Expertise Retrieval

Expertise Retrieval

Krisztian Balog

NTNU, Norway
krisztian.balog@idi.ntnu.no

Yi Fang

Purdue University, USA
yfang.purdue@gmail.com

Maarten de Rijke

University of Amsterdam, The Netherlands
derijke@uva.nl

Pavel Serdyukov

Yandex, Russia
pavser@yandex-team.ru

Luo Si

Purdue University, USA
lsi@purdue.edu

now
the essence of knowledge

Boston – Delft

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is K. Balog, Y. Fang, M. de Rijke, P. Serdyukov and L. Si, Expertise Retrieval, Foundation and Trends® in Information Retrieval, vol 6, nos 2–3, pp 127–256, 2012

ISBN: 978-1-60198-566-8

Foundations and Trends® in Information Retrieval
Volume 6 Issues 2–3, 2012

Editorial Board

Editorial Scope

Foundations and Trends® in Information Retrieval will publish survey and tutorial articles in the following topics:

- Applications of IR
- Architectures for IR
- Collaborative filtering and recommender systems
- Cross-lingual and multilingual IR
- Distributed IR and federated search
- Evaluation issues and test collections for IR
- Formal models and language models for IR
- IR on mobile platforms
- Indexing and retrieval of structured documents
- Information categorization and clustering
- Information extraction
- Information filtering and routing
- Metasearch, rank aggregation and data fusion
- Natural language processing for IR
- Performance issues for IR systems, including algorithms, data structures, optimization techniques, and scalability
- Question answering
- Summarization of single documents, multiple documents, and corpora
- Text mining
- Topic detection and tracking
- Usability, interactivity, and visualization issues in IR
- User modelling and user studies for IR
- Web search

Information for Librarians
Foundations and Trends® in Information Retrieval, 2012, Volume 6, 5 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Foundations and Trends® in
Information Retrieval
Vol. 6, Nos. 2–3 (2012) 127–256

DOI: 10.1561/1500000024

now
the essence of knowledge

Expertise Retrieval

Krisztian Balog[1], Yi Fang[2], Maarten de Rijke[3], Pavel Serdyukov[4] and Luo Si[5]

[1] *Department of Computer and Information Science, NTNU, Sem Sælands vei 7-9, Trondheim, 7491, Norway, krisztian.balog@idi.ntnu.no*
[2] *Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA, yfang.purdue@gmail.com*
[3] *Intelligent Systems Lab Amsterdam, University of Amsterdam, Science Park 904, Amsterdam, 1098 XH, The Netherlands, derijke@uva.nl*
[4] *Yandex, Lva Tolstogo, 16, Moscow, 119021, Russia, pavser@yandex-team.ru*
[5] *Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA, lsi@purdue.edu*

Abstract

People have looked for experts since before the advent of computers. With advances in information retrieval technology and the large-scale availability of digital traces of knowledge-related activities, computer systems that can fully automate the process of locating expertise have become a reality. The past decade has witnessed tremendous interest, and a wealth of results, in expertise retrieval as an emerging subdiscipline in information retrieval. This survey highlights advances in models and algorithms relevant to this field. We draw connections among methods proposed in the literature and summarize them in five groups of basic approaches. These serve as the building blocks for more advanced

models that arise when we consider a range of content-based factors that may impact the strength of association between a topic and a person. We also discuss practical aspects of building an expert search system and present applications of the technology in other domains, such as blog distillation and entity retrieval. The limitations of current approaches are also pointed out. We end our survey with a set of conjectures on what the future may hold for expertise retrieval research.

Contents

1 Introduction **1**

1.1 The Need for Expertise Retrieval 1
1.2 Challenges in Expertise Retrieval 3
1.3 Organization 4

2 Background **7**

2.1 Expertise Retrieval vs. Expertise Seeking 7
2.2 Early Work 8
2.3 Expertise Retrieval in Information Retrieval 9
2.4 Related Topics and Tasks 12

3 Expertise Retrieval Tasks **15**

3.1 Expertise Retrieval Systems 15
3.2 Two Tasks: Expert Finding and Expert Profiling 18

4 Evaluation **23**

4.1 Measures 23
4.2 Test Collections 24

5 Approaches to Expertise Retrieval **33**

5.1 A Brief Roadmap to Expertise Retrieval Approaches 33

5.2 Generative Probabilistic Models 37
5.3 Discriminative Probabilistic Models 47
5.4 Voting Models 55
5.5 Graph-based Models 60
5.6 Other Models 68
5.7 Comparison of Approaches 71

6 Advanced Components 73

6.1 Document-candidate Associations 73
6.2 Query Modeling, Relevance Feedback 79
6.3 Document Importance 82
6.4 Document Structure 85
6.5 External Evidence 87
6.6 Candidate Importance 89

7 Discussion 93

7.1 Practical Considerations 93
7.2 Limitations of Current Approaches 98
7.3 Relation Between Document Retrieval and Expert Finding 100
7.4 Broader Applicability 102
7.5 Emerging Challenges 104

8 Conclusions 109

Acknowledgments 115

References 117

1

Introduction

> Believe one who has proved it. Believe an expert.
> —Virgil (70 BC–19 BC), *Aeneid*

1.1 The Need for Expertise Retrieval

We call for an expert when we need someone to show us the right path to tackle a problem. There may be large volumes of information available around the problem at hand, but we need an expert to help us find our way. Sometimes the required knowledge is just not freely accessible in digital format. The information that is available might be hard to express in writing or it may be difficult to analyze.

Experts can be in demand not only for being asked for questions, but also for being assigned some role or job in an organizational setting. For instance, conference organizers may search for teams of reviewers, recruiters for talented employees, and consultants may look for other consultants to redirect inquiries and decrease the risk of losing clients.

Research on how to enable people to effectively share expertise can be traced back to at least the 1960s when studies in library and information science explored what sources of information knowledge workers

like researchers and engineers use [152]. Subsequent work identified complex information seeking strategies relying on a variety of information sources, including human experts [98, 180]. From results of this type of research grew the realization that the expertise of employees is a major value of an organization and that effective sharing of knowledge can lead to material gains [52, 63, 218].

How do we locate expertise? Relatively early on, the field of knowledge management developed, with the goal of using knowledge within an organization as well as possible. One focus was on developing information systems that could support search for expertise. Initial approaches were mainly focused on how to unify disparate and dissimilar databases of the organization into a single data warehouse that could easily be mined [79, 187]. Resulting tools relied on people to self-assess their skills against a predefined set of keywords, and often employed heuristics generated manually based on current working practice.

Despite the achievements made so far, the question of how to provide effective access to expertise is far from solved, and continues to be addressed from different viewpoints. It has been found that a standard document search engine may be of great help [99], but does not directly address this task: it returns documents, not people. Even in a professional environment, many of us still just "ask around" [101]. According to independent research carried out by Vanson Bourne, who assessed the current information capabilities and needs of 170 organizations in the United Kingdom with more than 1,000 employees, only 55 percent of professional service employees and a mere 27 percent of public sector employees are able to locate expertise using their current enterprise search systems, while 50 percent or more of those surveyed want to be able to locate expertise on a daily basis [167].

But today, as we increasingly live our professional lives online, evidence of expertise can be traced, mined, and organized. Over the past decade, this development, together with the increasingly distributed nature of our working environments, has led to renewed interest in two types of expertise retrieval system: *expert finding* systems (that help answer information needs such as "Find me someone who is an expert on X") and *expert profiling* systems (that help answer information needs such as "Tell me in which topics this person is an expert").

We refer to the general area of linking humans to expertise areas, and vice versa, as *expertise retrieval*.

Expertise retrieval (also known as expertise location or expertise identification) is traditionally regarded as a subject of research in information retrieval. And often, expertise retrieval is taken to mean "expertise retrieval *within a specific organization*." Expertise retrieval is a part of the functionality of a typical enterprise search system, which usually operates within the scope of a single company.

1.2 Challenges in Expertise Retrieval

Finding an expert is a challenging task because *expertise* is a loosely defined concept that is hard to formalize. It is common to refer to expertise as "tacit knowledge" [33], the type of knowledge that people carry in their minds and which is, therefore, difficult to access. It is often contrasted with "explicit knowledge," which is already captured, described, documented, and stored. However, the only way for an expert finding system to assess and access "tacit knowledge" in organizations is through artifacts of "explicit knowledge" (e.g., documents). As a consequence, expertise retrieval inherits many challenges from document retrieval, but there is more to it than just document retrieval.

