



















ferent embedding spaces. The different variants of ESR also perform rather differently on different queries. This work mainly focuses on how to make use of each edge type individually. An important future research direction is to study how to better utilize knowledge graph semantics in a more unified way, perhaps with the help of ranking labels.

This paper presents a new method of using knowledge graphs to improve the ranking of academic search. Although the details of this work are specific to Semantic Scholar, the techniques and lessons learned are general and can be applied to other full-text search engines.

## 8. ACKNOWLEDGMENTS

This research was supported by National Science Foundation (NSF) grant IIS-1422676 and a gift from the Allen Institute for Artificial Intelligence. Any opinions, findings, and conclusions in this paper are the authors' and do not necessarily reflect those of the sponsors.

## 9. REFERENCES

- [1] A. Arnold and W. W. Cohen. Information extraction as link prediction: Using curated citation networks to improve gene detection. In *International Conference on Wireless Algorithms, Systems, and Applications*, pages 541–550. Springer, 2009.
- [2] K. Balog, L. Azzopardi, and M. De Rijke. Formal models for expert finding in enterprise corpora. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR 2006)*, pages 43–50. ACM, 2006.
- [3] A. Berger and J. Lafferty. Information retrieval as statistical translation. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 1999)*, pages 222–229. ACM, 1999.
- [4] C. Caragea, F. A. Bulgarov, A. Godea, and S. Das Gollapalli. Citation-enhanced keyphrase extraction from research papers: A supervised approach. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1435–1446. Association for Computational Linguistics, 2014.
- [5] D. Carmel, M.-W. Chang, E. Gabrilovich, B.-J. P. Hsu, and K. Wang. ERD'14: Entity recognition and disambiguation challenge. In *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2014)*. ACM, 2014.
- [6] J. Chen, C. Xiong, and J. Callan. An empirical study of learning to rank for entity search. In *Proceedings of the 39th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2016)*. ACM, 2016. To Appear.
- [7] N. Craswell, A. P. de Vries, and I. Soboroff. Overview of the trec 2005 enterprise track. In *Proceedings of The 14th Text Retrieval Conference (TREC 2005)*, volume 5, pages 199–205, 2005.
- [8] J. Dalton, L. Dietz, and J. Allan. Entity query feature expansion using knowledge base links. In *Proceedings of the 37th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2014)*, pages 365–374. ACM, 2014.
- [9] L. Dietz, A. Kotov, and E. Meij. Utilizing knowledge bases in text-centric information retrieval. In *Proceedings of the 2016 ACM on International Conference on the Theory of Information Retrieval (ICTIR 2016)*, pages 5–5. ACM, 2016.
- [10] E. Gabrilovich, M. Ringgaard, and A. Subramanya. FACCI: Freebase annotation of ClueWeb corpora, Version 1 (Release date 2013-06-26, Format version 1, Correction level 0), June 2013.
- [11] J. Guo, Y. Fan, A. Qingyao, and W. Croft. A deep relevance matching model for ad-hoc retrieval. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (CIKM 2016)*, page To Appear. ACM, 2016.
- [12] F. Hasibi, K. Balog, and S. E. Bratsberg. Entity linking in queries: Tasks and evaluation. In *Proceedings of the Fifth ACM International Conference on The Theory of Information Retrieval (ICTIR 2015)*, pages 171–180. ACM, 2015.
- [13] P.-S. Huang, X. He, J. Gao, L. Deng, A. Acero, and L. Heck. Learning deep structured semantic models for web search using clickthrough data. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management (CIKM 2013)*, pages 2333–2338. ACM, 2013.
- [14] T. Joachims. Optimizing search engines using clickthrough data. In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2002)*, pages 133–142. ACM, 2002.
- [15] T.-Y. Liu. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, 3(3):225–331, 2009.
- [16] X. Liu and H. Fang. Latent entity space: a novel retrieval approach for entity-bearing queries. *Information Retrieval Journal*, 18(6):473–503, 2015.
- [17] X. Liu, P. Yang, and H. Fang. Entity came to rescue - Leveraging entities to minimize risks in web search. In *Proceedings of The 23st Text Retrieval Conference (TREC 2014)*. NIST, 2014.
- [18] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 27th Advances in Neural Information Processing Systems 2013 (NIPS 2013)*, pages 3111–3119, 2013.
- [19] H. Raviv, O. Kurland, and D. Carmel. Document retrieval using entity-based language models. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR 2016)*, pages 65–74. ACM, 2016.
- [20] A. Sinha, Z. Shen, Y. Song, H. Ma, D. Eide, B.-j. P. Hsu, and K. Wang. An overview of microsoft academic service (MAS) and applications. In *Proceedings of the 24th International Conference on World Wide Web (WWW 2015)*, pages 243–246. ACM, 2015.
- [21] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. Arnetminer: extraction and mining of academic social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD 2008)*, pages 990–998. ACM, 2008.
- [22] I. Vulić and M.-F. Moens. Monolingual and cross-lingual information retrieval models based on (bilingual) word embeddings. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2015)*, pages 363–372. ACM, 2015.
- [23] A. D. Wade, K. Wang, Y. Sun, and A. Gulli. Wsdm cup 2016: Entity ranking challenge. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining (WSDM 2016)*, pages 593–594. ACM, 2016.
- [24] X. Wei and W. B. Croft. Lda-based document models for ad-hoc retrieval. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR 2006)*, pages 178–185. ACM, 2006.
- [25] C. Xiong and J. Callan. EsdRank: Connecting query and documents through external semi-structured data. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM 2015)*, pages 951–960. ACM, 2015.
- [26] C. Xiong and J. Callan. Query expansion with Freebase. In *Proceedings of the fifth ACM International Conference on the Theory of Information Retrieval (ICTIR 2015)*, pages 111–120. ACM, 2015.
- [27] C. Xiong, J. Callan, and T.-Y. Liu. Bag-of-entity representation for ranking. In *Proceedings of the sixth ACM International Conference on the Theory of Information Retrieval (ICTIR 2016)*, pages 181–184. ACM, 2016.
- [28] Y. Xu, G. J. Jones, and B. Wang. Query dependent pseudo-relevance feedback based on Wikipedia. In *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009)*, pages 59–66. ACM, 2009.
- [29] J. Zhang, J. Tang, and J. Li. Expert finding in a social network. In *International Conference on Database Systems for Advanced Applications*, pages 1066–1069. Springer, 2007.