

# Supporting Search Engines with Knowledge and Context

Nikos Voskarides  
Amazon  
*nvvoskar@amazon.com*

## Abstract

Search engines leverage knowledge to improve information access. Such knowledge comes in different forms: unstructured knowledge (e.g., textual documents) and structured knowledge (e.g., relationships between real-world objects and topics). In order to effectively leverage knowledge, search engines should account for context, i.e., information about the user and query. In this thesis, we aim to support search engines in leveraging knowledge while accounting for context.

In the first part of this thesis, we study how to make structured knowledge more accessible to the user when the search engine proactively provides such knowledge as context to enrich search results. As a first task, we study how to retrieve descriptions of knowledge facts from a text corpus. Next, we study how to automatically generate knowledge fact descriptions. And finally, we study how to contextualize knowledge facts, that is, to automatically find facts related to a query fact.

In the second part of this thesis, we study how to improve interactive knowledge gathering. We focus on conversational search, where the user interacts with the search engine to gather knowledge over large unstructured knowledge repositories. We study multi-turn passage retrieval as an instance of conversational search and focus on query resolution, that is, add missing context from the conversation history to the current turn. We model query resolution as a term classification task and propose a method to address it.

In the final part of this thesis, we focus on search engine support for professional writers in the news domain. We study how to support such writers create event-narratives by exploring knowledge from a corpus of news articles. We propose a dataset construction procedure for this task that relies on existing news articles to simulate incomplete narratives and relevant articles. We study the performance of multiple rankers, lexical and semantic, and provide insights into the characteristics of this task.

**Awarded by:** University of Amsterdam, Amsterdam, The Netherlands **on** 5 February 2021.

**Supervised by:** Maarten de Rijke.

**Available at:** <https://hdl.handle.net/11245.1/78187b29-2403-4711-800a-0f92fcb9b15c>.

---

## Selected Publications

- Nikos Voskarides, Edgar Meij, Manos Tsagkias, Maarten de Rijke, and Wouter Weerkamp. Learning to explain entity relationships in knowledge graphs. In *ACL-IJCNLP 2015*, pages 564–574. ACL, July 2015.
- Nikos Voskarides, Edgar Meij, and Maarten de Rijke. Generating descriptions of entity relationships. In *ECIR 2017: 39th European Conference on Information Retrieval*, LNCS. Springer, April 2017.
- Nikos Voskarides, Edgar Meij, Ridho Reinanda, Abhinav Khaitan, Miles Osborne, Giorgio Stefanoni, Kambadur Prabhanjan, and Maarten de Rijke. Weakly-supervised contextualization of knowledge graph facts. In *SIGIR 2018: 41st international ACM SIGIR conference on Research and Development in Information Retrieval*, pages 765–774. ACM, July 2018.
- Nikos Voskarides, Dan Li, Pengjie Ren, Evangelos Kanoulas, and Maarten de Rijke. Query resolution for conversational search with limited supervision. In *SIGIR 2020: 43rd international ACM SIGIR conference on Research and Development in Information Retrieval*, pages 921–932. ACM, July 2020.
- Nikos Voskarides, Edgar Meij, Sabrina Sauer, and Maarten de Rijke. News article retrieval in context for event-centric narrative creation. In *ICTIR 2021: ACM SIGIR International Conference on the Theory of Information Retrieval*, pages 103–112. ACM, July 2021.