Expertise retrieval brings new challenges over and above the challenges usually associated with document retrieval. We list the key challenges:

- Candidate experts are usually not represented as retrievable units: they are identified indirectly through the texts associated with them, through authorship, mentions, or citations. We discuss this issue throughout the survey, when talking about test collections (Section 4), about advanced models (Section 6), and about practical considerations (Section 7.1).
- Moreover, expert names are often ambiguous: mentions might be incomplete and a single name may belong to multiple people, even within a single organization; this issue is discussed in Sections 2.4 and 6.1.
- Also, expertise evidence often comes from heterogeneous sources, not all of which are equally important: a brief email

probably carries a different weight than a technical standards document. The heterogeneous nature of expertise evidence is discussed throughout the survey, for instance in Sections 2.3, as part of our discussions of models for expertise retrieval in Section 5, in advanced components that consider document importance and structure (Sections 6.3 and 6.4), and as part of a discussion of practical considerations on expertise retrieval (Section 7.1).

- And finally, determining the strength of the association between a candidate expert and textual evidence of his or her expertise is a complex decision as well. It is the core focus of much of the modeling work presented in Sections 5 and 6.

The challenges listed above make expertise retrieval a multi-faceted research area and building a state-of-the-art expert finding system consists of many steps, each bringing its own scientific challenges.

While there are a number of tasks and problems related to expertise retrieval (see Section 2.4 for a collection of them), to maintain a clear focus throughout this survey, we center our attention around scenarios where an individual wants to contact an expert (as opposed to, for example, building a team of experts). Moreover, we primarily limit ourselves to *topical* aspects of the tasks, thereby largely abstracting away from *cognitive* and *social* considerations. While these are indeed interesting directions, they are research areas on their own and discussing them in detail is beyond the scope of this survey; we highlight, however, some of these works that are of particular relevance within the scope of our survey in Section 7.5.

1.3 Organization

With the multi-faceted nature of expertise retrieval in mind, we structure this survey in the following manner.

To begin, Section 2 focuses on the roots of research on expertise retrieval that served as an inspiration for the approaches developed in the 2000s that form the bulk of the material covered in this survey. Section 3 then introduces the primary tasks on which we focus in this

survey: the tasks of expertise retrieval and expertise profiling. Section 4 provides an overview of test collections and evaluation methodology commonly accepted in the research community. Section 5 continues with an overview of approaches and includes probabilistic models (generative and discriminative), voting models, graph-based models, as well as methods that do not fall under any of these headings. Extensions of these models are discussed in Section 6. Section 7 discusses practical considerations including the limitations of current expertise retrieval approaches and recent work aimed at addressing some of them. Finally, Section 8 concludes the survey and identifies some future research directions.

References

[1] M. S. Ackerman and D. W. Mcdonald, "Collaborative support for informal information in collective memory systems," *Information Systems Frontiers*, vol. 2, no. 3–4, pp. 333–347, 2000. ISSN 1387-3326.

[2] L. A. Adamic, J. Zhang, E. Bakshy, and M. S. Ackerman, "Knowledge sharing and Yahoo! Answers: Everyone knows something," in *Proceedings of the International World Wide Web Conference*, *(WWW '08)*, New York, NY, USA: ACM, pp. 665–674, 2008. ISBN 978-1-60558-085-2.

[3] A. V. Aho and M. J. Corasick, "Efficient string matching: An aid to bibliographic search," *Communications ACM*, vol. 18, pp. 333–340, June 1975.

[4] G. Amati, "Probabilistic models for information retrieval based on divergence from randomness," PhD thesis, Department of Computing Science, University of Glasgow, 2003.

[5] J. Artiles, A. Borthwick, J. Gonzalo, S. Sekine, and E. Amigó, "Weps-3 evaluation campaign: overview of the web people search clustering and attribute extraction tasks," in *Cross-Language Evaluation Forum 2010 Working Notes*, 2010.

[6] J. Artiles, J. Gonzalo, and S. Sekine, "The SemEval-2007 WePS evaluation: Establishing a benchmark for the web people search task," in *Proceedings of the International Workshop on Semantic Evaluations*, *(SemEval '07)*, Stroudsburg, PA, USA: ACL, pp. 64–69, 2007.

[7] J. Artiles, J. Gonzalo, and S. Sekine, "WePS 2 evaluation campaign: Overview of the web people search clustering task," in *Web People Search Evaluation Workshop (WePS 2009), WWW Conference*, 2009.

[8] J. Artiles, J. Gonzalo, and F. Verdejo, "A testbed for people searching strategies in the WWW," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '05)*, New York, NY, USA, pp. 569–570, 2005. ISBN 1-59593-034-5.

[9] S. Asmussen, *Applied Probability and Queues*. Wiley, 1987. ISBN 0387002111.

[10] P. Bailey, N. Craswell, A. P. de Vries, and I. Soboroff, "Overview of the TREC 2007 enterprise track," in *Proceedings of the Text REtrieval Conference, (TREC '07)*, Gaithersburg, MD, 2008.

[11] P. Bailey, N. Craswell, I. Soboroff, and A. de Vries, "The CSIRO enterprise search test collection," *ACM SIGIR Forum*, vol. 41, pp. 42–45, 2007.

[12] K. Balog, "People search in the enterprise," PhD thesis, University of Amsterdam, June 2008.

[13] K. Balog, L. Azzopardi, and M. de Rijke, "Formal models for expert finding in enterprise corpora," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '06)*, New York, NY, USA, pp. 43–50, 2006. ISBN 1-59593-369-7.

[14] K. Balog, L. Azzopardi, and M. de Rijke, "A language modeling framework for expert finding," *Information Processing and Management*, vol. 45, pp. 1–19, January 2009. ISSN 0306-4573.

[15] K. Balog, L. Azzopardi, and M. de Rijke, "Resolving person names in web people search," in *Weaving Services, Locations, and People on the World Wide Web*, Springer, pp. 301–323, 2009. ISBN 978-3-642-00569-5.

[16] K. Balog, T. Bogers, L. Azzopardi, M. de Rijke, and A. van den Bosch, "Broad expertise retrieval in sparse data environments," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '07)*, New York, NY, USA, pp. 551–558, 2007. ISBN 978-1-59593-597-7.

[17] K. Balog, M. Bron, and M. de Rijke, "Query modeling for entity search based on terms, categories, and examples," *ACM Transactions on Information Systems*, vol. 29, November 2011. ISSN 1046-8188.

[18] K. Balog and M. de Rijke, "Finding experts and their details in e-mail corpora," in *Proceedings of the International World Wide Web Conference, (WWW '06)*, New York, NY, USA, pp. 1035–1036, 2006. ISBN 1-59593-323-9.

[19] K. Balog and M. de Rijke, "Determining expert profiles (with an application to expert finding)," in *Proceedings of the International Joint Conference on Artifical Intelligence, (IJCAI '07)*, San Francisco, USA, pp. 2657–2662, 2007.

[20] K. Balog and M. de Rijke, "Finding similar experts," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '07)*, New York, NY, USA, pp. 821–822, 2007. ISBN 978-1-59593-597-7.

[21] K. Balog and M. de Rijke, "Associating people and documents," in *Proceedings of the European Conference on IR Research*, (C. M. et al., ed.), *(ECIR '08)*, Berlin, Heidelberg, pp. 296–308, 2008. ISBN 978-3-540-78645-0.

[22] K. Balog and M. de Rijke, "Non-local evidence for expert finding," in *Proceeding of the ACM International Conference on Information and Knowledge*

Management, *(CIKM '08)*, New York, NY, USA, pp. 489–498, October 2008. ISBN 978-1-59593-991-3.

[23] K. Balog and M. de Rijke, "Combining candidate and document models for expert search," in *Proceedings of the Text REtrieval Conference*, *(TREC '08)*, Gaithersburg, MD, 2009.

[24] K. Balog, M. de Rijke, and W. Weerkamp, "Bloggers as experts: Feed distillation using expert retrieval models," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, *(SIGIR '08)*, New York, NY, USA, pp. 753–754, 2008. ISBN 978-1-60558-164-4.

