Abstract

The Web is the largest repository of knowledge in the world. Everyday people contribute to make it bigger by generating new web data. Data never sleeps. Every minute someone writes a new blog post, uploads a video or comments on an article. Usually people rely on Web Search Engines for satisfying their *information needs*: they formulate their needs as text queries and they expect a list of highly relevant documents answering their requests. Being able to manage this massive volume of data, ensuring high quality and performance, is a challenging topic that we tackle in this thesis.

In this dissertation we focus on the *Web of Data*: a recent approach, originated from the Semantic Web community, consisting in a collective effort to augment the existing Web with semistructured-data. We propose to manage the *data explosion* shifting from a retrieval model based on documents to a model enriched with *entities*, where an entity can describe a person, a product, a location, a company, through *semi-structured information*.

In our work, we combine the Web of Data with an important source of knowledge: *query logs*, which record the interactions between the Web Search Engine and the users. Query log mining aims at extracting valuable knowledge that can be exploited to enhance users' search experience. According to this vision, this dissertation aims at improving Web Search Engines toward the mutual use of query logs and entities.

The contributions of this work are the following: we show how historical usage data can be exploited for improving performance during the snippet generation process. Secondly, we propose a query recommender system that, by combining entities with queries, leads to significant improvements to the quality of the suggestions. Furthermore, we develop a new technique for estimating the *relatedness* between two entities, i.e., their semantic similarity. Finally, we show that entities may be useful for automatically building explanatory statements that aim at helping the user to better understand if, and why, the suggested item can be of her interest.