















## 6 Conclusions and Future Work

In this paper, we addressed the problem of generating training data for LTR processes in an expert seeking application. The implemented framework can be used to semi-automatically generate training data through a hierarchical system of rules. The evaluation shows that the single most important factor for the quality of the ranking function is the number of training pairs that are used to learn the ranking function. Future work that can be addressed is the implementation of online-learning, where the ranking function is updated through user feedback from previous searches. Our future work also includes using our software with other company knowledge bases, and testing and evaluating the framework with respect to standard information retrieval evaluation methods.

**Acknowledgments.** This work has been partly supported by the FP7 EU project SMART VORTEX.

## References

1. Agresti, A.: Analysis of ordinal categorical data. Wiley series in probability and mathematical statistics: Applied probability and statistics, Wiley (1984)
2. Albertoni, R., De Martino, M.: Semantic similarity of ontology instances tailored on the application context. In: OTM Conferences. pp. 1020–1038. LNCS Vol. 4275, Springer (2006)
3. Beierle, F.: Using a System of Rules to Generate Training Data for Learning-to-Rank Processes in an Expert Seeking Application. Master’s thesis, University of Hagen (2014)
4. Beierle, F., Engel, F., Hemmje, M.: Generation of training data for learning-to-rank processes in an expert seeking application. In: Informatiktage 2014 - Fachwissenschaftlicher Informatik-Kongress. Lecture Notes in Informatics (LNI), vol. S-13, pp. 97–100. Köllen Druck+Verlag (2014)
5. Dali, L., Fortuna, B., Tran, T., Mladenović, D.: Query-independent learning to rank for RDF entity search. The Semantic Web pp. 484–498 (2012)
6. Engel, F., Juchmes, M., Hemmje, M.: Expert search in semantic annotated enterprise data: integrating query- dependent and independent relevance factors. In: LWA 2013 - Lernen, Wissen & Adaptivität. Workshop Proceedings. pp. 41–44. Bamberg (2013)
7. Hofmann, K., Balog, K., Bogers, T., Rijke, M.d.: Contextual factors for finding similar experts. *Journal of the American Society for Information Science & Technology* 61(5), 994–1014 (2010)
8. Joachims, T.: Optimizing search engines using clickthrough data. In: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 133–142. KDD '02, ACM, New York, NY, USA (2002)
9. Macdonald, C., Santos, R.L.T., Ounis, I.: The whens and hows of learning to rank for web search. *Inf. Retr.* 16(5), 584–628 (2013)
10. Nevo, D., Benbasat, I., Wand, Y.: The knowledge demands of expertise seekers in two different contexts: Knowledge allocation versus knowledge retrieval. *Decis. Support Syst.* 53(3), 482–489 (2012)
11. Woudstra, L., van den Hooff, B., Schouten, A.P.: Dimensions of quality and accessibility: Selection of human information sources from a social capital perspective. *Information Processing & Management* 48(4), 618–630 (2012)