[25] K. Balog, A. P. de Vries, P. Serdyukov, P. Thomas, and T. Westerveld, "Overview of the TREC 2009 Entity Track," in *Proceedings of the Text REtrieval Conference*, *(TREC '09)*, Gaithersburg, MD, February 2010.

[26] K. Balog, K. Hofmann, W. Weerkamp, and M. de Rijke, "The University of Amsterdam at the TREC 2007 enterprise track," in *Proceedings of the Text REtrieval Conference*, *(TREC '06)*, Gaithersburg, MD, 2007.

[27] K. Balog, K. Hofmann, W. Weerkamp, and M. de Rijke, "Query and document models for enterprise search," in *Proceedings of the Text REtrieval Conference*, *(TREC '07)*, Gaithersburg, MD, 2008.

[28] K. Balog, P. Serdyukov, and A. P. de Vries, "Overview of the TREC 2010 entity track," in *Proceedings of the Text REtrieval Conference*, *(TREC '10)*, Gaithersburg, MD, February 2011.

[29] K. Balog, I. Soboroff, P. Thomas, N. Craswell, A. P. de Vries, and P. Bailey, "Overview of the TREC 2008 enterprise track," in *Proceedings of the Text REtrieval Conference*, *(TREC '08)*, Gaithersburg, MD, 2009.

[30] K. Balog, W. Weerkamp, and M. de Rijke, "A few examples go a long way: Constructing query models from elaborate query formulations," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, *(SIGIR '08)*, New York, NY, USA, pp. 371–378, 2008. ISBN 978-1-60558-164-4.

[31] B. Bansal, "Large scale computing @ linkedin," *Talk*, 2011. URL http://cacs.usc.edu/education/cs596/usc_bbansal.pdf.

[32] S. Bao, H. Duan, Q. Zhou, M. Xiong, Y. Cao, and Y. Yu, "Research on expert search at enterprise track of TREC 2006," in *Proceedings of the Text REtrieval Conference*, *(TREC '06)*, Gaithersburg, MD, 2007.

[33] P. Baumard, *Tacit knowledge in organizations*. Thousand Oaks, CA, USA: Sage Publications, 2001.

[34] I. Becerra-Fernandez, "The role of artificial intelligence technologies in the implementation of people-finder knowledge management systems," *Knowledge-Based Systems*, vol. 13, no. 5, pp. 315–320, 2000. ISSN 0950-7051.

[35] I. Becerra-Fernandez, "Searching for experts on the web: A review of contemporary expertise locator systems," *ACM Transactions on Internet Technology*, vol. 6, no. 4, pp. 333–355, 2006. ISSN 1533-5399.

[36] J. Bennett and S. Lanning, "The Netflix prize," in *Proceedings of KDD Cup and Workshop*, vol. 2007 of *(KDD '07)*, New York, NY, USA: ACM, p. 35, 2007. ISBN 978-1-59593-609-7.

[37] R. Berendsen, B. Kovachev, E. Nastou, M. de Rijke, and W. Weerkamp, "Result disambiguation in web people search," in *Proceedings of the European Conference on Information Retrieval, (ECIR '12)*, Berlin, Heidelberg, 2012.

[38] R. Berendsen, E. Tsagkias, M. de Rijke, and E. Meij, "Generating pseudo test collections for learning to rank scientific articles," in *CLEF 2012: Conference and Labs of the Evaluation Forum*, Rome, Italy, 2012.

[39] R. Blanco, H. Halpin, D. M. Herzig, P. Mika, J. Pound, H. S. Thompson, and T. T. Duc, "Entity search evaluation over structured web data," in *Proceedings of the International Workshop on Entity-Oriented Search, (EOS '11)*, pp. 65–71, 2011.

[40] D. Blei, A. Ng, and M. Jordan, "Latent Dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003. ISSN 1532-4435.

[41] S. P. Borgatti and R. Cross, "A relational view of information seeking and learning in social networks," *Management Science*, vol. 49, no. 4, pp. 432–445, 2003. ISSN 0025-1909.

[42] M. Bron, K. Balog, and M. de Rijke, "Ranking related entities: Components and analyses," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '10)*, New York, NY, USA, pp. 1079–1088, 2010. ISBN 978-1-4503-0099-5.

[43] M. Bron, B. Huurnink, and M. de Rijke, "Linking archives using document enrichment and term selection," in *Proceedings of the International Conference on Theory and Practice of Digital Libraries, (TPDL '11)*, Heidelberg, Germany, 2011. ISBN 978-3-642-24468-1.

[44] C. Burgess, K. Livesay, and K. Lund, "Explorations in context space: Words, sentences, discourse," *Discourse Processes*, vol. 25, pp. 211–257, 1998.

[45] J. Callan, "Distributed information retrieval," *Advances in Information Retrieval: Recent Research from the Center for Intelligent Information Retrieval*, vol. 7, pp. 127–150, 2002.

[46] C. S. Campbell, P. P. Maglio, A. Cozzi, and B. Dom, "Expertise identification using email communications," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '03)*, New York, NY, USA, pp. 528–531, 2003. ISBN 1-58113-723-0.

[47] Y. Cao, J. Liu, S. Bao, and H. Li, "Research on expert search at enterprise track of TREC 2005," in *Proceedings of the Text REtrieval Conference, (TREC '05)*, Gaithersburg, MD, 2006.

[48] S. Cetintas, M. Rogati, L. Si, and Y. Fang, "Identifying similar people in professional social networks with discriminative probabilistic models," in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '11)*, New York, NY, USA, pp. 1209–1210, 2011. ISBN 978-1-4503-0757-4.

[49] O. Chapelle and M. Wu, "Gradient descent optimization of smoothed information retrieval metrics," *Information retrieval*, vol. 13, no. 3, pp. 216–235, 2010. ISSN 1386-4564.

[50] L. Charlin, R. Zemel, and C. Boutilier, "A framework for optimizing paper matching," in *Proceedings of the Conference on Uncertainty in Artificial Intelligence, (UAI'11)*, pp. 86–95, 2011. ISBN 978-0-9749039-7-2.

[51] H. Chen, H. Shen, J. Xiong, S. Tan, and X. Cheng, "Social network structure behind the mailing lists: Ict-iiis at TREC 2006 expert finding track," in *Proceedings of the Text REtrieval Conference*, *(TREC '05)*, Gaithersburg, MD, 2006.

[52] D. Constant, L. Sproull, and S. Kiesler, "The kindness of strangers: The usefulness of electronic weak ties for technical advice," *Organization Science*, vol. 7, no. 2, pp. 119–135, 1996.

[53] W. Cooper, "Exploiting the maximum entropy principle to increase retrieval effectiveness," *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 34, no. 1, pp. 31–39, 1983. ISSN 0002-8231.

[54] N. Craswell, A. de Vries, and I. Soboroff, "Overview of the TREC-2005 enterprise track," in *Proceedings of the Text REtrieval Conference*, *(TREC '05)*, Gaithersburg, MD, 2006.

[55] N. Craswell and D. Hawking, "Overview of the TREC-2002 web track," in *Proceedings of the Text REtrieval Conference Proceedings*, *(TREC '01)*, Gaithersburg, MD, 2002.

[56] N. Craswell, D. Hawking, A. M. Vercoustre, and P. Wilkins, "P@noptic expert: Searching for experts not just for documents," in *Ausweb Poster Proceedings*, 2001.

[57] N. Craswell, D. Hawking, R. Wilkinson, and M. Wu, "Overview of the TREC 2003 web track," in *Proceedings of the Text REtrieval Conference Proceedings*, vol. 2003 *of TREC '02*, Gaithersburg, MD, pp. 78–92, 2003.

[58] N. Craswell, S. Robertson, H. Zaragoza, and M. Taylor, "Relevance weighting for query independent evidence," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, *(SIGIR '05)*, New York, NY, USA, pp. 416–423, 2005. ISBN 1-59593-034-5.

[59] A. Csomai and R. Mihalcea, "Linking documents to encyclopedic knowledge," *IEEE Intelligent Systems*, vol. 23, pp. 34–41, September 2008. ISSN 1541-1672.

