Task-Based Support in Search Engines

Darío Garigliotti

Abstract

Web search has become a key technology on which people rely daily for getting information about almost everything. The evolution of the search experience has also shaped the expectations of people about it. Many users seem to expect today's web search engines to behave like a kind of "wise interpreter," capable of understanding the meaning behind a search query, realizing its current context, and responding to it directly and appropriately. Search by meaning, or semantic search, encompasses a large portion of information retrieval (IR) research devoted to study more meaningful representations of the information need expressed by the user query. Entity cards, direct displays, and verticals are examples of how major commercial search engines have indeed responded to user expectations, capitalizing on query understanding. Search is usually performed with a specific goal underlying the query. In many cases, this goal consists of a nontrivial task to be completed. Current search engines support a small set of basic tasks, and most of the knowledge-intensive workload for supporting more complex tasks is left to the user.

Task-based search can be viewed as an information access paradigm that aims to enhance search engines with functionalities for recognizing the underlying tasks in searches and providing support for task completion. The research presented in this thesis focuses on utilizing and extending methods and techniques from semantic search in the next stage of the evolution of search engines, namely, to support users in achieving their tasks. Our work can be grouped in three grand themes: (1) *Entity type information for entity retrieval:* we conduct a systematic evaluation and analysis of methods for type-aware entity retrieval, in terms of three main dimensions. Also, we revisit the problem of hierarchical target type identification, present a state-of-the-art supervised learning method, and analyze the usage of automatically identified target entity types for type-aware entity retrieval; (2) *Entity-oriented search intents:* we propose a categorization scheme for entity-oriented search intents, and study the distributions of entity intent categories per entity type. We further develop a method for constructing a knowledge base of entity-oriented search intents; and (3) *Task-based search:* we design a probabilistic generative framework for task-based query suggestion, and principledly estimate each of its components. Furthermore, we introduce the problems of query-based task recommendation and mission-based task recommendation, and establish respective methods as suitable baselines.

Selected Publications

Darío Garigliotti, Faegheh Hasibi, and Krisztian Balog. Identifying and exploiting target entity type information for ad hoc entity retrieval. *Information Retrieval Journal*, 22(3-4):285–323, 2019.

Darío Garigliotti and Krisztian Balog. IntentsKB: A Knowledge Base of Entity-Oriented Search Intents. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM '18)*, pages 1659–1662, 2018a.

Darío Garigliotti and Krisztian Balog. Towards an understanding of entity-oriented search intents. In *European Conference on Information Retrieval (ECIR '18)*, pages 644–650. Springer, 2018b.

Darío Garigliotti and Krisztian Balog. On Type-Aware Entity Retrieval. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR '17)*, pages 27–34, 2017a.

Darío Garigliotti and Krisztian Balog. Generating Query Suggestions to Support Task-Based Search. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17)*, pages 1153–1156, 2017b.