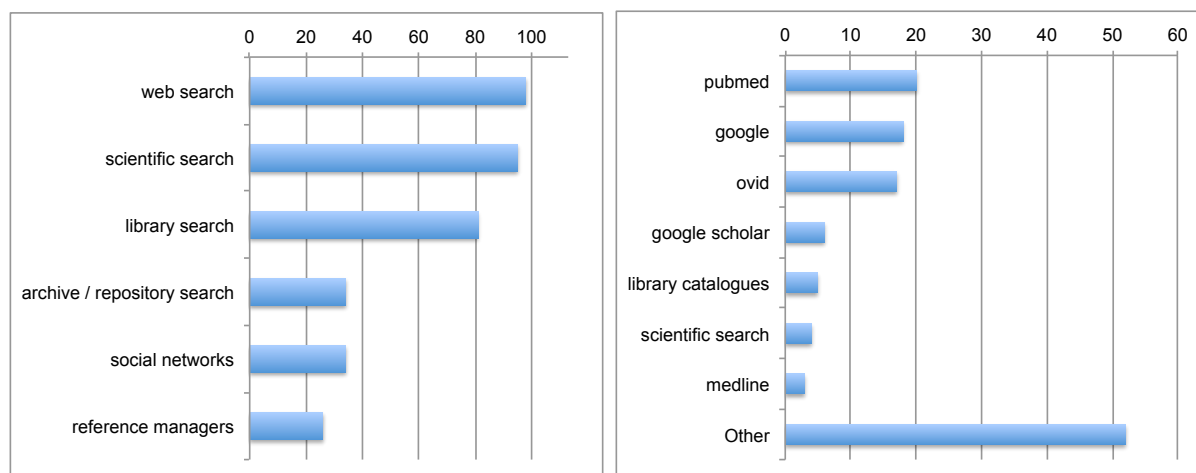
As mentioned, the workshop organizers conducted a survey to ensure that we would have sufficient eye for the practitioners’ view on search during our discussions.

3.1 Discussion and implications of the survey results

The survey attracted 113 respondents. 56 of the participants were librarians and information specialists; 12 were researchers or professors. All other job titles were mentioned three times or fewer (a surprisingly long tail). In this report, we show the results for three questions that we consider to be the most relevant: (1) which search tasks do professionals undertake; (2) which search systems do they use; (3) what would they like us as academic community to focus our research on.

Figure 1 shows the distribution of answers to the question which type of search tasks the respondents undertake. We consider the answers ‘compliance’ and ‘re-finding’ to be of particular interest. Both are mentioned by the majority of respondents and are clearly distinct from Web search (unlike the top 3 tasks: quick lookup, reference check, and fact check). Refinding has long been known to be an important issue. Dumais and colleagues at Microsoft Research observed that much knowledge work involves integrating and re-using information that has previously been created or accessed. They proposed the *Stuff I’ve Seen* (SIS) system that addresses this very issue [5]. Re-finding is even more of a problem within a professional search context than in general Web search given the difficulty in finding the right information in the first place, a commonly reported problem in enterprise search [11].

Compliance is a core information professional issue. Compliance search, finding all critical information on a topic, can be seen a type of *discovery* task known to prominently feature in enterprise search activities. The Enterprise Search survey conducted by the Association for Information and Image Management (AIIM) found, for example, that 50% of their respondents report that they deal with internal compliance audits and 44% with pre-trial



(a) When you search for information for professional use, which search engines/systems do you use frequently? (open question)
 (b) Which search system do you use the most frequently for work-related search? (open question)

Figure 2: Search systems used by the participants.

legal discovery [15]. Other tasks in which ‘finding all critical information’ is important are systematic review and prior art search.

Figure 2 shows an overview of the responses to the two questions addressing the search systems used by the participants. The options for the question under Figure 2(a) were overarching categories of search systems. With the (open) question under Figure 2(b) we aimed at one particular search system, but some participants answered with an type of system instead (‘scientific search’, ‘library catalogues’). Figure 2(b) shows a surprisingly long tail. This indicates that professional searchers have individual preferences as to which systems they use, reflecting the diversity of ‘the’ professional search community: a key challenge for the academic community is to find unifying themes and patterns, and in executing cross-cutting studies.