[60] S. Cucerzan, "Large-scale named entity disambiguation based on wikipedia data," in *Proceedings of Empirical Methods in Natural Language Processing*, *(EMNLP '07)*, Stroudsburg, PA, USA: ACL, pp. 708–716, 2007.

[61] R. Cummins, M. Lalmas, and C. O'Riordan, "Learning aggregation functions for expert search," in *Proceedings of the European Conference on Artificial Intelligence*, *(ECAI '10)*, pp. 535–540, 2010. ISBN 978-1-60750-605-8.

[62] S. Das, P. Mitra, and C. Giles, "Learning to rank homepages for researcher-name queries," *The International Workshop on Entity-Oriented Search*, pp. 53–58, 2011.

[63] T. Davenport and L. Prusak, *Working Knowledge: How Organizations Manage What They Know*. Harvard Business Press, 2000. ISBN 0875846556.

[64] M. de Rijke, K. Balog, T. Bogers, and A. van den Bosch, "On the evaluation of entity profiles," in *Proceedings of the 2010 Conference on Multilingual and Multimodal Information Access Evaluation*, *(CLEF '10)*, pp. 94–99, September 2010. ISBN 978-3-642-15997-8.

[65] A. de Vries, A.-M. Vercoustre, J. A. Thom, N. Craswell, and M. Lalmas, "Overview of the INEX 2007 entity ranking track," in *Proceedings of the Initiative on the Evaluation of XML Retrieval, (INEX '07)*, Berlin, Heidelberg, pp. 245–251, 2008. ISBN 978-3-540-85901-7.

[66] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107–113, 2008. ISSN 0001-0782.

[67] G. Demartini, A. de Vries, T. Iofciu, and J. Zhu, "Overview of the INEX 2008 entity ranking track," in *Proceedings of the Initiative on the Evaluation of XML Retrieval, (INEX '08)*, Berlin, Heidelberg, pp. 243–252, 2009. ISBN 978-3-642-03760-3.

[68] G. Demartini, J. Gaugaz, and W. Nejdl, "A vector space model for ranking entities and its application to expert search," *Advances in Information Retrieval*, pp. 189–201, 2009.

[69] G. Demartini, T. Iofciu, and A. de Vries, "Overview of the INEX 2009 entity ranking track," in *Proceedings of the Initiative on the Evaluation of XML Retrieval*, vol. 6203 of *(INEX '09)*, (S. Geva, J. Kamps, and A. Trotman, eds.), Berlin, Heidelberg, pp. 254–264, 2010. ISBN 978-3-642-14555-1.

[70] G. Demartini and C. Niederée, "Finding experts on the semantic desktop," in *Proceedings of the International Semantic Web Conference, (ISWC '08)*, 2008. ISBN 978-3-540-88563-4.

[71] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1–38, 1977. ISSN 1467-9868.

[72] H. Deng, I. King, and M. R. Lyu, "Formal models for expert finding on DBLP bibliography data," in *Proceedings of the IEEE International Conference on Data Mining, (ICDM '08)*, Washington, DC, USA, pp. 163–172, 2008.

[73] H. Deng, I. King, and M. R. Lyu, "Enhancing expertise retrieval using community-aware strategies," in *Proceeding of the ACM International Conference on Information and Knowledge Management, (CIKM '09)*, New York, NY, USA, pp. 1733–1736, 2009. ISBN 978-1-60558-512-3.

[74] J. Dennis and R. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial Mathematics, 1996. ISBN 0898713641.

[75] A. Doan and A. Y. Halevy, "Semantic-integration research in the database community," *AI Magazine*, vol. 26, no. 1, pp. 83–94, 2005. ISSN 0738-4602.

[76] B. Dom, I. Eiron, A. Cozzi, and Y. Zhang, "Graph-based ranking algorithms for e-mail expertise analysis," in *Proceedings of the ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, (DMKD '03)*, New York, NY, USA, pp. 42–48, 2003.

[77] X. Dong, A. Halevy, and J. Madhavan, "Reference reconciliation in complex information spaces," in *Proceedings of the ACM SIGMOD International Conference on Managementof Data, (SIGMOD '05)*, New York, NY, USA, pp. 85–96, 2005. ISBN 1-59593-060-4.

[78] H. Duan, Q. Zhou, Z. Lu, O. Jin, S. Bao, and Y. C. Y. Yu, "Research on enterprise track of TREC 2007 at SJTU APEX lab," in *Proceedings of the Text REtrieval Conference, (TREC '06)*, Gaithersburg, MD, 2007.

[79] ECSCW'99 Workshop, *Beyond Knowledge Management: Managing Expertise*. 1999. URL http://www.informatik.uni-bonn.de/~prosec/ECSCW-XMWS/.

[80] K. Ehrlich, C. Lin, and V. Griffiths-Fisher, "Searching for experts in the enterprise: Combining text and social network analysis," in *Proceedings of the 2007 International ACM SIGGROUP Conference on Supporting Group Work, (GROUP '07)*, New York, NY, USA, pp. 117–126, 2007. ISBN 978-1-59593-845-9.

[81] K. Ehrlich and N. S. Shami, "Searching for expertise," in *Proceeding of the Annual SIGCHI Conference on Human Factors in Computing Systems, (CHI '08)*, New York, NY, USA, pp. 1093–1096, 2008. ISBN 978-1-60558-011-1.

[82] A. K. Elmagarmid, P. G. Ipeirotis, and V. S. Verykios, "Duplicate record detection: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, pp. 1–16, January 2007. ISSN 1041-4347.

[83] H. Fang and C. Zhai, "Probabilistic models for expert finding," in *Proceedings of the European Conference on IR Research, (ECIR'07)*, Berlin, Heidelberg, pp. 418–430, 2007. ISBN 978-3-540-71494-1.

[84] Y. Fang, L. Si, and A. Mathur, "Facfinder: Search for expertise in academic institutions," Technical Report: SERC-TR-294, Department of Computer Science, Purdue University, 2008.

[85] Y. Fang, L. Si, and A. Mathur, "Ranking experts with discriminative probabilistic models," in *SIGIR Workshop on Learning to Rank for Information Retrieval, (LR4IR '09)*, 2009.

[86] Y. Fang, L. Si, and A. Mathur, "Discriminative graphical models for faculty homepage discovery," *Information Retrieval*, vol. 13, no. 6, pp. 618–635, 2010. ISSN 1386-4564.

[87] Y. Fang, L. Si, and A. Mathur, "Discriminative probabilistic models for expert search in heterogeneous information sources," *Information Retrieval*, vol. 14, no. 2, pp. 158–177, 2011. ISSN 1386-4564.

[88] Y. Fang, L. Si, and A. P. Mathur, "Discriminative models of integrating document evidence and document-candidate associations for expert search," in *Proceeding of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '10)*, New York, NY, USA, pp. 683–690, 2010. ISBN 978-1-4503-0153-4.

[89] Y. Fang, N. Somasundaram, L. Si, J. Ko, and A. Mathur, "Analysis of an expert search query log," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '11)*, New York, NY, USA, pp. 1189–1190, 2011. ISBN 978-1-4503-0757-4.

[90] S. Fissaha Adafre and M. de Rijke, "Discovering missing links in wikipedia," in *Proceedings of the International Workshop on Link Discovery, (LinkKDD '05)*, New York, NY, USA, pp. 90–97, 2005. ISBN 1-59593-135-X.

[91] S. Fissaha Adafre, M. de Rijke, and E. T. K. Sang, "Entity retrieval," in *Recent Advances in Natural Language Processing, (RANLP '07)*, McGraw-Hill, September 2007. ISBN 978-90-272-4825-1.

[92] E. A. Fox and J. A. Shaw, "Combination of multiple searches," in *Text REtrieval Conference (TREC-2)*, NIST, pp. 243–252, 1994.

[93] Y. Fu, R. Xiang, Y. Liu, M. Zhang, and S. Ma, "Finding experts using social network analysis," in *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, (WI '07)*, Washington, DC, USA, pp. 77–80, 2007. ISBN 0-7695-3026-5.

[94] Z. Guan, G. Miao, R. McLoughlin, X. Yan, and D. Cai, "Co-occurrence based diffusion for expert search on the web," *IEEE Transactions on Knowledge and Data Engineering*, 2012. ISSN 1041-4347.

[95] H. Halpin, D. M. Herzig, P. Mika, R. Blanco, J. Pound, H. S. Thompson, and T. T. Duc, "Evaluating ad-hoc object retrieval," in *Proceedings of the International Workshop on Evaluation of Semantic Technologies, (IWEST '10)*, 2010.

[96] T. Heath, E. Motta, and M. Petre, "Person to person trust factors in word of mouth recommendation," in *Proceedings of the CHI2006 Workshop on Reinventing Trust, Collaboration, and Compliance in Social Systems, (Reinvent '06)*, 2006.

[97] R. Herbrich, T. Graepel, and K. Obermayer, "Large margin rank boundaries for ordinal regression," *Advances in Large Margin Classifiers*, vol. 88, no. 2, pp. 115–132, 2000.

[98] M. Hertzum, "People as carriers of experience and sources of commitment: Information seeking in a software design project," *New Review of Information Behaviour Research*, vol. 1, pp. 135–149, 2000. ISSN 1471-6313.

[99] M. Hertzum and A. M. Pejtersen, "The information-seeking practices of engineers: searching for documents as well as for people," *Information Processing and Management*, vol. 36, no. 5, pp. 761–778, 2000. ISSN 0306-4573.

[100] K. Hofmann, K. Balog, T. Bogers, and M. de Rijke, "Integrating contextual factors into topic-centric retrieval models for finding similar experts," in *Proceedings of the SIGIR Workshop on Future Challenges in Expertise Retrieval, (fCHER '08)*, 2008.

[101] K. Hofmann, K. Balog, T. Bogers, and M. de Rijke, "Contextual factors for finding similar experts," *Journal of the American Society for Information Science and Technology*, vol. 61, pp. 994–1014, May 2010. ISSN 0002-8231.

[102] T. Hofmann, "Probabilistic latent semantic indexing," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '99)*, New York, NY, USA, pp. 50–57, 1999. ISBN 1-58113-096-1.

[103] A. Hogan and A. Harth, "The expertfinder corpus 2007 for the benchmarking and development of expert-finding systems," in *Proceedings of the International ExpertFinder Workshop, (EFW '07)*, 2007.

[104] D. Horowitz and S. Kamvar, "The anatomy of a large-scale social search engine," in *Proceedings of the International World Wide Web Conference, (WWW '10)*, New York, NY, USA, pp. 431–440, 2010. ISBN 978-1-60558-799-8.

[105] J. Jiang, S. Han, and W. Lu, "Expertise retrieval using search engine results," in *Proceedings of the SIGIR Workshop on Future Challenges in Expertise Retrieval, (fCHER '08)*, ACM, 2008.

[106] J. Jiang, W. Liu, X. Rong, and Y. Gao, "Adapting language modeling methods for expert search to rank Wikipedia entities," in *Proceedings of the Initiative on the Evaluation of XML Retrieval, (INEX '08)*, Berlin, Heidelberg, pp. 264–272, 2009. ISBN 978-3-642-03760-3.

[107] J. Jiang, W. Lu, and D. Liu, "CSIR at TREC 2007 expert search task," in *Proceedings of the Text REtrieval Conference, (TREC '06)*, Gaithersburg, MD, 2007.

[108] V. Jijkoun, M. A. Khalid, M. Marx, and M. de Rijke, "Named entity normalization in user generated content," in *Proceedings of the workshop on Analytics for Noisy Unstructured Text Data, (AND '08)*, New York, NY, USA: ACM, pp. 23–30, 2008.

[109] R. Kaptein and J. Kamps, "Finding entities in Wikipedia using links and categories," in *Proceedings of the Initiative on the Evaluation of XML Retrieval, (INEX '08)*, Berlin, Heidelberg: Springer Verlag, pp. 273–279, 2009. ISBN 978-3-642-03760-3.

[110] R. Kaptein, P. Serdyukov, A. de Vries, and J. Kamps, "Entity ranking using Wikipedia as a pivot," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '10)*, New York, NY, USA, pp. 69–78, 2010. ISBN 978-1-4503-0099-5.

[111] M. Karimzadehgan, R. W. White, and M. Richardson, "Enhancing expert finding using organizational hierarchies," in *Proceedings of the European Conference on IR Research, (ECIR'09)*, Berlin, Heidelberg, pp. 177–188, 2009. ISBN 978-3-642-00957-0.

[112] M. Karimzadehgan and C. Zhai, "Constrained multi-aspect expertise matching for committee review assignment," in *Conference on Information and Knowledge Management (CIKM'09)*, pp. 1697–1700, 2009. ISBN 978-1-59593-991-3.

[113] M. Karimzadehgan, C. Zhai, and G. Belford, "Multi-aspect expertise matching for review assignment," in *Proceeding of the ACM International Conference on Information and Knowledge Management, (CIKM '08)*, New York, NY, USA, pp. 1113–1122, 2008. ISBN 978-1-59593-991-3.

[114] M. Khalid, V. Jijkoun, and M. de Rijke, "The impact of named entity normalization on information retrieval for question answering," in *Proceedings of the European Conference on IR Research, (ECIR '08)*, Berlin, Heidelberg, pp. 705–710, 2008. ISBN 978-3-540-78645-0.

[115] J. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM (JACM)*, vol. 46, no. 5, pp. 604–632, 1999. ISSN 0004-5411.

[116] J. Ko, E. Nyberg, and L. Si, "A probabilistic graphical model for joint answer ranking in question answering," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '07)*, New York, NY, USA, pp. 343–350, 2007. ISBN 978-1-59593-597-7.

[117] G. Kukla, P. Kazienko, P. Bródka, and T. Filipowski, "Soclake: social latent knowledge explorator," *The Computer Journal*, vol. 55, no. 3, pp. 258–276, 2012. ISSN 0010-4620.

[118] L. Kuncheva and C. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine learning*, vol. 51, no. 2, pp. 181–207, 2003.

[119] T. Lappas, K. Liu, and E. Terzi, "Finding a team of experts in social networks," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (KDD '09)*, New York, NY, USA, pp. 467–476, 2009. ISBN 978-1-60558-495-9.

[120] V. Lavrenko and W. B. Croft, "Relevance based language models," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '01)*, New York, NY, USA, pp. 120–127, 2001. ISBN 1-58113-331-6.

[121] J. H. Lee, "Combining multiple evidence from different properties of weighting schemes," in *Special Interest Group on Informations Retrieval SIGIR'95*, pp. 180–188, 1995.

[122] J.-Z. Li, J. Tang, J. Zhang, Q. Luo, Y. Liu, and M. Hong, "Eos: expertise oriented search using social networks," in *Proceedings of the International World Wide Web Conference, (WWW '07)*, New York, NY, USA, pp. 1271–1272, 2007. ISBN 978-1-59593-654-7.

[123] R. Liebregts and T. Bogers, "Design and implementation of a university-wide expert search engine," in *Proceedings of the European Conference on IR Research, (ECIR'09)*, Berlin, Heidelberg, pp. 587–594, April 2009. ISBN 978-3-642-00957-0.

[124] J. Liker and D. Meier, *Toyota Talent: Developing Your People the Toyota Way*. McGraw-Hill, 2007. ISBN 0071477454.

[125] B. Lin, K. D. Rosa, R. Shah, and N. Agarwal, "Lads: Rapid development of a learning-to-rank based related entity finding system using open advancement," in *Proceedings of the International Workshop on Entity-Oriented Search, (EOS '11)*, pp. 14–19, 2011.

[126] T. Liu, "Learning to rank for information retrieval," *Foundations and Trends in Information Retrieval*, vol. 3, no. 3, pp. 225–331, 2009. ISSN 1554-0669.

[127] X. Liu and W. B. Croft, "Cluster-based retrieval using language models," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '04)*, New York, NY, USA, pp. 186–193, 2004. ISBN 1-58113-881-4.

[128] X. Liu, W. B. Croft, and M. Koll, "Finding experts in community-based question-answering services," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '05)*, New York, NY, USA, pp. 315–316, 2005. ISBN 1-59593-140-6.

[129] C. Macdonald, "The voting model for people search," PhD thesis, University of Glasgow, Feburary 2009.