Figure 3 shows the counts for the options that we gave the participants for the question “What would you like us, as Information Retrieval researchers, to focus our research on?” (multiple answers could be selected). The figure indicates that the majority of respondents want to see more integrated search environments. This finding is not at all surprising given that the silo-based architecture in professional search contexts is frequently being flagged up as a problem. For example, in the context of enterprise search, thousands of different databases may be of relevance for different tasks [4]. The integration across different repositories, the need for access control, the different structures underlying the databases (from unstructured to fully structured) makes this a non-trivial exercise [11]. With reference to the AIIM Enterprise Search survey mentioned earlier, only 11% of organizations reported to have a fully joined up search approach and another 18% across departmental content [15].

The need for integrated search environments appears to first of all highlight a very practical engineering problem, but its impact is wider and touches upon many topics of academic interest, including federation, merging, presentation etc. of different data sources (see e.g. [2]).

After a discussion of the survey results we decided on two topics for the break-out sessions:

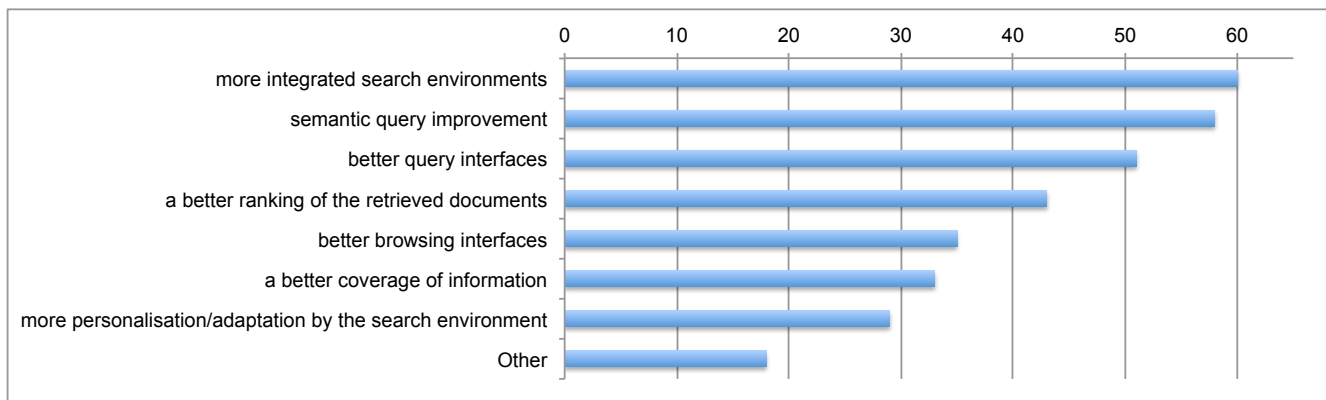


Figure 3: Distribution of answers for the question “What would you like us, as Information Retrieval researchers, to focus our research on?” The full answer for more integrated search environments is “more integrated search environments (to be able to find everything behind one interface)”

(1) commonalities between domains, triggered by the long tail of responses to questions related to job title, field, and search engines used; and (2) data and evaluation, because we can only work towards improving professional search applications if we have access to representative data and evaluation protocols. These two topics are discussed in the next two subsections.

3.2 Commonalities between domains

Most IR research addresses web search or zooms in to a highly specific domain. We are interested in professional search as a common problem area. The reality is much more challenging than the cases we end up studying in detail. Can professional search be seen as a single problem domain? We first discussed definitions of professional search, and then addressed commonalities and differences in the search process.

Definition. Definitions of professional search in the scientific literature seem to focus on descriptive and behavioural characteristics [17, 8, 9, 10, 19]: *professional search takes place in the work context, by specialists, and using specialist sources, often with controlled vocabularies.*

The search process. A common aspect of professional search across domains is the need for professionals to be in control, explaining the predominance of Boolean search in, e.g., prior art search and systematic review. Information professionals often still prefer Boolean search for transparency and the need to reproduce their results later on, in specific cases in the courtroom.