[130] C. Macdonald, D. Hannah, and I. Ounis, "High quality expertise evidence for expert search," in *Proceedings of the European Conference on IR Research, (ECIR '08)*, Berlin, Heidelberg, pp. 283–295, 2008. ISBN 978-3-540-78645-0.

[131] C. Macdonald and I. Ounis, "Searching for expertise using the Terrier platform," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '06)*, New York, NY, USA, pp. 732–732, 2006. ISBN 1-59593-369-7.

[132] C. Macdonald and I. Ounis, "Voting for candidates: Adapting data fusion techniques for an expert search task," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '06)*, New York, NY, USA, pp. 387–396, 2006. ISBN 1-59593-433-2.

[133] C. Macdonald and I. Ounis, "Expertise drift and query expansion in expert search," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '07)*, New York, NY, USA, pp. 341–350, 2007. ISBN 978-1-59593-803-9.

[134] C. Macdonald and I. Ounis, "Using relevance feedback in expert search," in *Proceedings of the European Conference on IR Research, (ECIR'07)*, Berlin, Heidelberg, pp. 431–443, 2007. ISBN 978-3-540-71494-1.

[135] C. Macdonald and I. Ounis, "Expert search evaluation by supporting documents," in *Proceedings of the European Conference on IR Research, (ECIR '08)*, Berlin, Heidelberg, pp. 555–563, 2008. ISBN 978-3-540-78645-0.

[136] C. Macdonald and I. Ounis, "Key blog distillation: Ranking aggregates," in *Proceeding of the ACM International Conference on Information and Knowledge Management, (CIKM '08)*, New York, NY, USA, pp. 1043–1052, 2008. ISBN 978-1-59593-991-3.

[137] C. Macdonald and I. Ounis, "The influence of the document ranking in expert search," in *Proceeding of the ACM International Conference on Information and Knowledge Management, (CIKM '09)*, New York, NY, USA, pp. 1983–1986, 2009. ISBN 978-1-60558-512-3.

[138] C. Macdonald and I. Ounis, "Searching for expertise: experiments with the voting model," *Computer Journal: Special Focus on Profiling Expertise and Behaviour*, vol. 52, no. 7, pp. 729–748, 2009. ISSN 0010-4620.

[139] C. Macdonald and I. Ounis, "The influence of the document ranking in expert search," *Information Processing and Management*, vol. 74, no. 3, pp. 376–390, 2010. ISSN 0306-4573.

[140] C. Macdonald and I. Ounis, "Learning models for ranking aggregates," in *Proceedings of the European Conference on IR Research, (ECIR '11)*, Berlin, Heidelberg, pp. 517–529, 2011. ISBN 978-3-642-20160-8.

[141] C. Macdonald, R. L. Santos, I. Ounis, and I. Soboroff, "Blog track research at TREC," *SIGIR Forum*, vol. 44, pp. 58–75, August 2010. ISSN 0163-5840.

[142] C. Macdonald and R. W. White, "Usefulness of click-through data in expert search," in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '09)*, New York, NY, USA, pp. 816–817, 2009. ISBN 978-1-60558-483-6.

[143] R. Maclin and D. Opitz, "Popular ensemble methods: An empirical study," *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, 1999.

[144] G. S. Mann and D. Yarowsky, "Unsupervised personal name disambiguation," in *Proceedings of the Conference on Natural Language Learning at HLT-NAACL*, Stroudsburg, PA, USA, pp. 33–40, 2003.

[145] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008. ISBN 0521865719.

[146] M. Maybury, "Expert finding systems," Technical Report: MTR-06B000040, 2006.

[147] D. W. McDonald, "Evaluating expertise recommendations," in *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, (GROUP '01)*, New York, NY, USA, pp. 214–223, 2001.

[148] D. W. McDonald and M. S. Ackerman, "Expertise recommender: a flexible recommendation system and architecture," in *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, (CSCW '00)*, New York, NY, USA, pp. 231–240, 2000.

[149] E. Meij, M. Bron, L. Hollink, B. Huurnink, and M. de Rijke, "Mapping queries to the linking open data cloud: A case study using dbpedia," *Journal of Web Semantics*, vol. 9, no. 4, pp. 418–433, 2011. ISSN 1570-8268.

[150] E. Meij, M. Bron, B. Huurnink, L. Hollink, and M. de Rijke, "Learning semantic query suggestions," in *Proceedings of the International Semantic Web Conference, (ISWC '09)*, pp. 415–430, 2009. ISBN 978-3-642-04929-3.

[151] E. Meij, W. Weerkamp, and M. de Rijke, "Adding semantics to microblog posts," in *Proceedings of the ACM International Conference on Web Search and Data Mining, (WSDM '12)*, New York, NY, USA, 2012.

[152] H. Menzel, *Review of Studies in the Flow of Information Among Scientists*. New York: Columbia University, Bureau of Applied Social Research, 1960.

[153] H. Menzel, "Information needs and uses in science and technology," *Annual Review of Information Science and Technology*, vol. 1, pp. 41–69, 1966.

[154] R. Mihalcea and A. Csomai, "Wikify!: Linking documents to encyclopedic knowledge," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '07)*, New York, NY, USA, pp. 233–241, 2007. ISBN 978-1-59593-803-9.

[155] D. Milne and I. H. Witten, "Learning to link with Wikipedia," in *Proceeding of the ACM International Conference on Information and Knowledge Management, (CIKM '08)*, New York, NY, USA, pp. 509–518, 2008. ISBN 978-1-59593-991-3.

[156] D. Mimno and A. McCallum, "Expertise modeling for matching papers with reviewers," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (KDD '07)*, New York, NY, USA, pp. 500–509, 2007. ISBN 978-1-59593-609-7.

[157] A. Mockus and J. D. Herbsleb, "Expertise browser: A quantitative approach to identifying expertise," in *Proceedings of the International Conference on Software Engineering, (ICSE '02)*, New York, NY, USA, pp. 503–512, 2002.

[158] C. Moreira, "Learning to rank academic experts," Master Thesis, Technical University of Lisbon, 2011.

[159] C. Moreira, P. Calado, and B. Martins, "Learning to rank for expert search in digital libraries of academic publications," *Progress in Artificial Intelligence*, pp. 431–445, 2011.

[160] M. R. Morris, J. Teevan, and K. Panovich, "A comparison of information seeking using search engines and social networks," in *Proceedings of the International Conference on Weblogs and Social Media, (ICWSM '10)*, pp. 291–294, 2010.

[161] R. Nallapati, "Discriminative models for information retrieval," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '04)*, New York, NY, USA, pp. 64–71, 2004. ISBN 1-58113-881-4.

[162] R. Neumayer, K. Balog, and K. Nørvåg, "On the modeling of entities for ad-hoc entity search in the web of data," in *European Conference on Information Retrieval (ECIR 2012)*, pp. 133–145, 2012.

[163] A. Ng and M. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes," in *Proceedings of the Advances in Neural Information Processing Systems, (NIPS '01)*, MIT Press, pp. 841–848, 2001. ISBN 0-262-02550-7.

[164] M. G. Noll, C. m. Au Yeung, N. Gibbins, C. Meinel, and N. Shadbolt, "Telling experts from spammers: expertise ranking in folksonomies," in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '09)*, New York, NY, USA, pp. 612–619, 2009. ISBN 978-1-60558-483-6.

[165] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," Technical Report 1999–66, 1999.

[166] A. Pal, S. Chang, and J. Konstan, "Evolution of experts in question answering communities," in *Proceedings of the International AAAI Conference on Weblogs and Social Media*, pp. 274–281, 2012.

[167] M. Peacock, "The search for expert knowledge continues," *CMSWire*, 2009. http://www.cmswire.com/cms/enterprise-cms/the-search-for-expert-knowledge-continues-004594.php.

[168] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988. ISBN 1558604790.

[169] J. Pehcevski, A.-M. Vercoustre, and J. A. Thom, "Exploiting locality of wikipedia links in entity ranking," in *Proceedings of the European Conference on IR Research, (ECIR '08)*, Glasgow, UK, pp. 258–269, 2008.

[170] D. Petkova and W. B. Croft, "Hierarchical language models for expert finding in enterprise corpora," in *Proceedings of the IEEE International Conference on Tools with Artificial Intelligence, (ICTAI '06)*, Washington, USA, pp. 599–608, 2006.

[171] D. Petkova and W. B. Croft, "Proximity-based document representation for named entity retrieval," in *Proceedings of the ACM International Conference on Information and Knowledge Management, (CIKM '07)*, New York, NY, USA, pp. 731–740, 2007. ISBN 978-1-59593-803-9.

[172] V. Plachouras, "Diversity in expert search," in *Proceedings of the Workshop on Diversity in Document Retrieval*, (C. Macdonald, J. Wang, and C. Clarke, eds.), pp. 63–67, 2011.

[173] J. Pound, P. Mika, and H. Zaragoza, "Ad-hoc object retrieval in the web of data," in *Proceedings of the International World Wide Web Conference, (WWW '10)*, New York, NY, USA, pp. 771–780, 2010. ISBN 978-1-60558-799-8.

[174] F. Riahi, Z. Zolaktaf, M. Shafiei, and E. Milios, "Finding expert users in community question answering," in *Proceedings of the International Conference on World Wide Web — Workshop on Community Question Answering on the Web, (CQA'12)*, pp. 791–798, 2012. ISBN 978-1-4503-1229-5.

[175] M. Richardson and R. W. White, "Supporting synchronous social q&a throughout the question lifecycle," in *Proceedings of the International World Wide Web Conference, (WWW '11)*, New York, NY, USA, pp. 755–764, 2011. ISBN 978-1-4503-0632-4.

[176] S. Robertson and K. Jones, "Relevance weighting of search terms," *Journal of the American Society for Information Science and Technology*, vol. 27, no. 3, pp. 129–146, 1976. ISSN 0002-8231.

[177] S. E. Robertson, "The probability ranking principle in IR," *Journal of Documentation*, vol. 33, no. 4, pp. 294–304, 1977. ISSN 0022-0418.

[178] J. Rocchio, *Relevance Feedback in Information Retrieval.* Englewood Cliffs, N.J.: Prentice Hall, 1971.

[179] M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth, "The author-topic model for authors and documents," in *Proceedings of the Conference on Uncertainty in Artificial Intelligence, (UAI '04)*, pp. 487–494, 2004. ISBN 0-9749039-0-6.

[180] V. Rosenberg, "Factors affecting the preferences of industrial personnel for information gathering methods," *Information Storage and Retrieval*, vol. 3, no. 3, pp. 119–127, 1967. ISSN 0020-0271.

[181] G. Salton, A. Wong, and C. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975. ISSN 0001-0782.

[182] M. Sanderson, "Test collection based evaluation of information retrieval systems," *Foundations and Trends in Information Retrieval*, vol. 4, no. 4, pp. 247–375, 2010. ISSN 1554-0669.

[183] R. L. T. Santos, C. Macdonald, and I. Ounis, "Voting for related entities," in *Adaptivity, Personalization and Fusion of Heterogeneous Information, (RIAO '10)*, Paris, France, pp. 1–8, 2010.

[184] R. L. T. Santos, C. Macdonald, and I. Ounis, "Mimicking web search engines for expert search," *Information Processing and Management*, vol. 47, no. 4, pp. 467–481, 2011. ISSN 0306-4573.

[185] T. Saracevic and P. B. Kantor, "A study of information seeking and retrieving. III. searchers, searches, overlap," *Journal of the American Society for Information Science and Technology*, vol. 39, pp. 197–216, 1988.

[186] M. F. Schwartz and D. C. M. Wood, "Discovering shared interests using graph analysis," *Communications of the ACM*, vol. 36, pp. 78–89, August 1993. ISSN 0001-0782.

[187] D. Seid and A. Kobsa, *Demoir: A Hybrid Architecture for Expertise Modeling and Recommender Systems.* 2000.

[188] P. Serdyukov, R. Aly, and H. Djoerd, "University of Twente at the TREC 2008 enterprise track: Using the global web as an expertise evidence source," in *Proceedings of the Text REtrieval Conference*, *(TREC '08)*, Gaithersburg, MD, 2009.

[189] P. Serdyukov, S. Chernov, and W. Nejdl, "Enhancing expert search through query modeling," in *Proceedings of the European Conference on IR Research*, *(ECIR'07)*, Berlin, Heidelberg, pp. 737–740, 2007. ISBN 978-3-540-71494-1.

[190] P. Serdyukov and D. Hiemstra, "Being omnipresent to be almighty: The importance of the global web evidence for organizational expert finding," in *Proceedings of the SIGIR Workshop on Future Challenges in Expertise Retrieval*, *(fCHER '08)*, pp. 17–24, 2008.

[191] P. Serdyukov and D. Hiemstra, "Modeling documents as mixtures of persons for expert finding," in *Proceedings of the European Conference on IR Research*, *(ECIR '08)*, Berlin, Heidelberg, pp. 309–320, 2008. ISBN 978-3-540-78645-0.

[192] P. Serdyukov, H. Rode, and D. Hiemstra, "Modeling multi-step relevance propagation for expert finding," in *Proceeding of the ACM International Conference on Information and Knowledge Management*, *(CIKM '08)*, New York, NY, USA, pp. 1133–1142, 2008. ISBN 978-1-59593-991-3.

[193] P. Serdyukov, M. Taylor, V. Vinay, M. Richardson, and R. W. White, "Automatic people tagging for expertise profiling in the enterprise," in *Proceedings of the European Conference on IR Research*, *(ECIR '11)*, Berlin, Heidelberg, pp. 399–410, 2011. ISBN 978-3-642-20160-8.

[194] S. N. Shami, K. Ehrlich, and D. R. Millen, "Pick me!: Link selection in expertise search results," in *Proceeding of the Annual SIGCHI Conference on Human Factors in Computing Systems*, *(CHI '08)*, New York, NY, USA, pp. 1089–1092, 2008. ISBN 978-1-60558-011-1.

[195] S. N. Shami, C. Y. Yuan, D. Cosley, L. Xia, and G. Gay, "That's what friends are for: facilitating 'who knows what' across group boundaries," in *Proceedings of the 2007 International ACM Conference on Supporting Group Work*, *(GROUP '07)*, New York, NY, USA, pp. 379–382, 2007. ISBN 978-1-59593-845-9.

[196] M. Shokouhi and L. Si, "Federated search," *Foundations and Trends in Information Retrieval*, vol. 5, no. 1, 2011. ISSN 1554-0669.

[197] E. Smirnova and K. Balog, "A user-oriented model for expert finding," in *Proceedings of the European Conference on IR Research*, *(ECIR '11)*, Berlin, Heidelberg, pp. 580–592, 2011. ISBN 978-3-642-20160-8.

[198] I. Soboroff, A. de Vries, and N. Crawell, "Overview of the TREC-2006 enterprise track," in *Proceedings of the Text REtrieval Conference*, *(TREC '06)*, Gaithersburg, MD, 2007.

[199] D. Song and P. D. Bruza, "Towards context sensitive information inference," *Journal of the American Society for Information Science and Technology*, vol. 54, pp. 321–334, February 2003.

[200] Y. Song, J. Huang, I. C. Councill, J. Li, and C. L. Giles, "Efficient topic-based unsupervised name disambiguation," in *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, *(JCDL '07)*, New York, NY, USA, pp. 342–351, 2007. ISBN 978-1-59593-644-8.

[201] P. Sorg and P. Cimiano, "Finding the right expert: Discriminative models for expert retrieval," in *Proceedings of the International Conference on Knowledge Discovery and Information Retrieval, (KDIR '11)*, 2011.

[202] P. Sorg, P. Cimiano, A. Schultz, and S. Sizov, "Overview of the cross-lingual expert search (cries) pilot challenge," in *Workshop of Cross-Language Evaluation Forum, (CLEF '10)*, September 2010. ISBN 978-88-904810-0-0.

[203] H. Su, J. Tang, and W. Hong, "Learning to diversify expert finding with subtopics," in *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, (PAKDD'12)*, 2012. ISBN 978-3-642-30216-9.

[204] C. Sutton and A. McCallum, "An introduction to conditional random fields for relational learning," *Introduction to Statistical Relational Learning*, p. 93, 2007.

[205] J. Tang, J. Sun, C. Wang, and Z. Yang, "Social influence analysis in large-scale networks," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (KDD '09)*, New York, NY, USA, pp. 807–816, 2009. ISBN 978-1-60558-495-9.