In addition, the search by professionals is highly iterative and there is a tolerance to low precision, especially in early stages of the work process. Professionals formulate lengthy sequences of queries, often to multiple different databases, while they follow an exploratory search process [7]. This was also illustrated by the example search tasks provided by participants of our survey. Three replies to the question “Describe which actions you took to

search” read:

- “Searched MEDLINE and Embase on the Ovid platform 00 used controlled vocabulary and text words for the concepts: fever, ibuprofen, paracetamol, child. Did not use age limit with database for MEDLINE as I was including the epub ahead of print, in process and non-indexed subject headings.”
- “Looked for high quality evidence as per the Evidence pyramid - Systems, Syntheses, Systematic Review/Meta-Analysis; RCTs; Cohort; Case Control; Case Series; at the bottom: Expert Opinion”
- “systematic search (analysing research question, selecting key elements, formulating search terms); search in library catalogue, migratierecht.nl, vluchtweb.nl; Web of Sciences; Google scholar and google”

Another challenge for search tasks spanning multiple and diverse indexes is that not all data sources are equally accessible. Some require subscription, many have specialized interfaces (‘advanced search’), that are different in options and user interface, and collecting a set of results from multiple sources remains an involving process due to the lack of interoperability between systems.

We specifically discussed the topic of controlled vocabularies, which are commonly used in some specialized domains. Thesauri and ontologies that constitute a controlled vocabulary are commonly used in the medical and legal domains, but for example not in e-discovery . While not always made explicit in a controlled vocabulary, the scope of professional search tasks is clearly more constrained than “all information” (the web). Aspects of restricted scope also play a role in enterprise search, email search, and community search. There seems a lot of opportunity for the usage of context informed by document usage and search history.

Conclusion. Looking into the survey and our discussions we conclude that (1) we can define professional search using a set of common characteristics; (2) these common characteristics mainly refer to aspects of the search process; (3) there are definitely differences between tasks/domains, but we do not know sufficient yet about the particularities of the diverse domains to give a definite answer to the question ‘Can professional search be seen as a single problem domain?’

3.3 Data and evaluation

There are only very limited test collections available for professional search; in some domains there are no test collections at all, or they are only partly relevant. For example, there are collections without relevance assessments; information needs without collection; or only one type of documents instead of the complete variety that is representative of the problem. We addressed challenges in creating test collections for professional search, and made a first attempt to define what we need as test collections.

Challenges in creating test collections for professional search. The first challenge is that there is a large diversity in information types for professional search problems. Typical examples of domains with many multiple different document types are enterprise search, legal search, and academic search. In addition, many domains are dealing with privacy-sensitive data. The second challenge is that a document collection can only be used

as test collection if queries and relevance assessments are added. The collection of relevance assessments poses challenges because experts are expensive, and because for highly specific search tasks the relevance assessments are even more subjective (context-dependent) than for web search. Even if we knew how to collect benchmark data for research in professional search, a central question remains how generalizable the data is (see Section 3.2).

What do we need? The academic field of professional search is in need for benchmark data, as well as user observation studies and log data. The data should reflect the actual search problems and the topics should be temporarily grounded. Some of the useful, rich, collections are relatively small. Examples are the iSearch collections for academic search [14] (65 topics) and the Queensland dataset for case law search [13] (12 topics). For some purposes this is sufficient, but for quantitative system comparisons we need more topics, and a diversity of topics. As for now, it is still an open question what constitutes a sufficiently diverse data set to evaluate our methods.

Conclusion. In the case of professional search there is so much we do not know yet. Also small, qualitative data can be valuable. To put it even stronger: *We are hungry for anecdotes*, because we need to learn, case by case, how to approach the study of difficult search tasks in a realistic setting. Although qualitative studies are definitely interesting for the field, they tend to be difficult to publish, and perhaps we need to alter our own reviewing practices to make more space for exploratory work that ‘just’ sets the scene.

4 Outlook/follow-up

Concluding this workshop summary, we want to emphasize the importance of the area of research.

Professional search has the key benefit that the task to be solved is, usually, clear; at least to the person who carries out the searches. And, the gains of doing better can be quantified in terms of increased profits of their business. But, looking back to the enlightening contribution presented by David Lewis, there are so many things that we academics assume have been solved, for which, as soon as you take the problem out of the context of a very small set of nicely curated test collections, those presumably simple steps turn out as huge, open and unsolved problems: “we don’t know s**** (search)”.³ Addressing problems arising in professional search will thus have much wider impact, and lead to insights that improve our understanding of fundamental components that matter for every information access problem that we might care about.

So, join us in our quest, leave the Web to the big corporations, and let academia focus on tackling problems in professional search!

References

- [1] Abdellatif, T.M., Capretz, L.F., Ho, D.: Searching for Relevant Lessons Learned Using Hybrid Information Retrieval Classifiers: A Case Study in Software Engineering. In: ProfS/KG4IR/Data: Search@ SIGIR. pp. 12–17 (2018)

³We highly recommend the related short paper presented at the new DESIRES conference [12].