[206] J. Tang, J. Zhang, R. Jin, Z. Yang, K. Cai, L. Zhang, and Z. Su, "Topic level expertise search over heterogeneous networks," *Machine Learning Journal*, vol. 82, no. 2, pp. 211–237, 2011.

[207] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su, "Arnetminer: Extraction and mining of academic social networks," in *Proceeding of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (KDD '08)*, New York, NY, USA, pp. 990–998, 2008. ISBN 978-1-60558-193-4.

[208] W. Tang, J. Tang, T. Lei, C. Tan, B. Gao, and T. Li, "On optimization of expertise matching with various constraints," *Neurocomputing*, vol. 76, no. 1, pp. 71–83, 2012.

[209] D. Tax, M. V. Breukelen, R. Duin, and J. Kittler, "Combining multiple classifiers by averaging or by multiplying?," *Pattern Recognition*, vol. 33, no. 9, pp. 1475–1485, 2000. ISSN 0031-3203.

[210] L. Terveen and D. W. Mcdonald, "Social matching: A framework and research agenda," *ACM Transactions on Computer-Human Interaction*, vol. 12, no. 3, pp. 401–434, 2005. ISSN 1073-0516.

[211] T. Tsikrika, P. Serdyukov, H. Rode, T. Westerveld, R. Aly, D. Hiemstra, and A. de Vries, "Structured document retrieval, multimedia retrieval, and entity ranking using PF/Tijah," in *Proceedings of the Initiative on the Evaluation of XML Retrieval*, vol. 4862 *of (INEX '07)*, Berlin, Heidelberg: Springer Verlag, pp. 306–320, March 2008. ISBN 978-3-540-85901-7.

[212] Y. Tu, N. Johri, D. Roth, and J. Hockenmaier, "Citation author topic model in expert search," in *Proceedings of the International Conference on Computational Linguistics: Posters, (COLING '10)*, pp. 1265–1273, 2010.

[213] H. Turtle and W. B. Croft, "Evaluation of an inference network-based retrieval model," *ACM Transactions on Information Systems*, vol. 9, pp. 187–222, 1991.

[214] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer Verlag, 2000. ISBN 0387987800.

[215] W. Weerkamp, K. Balog, and M. de Rijke, "Blog feed search with a post index," *Information Retrieval*, vol. 14, pp. 515–545, 2011. ISSN 1386-4564.

[216] W. Weerkamp and M. de Rijke, "Credibility improves topical blog post retrieval," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, *(ACL'08: HLT)*, Stroudsburg, PA, USA, pp. 923–931, June 2008. ISBN 978-1-932432-04-6.

[217] J. Weng, E.-P. Lim, J. Jiang, and Q. He, "Twitterrank: finding topic-sensitive influential twitterers," in *Proceedings of the ACM International Conference on Web Search and Data Mining*, *(WSDM '10)*, New York, NY, USA, pp. 261–270, 2010. ISBN 978-1-60558-889-6.

[218] K. M. Wiig, "Knowledge management: Where did it come from and where will it go?," *Expert Systems with Applications*, vol. 13, pp. 1–14, July 1997.

[219] W. E. Winkler, "The state of record linkage and current research problems," Technical Report, Statistical Research Division, U.S. Census Bureau, 1999.

[220] 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, pp. 1267–1278, 2008. ISSN 0306-4573.

[221] Y. Xue, "Thuir at TREC2008: enterprise track," in *Proceedings of the Text REtrieval Conference*, *(TREC '07)*, Gaithersburg, MD, 2008.

[222] S. Yahyaei and C. Monz, "Applying maximum entropy to known-item email retrieval," in *Proceedings of the European Conference on IR Research*, *(ECIR '08)*, Berlin, Heidelberg, pp. 406–413, 2008. ISBN 978-3-540-78645-0.

[223] Z. Yang, J. Tang, B. Wang, J. Guo, J. Li, and S. Chen, "Expert2bole: From expert finding to bole search," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *(KDD '09)*, New York, NY, USA, pp. 1–4, 2009. ISBN 978-1-60558-495-9.

[224] S. Yarosh, T. Matthews, and M. Zhou, "Asking the right person: supporting expertise selection in the enterprise," in *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems*, *(CHI'12)*, pp. 2247–2256, 2012.

[225] D. Yimam-Seid and A. Kobsa, "Expert finding systems for organizations: Problem and domain analysis and the demoir approach," *Journal of Organizational Computing and Electronic Commerce*, vol. 13, no. 1, pp. 1–24, 2003. ISSN 1091-9392.

[226] E. Yom-Tov, S. Fine, D. Carmel, and A. Darlow, "Learning to estimate query difficulty: Including applications to missing content detection and distributed information retrieval," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, *(SIGIR '05)*, New York, NY, USA, pp. 512–519, 2005. ISBN 1-59593-034-5.

[227] C. Zhai, "Statistical language models for information retrieval a critical review," *Foundations and Trends in Information Retrieval*, vol. 2, pp. 137–213, March 2008. ISSN 1554-0669.

[228] C. Zhai and J. Lafferty, "Model-based feedback in the language modeling approach to information retrieval," in *Proceedings of the ACM International Conference on Information and Knowledge Management*, *(CIKM '01)*, New York, NY, USA, pp. 403–410, 2001. ISBN 1-58113-436-3.

[229] C. Zhai and J. Lafferty, "A study of smoothing methods for language models applied to information retrieval," *ACM Transactions on Information Systems*, vol. 22, pp. 179–214, April 2004.

[230] J. Zhang and M. S. Ackerman, "Searching for expertise in social networks: a simulation of potential strategies," in *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, *(GROUP '05)*, New York, NY, USA, pp. 71–80, 2005. ISBN 1-59593-223-2.

[231] J. Zhang, M. S. Ackerman, and L. Adamic, "Expertise networks in online communities: Structure and algorithms," in *Proceedings of the International World Wide Web Conference*, *(WWW '07)*, New York, NY, USA, pp. 221–230, 2007. ISBN 978-1-59593-654-7.

[232] J. Zhang, J. Tang, and J. Li, "Expert finding in a social network," *Advances in Databases: Concepts, Systems and Applications*, pp. 1066–1069, 2007.

[233] J. Zhang, J. Tang, L. Liu, and J. Li, "A mixture model for expert finding," in *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*, *(PAKDD '08)*, pp. 466–478, 2008. ISBN 978-3-540-68124-3.

[234] M. Zhang, R. Song, C. Lin, S. Ma, Z. Jiang, Y. Jin, Y. Liu, and L. Zhao, "Expansion-based technologies in finding relevant and new information: THU TREC2002 novelty track experiments," in *Proceedings of the Text Retrieval Conference (TREC)*, pp. 586–590, 2002.

[235] J. Zhu, "A study of the relationship between *ad-hoc* retrieval and expert finding in enterprise environment," in *Proceeding of the ACM Workshop on Web Information and Data Management*, *(WIDM '08)*, New York, NY, USA, pp. 25–30, 2008.

[236] J. Zhu, X. Huang, D. Song, and S. Ruger, "Integrating multiple document features in language models for expert finding," *Knowledge and Information Systems*, vol. 23, no. 1, pp. 29–54, 2010. ISSN 0219-1377.

[237] J. Zhu, D. Song, and S. Rüger, "Integrating document features for entity ranking," in *Proceedings of the Initiative on the Evaluation of XML Retrieval*, *(INEX '07)*, Berlin, Heidelberg, pp. 336–347, 2008. ISBN 978-3-540-85901-7.

[238] J. Zhu, D. Song, S. Rüger, and X. Huang, "Modeling document features for expert finding," in *Proceeding of the ACM International Conference on Information and Knowledge Management*, *(CIKM '08)*, New York, NY, USA, pp. 1421–1422, 2008. ISBN 978-1-59593-991-3.

[239] J. Zhu, D. Song, and S. M. Rüger, "Integrating multiple windows and document features for expert finding," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 4, pp. 694–715, 2009. ISSN 0002-8231.

[240] J. Zhu, D. Song, S. M. Rüger, M. Eisenstadt, and E. Motta, "The open university at trec 2006 enterprise track expert search task," in *Proceedings of the Text REtrieval Conference*, *(TREC '06)*, Gaithersburg, MD, 2007.