-
- [2] Arguello, J.: Aggregated Search. *Foundations and Trends in Information Retrieval* **10**, 365–502 (2017)
- [3] Arora, P., Hossari, M., Maldonado, A., Conran, C., Jones, G.J., Paulus, A., Klostermann, J., Dirschl, C.: Challenges in the Development of Effective Systems for Professional Legal Search. In: *ProfS/KG4IR/Data: Search@ SIGIR*. pp. 29–34 (2018)
- [4] Cortez, E., Bernstein, P.A., He, Y., Novik, L.: Annotating Database Schemas to Help Enterprise Search. *PVLDB* **8**(12), 1936–1939 (2015)
- [5] Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., Robbins, D.C.: Stuff i’ve seen: A system for personal information retrieval and re-use. In: *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval*. pp. 72–79. *SIGIR ’03*, ACM (2003)
- [6] Ghelani, N., Cormack, G.V., Smucker, M.D.: Refresh Strategies in Continuous Active Learning. In: *ProfS/KG4IR/Data: Search@ SIGIR*. pp. 18–23 (2018)
- [7] He, J., Bron, M., de Vries, A.P.: Characterizing stages of a multi-session complex search task through direct and indirect query modifications. In: *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*. pp. 897–900. ACM (2013)
- [8] Jankowski, T.A.: *Expert Searching in the Google Age*. Rowman & Littlefield (2016)
- [9] Kim, Y., Seo, J., Croft, W.B.: Automatic boolean query suggestion for professional search. In: *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. pp. 825–834. ACM (2011)
- [10] Koster, C., Oostdijk, N., Verberne, S., D’hondt, E.: Challenges in professional search with phasar. In: *Proceedings of the Dutch-Belgian Information Retrieval workshop (DIR 2009)*. pp. 101–102. Enschede, Nederland:[sn] (2009)
- [11] Kruschwitz, U., Hull, C.: Searching the enterprise. *Foundations and Trends in Information Retrieval* **11**(1), 1–142 (2017). <https://doi.org/10.1561/15000000053>, <http://dx.doi.org/10.1561/15000000053>
- [12] Lewis, D.: We don’t know s***** (search). In: Alonso, O., Silvello, G. (eds.) *Proceedings of the First Biennial Conference on Design of Experimental Search & Information Retrieval Systems*, Bertinoro, Italy, August 28-31, 2018. *CEUR Workshop Proceedings*, vol. 2167, p. 107. *CEUR-WS.org* (2018), <http://ceur-ws.org/Vol-2167/short10.pdf>
- [13] Locke, D., Zuccon, G.: A test collection for evaluating legal case law search. In: *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. pp. 1261–1264. *SIGIR ’18*, ACM, New York, NY, USA (2018). <https://doi.org/10.1145/3209978.3210161>, <http://doi.acm.org/10.1145/3209978.3210161>
- [14] Lykke, M., Larsen, B., Lund, H., Ingwersen, P.: Developing a test collection for the evaluation of integrated search. In: *European Conference on Information Retrieval*. pp. 627–630. Springer (2010)
- [15] Miles, D.: *AIIM Industry Watch: Search and Discovery - Exploiting Knowledge, Minimizing Risk*. AIIM (2014)
- [16] Palshikar, G.K., Srivastava, R., Shah, M., Pawar, S.: Automatic Shortlisting of Candidates in Recruitment. In: *ProfS/KG4IR/Data: Search@ SIGIR*. pp. 5–11 (2018)

-
- [17] Russell-Rose, T., Chamberlain, J., Azzopardi, L.: Information retrieval in the workplace: A comparison of professional search practices. *Information Processing & Management* **54**(6), 1042–1057 (2018)
 - [18] Schoegje, T., Pieters, T., van den Broek, E.L.: Validating the Importance of Work Tasks for Professional Search. In: *ProfS/KG4IR/Data: Search@ SIGIR*. pp. 24–28 (2018)
 - [19] Tait, J.I.: An introduction to professional search. In: *Professional search in the modern world*, pp. 1–5. Springer (2014)
 - [20] Verberne, S.: Explainable IR for Personalizing Professional Search. In: *ProfS/KG4IR/Data: Search@ SIGIR*. pp. 35–42 (2018)