



*Entity Associations for Search.*

R. Reinanda

# Summary

In this thesis, we investigate the broad problem of computing entity associations for search. Specifically, we consider three types of entity association: *entity-entity*, *entity-document*, and *entity-aspect* associations. We touch upon various domains, starting with specific domains such as the humanities and business, and ending in Web search.

The first type of association is *entity-entity* association. We begin our investigation by considering entity networks as a means of exploring document collections, and formulate the task of entity network extraction as ranking related entities. We combine approaches based on association finding and relation extraction to address the task. In our second study we go in a different direction, and focus on establishing temporal boundaries between pairs of entities having confirmed the relations between them. Finally, we bring the type of the relations to the forefront in our third study. We consider the task of recommending entities related to a query entity given their direct and indirect connections in a knowledge graph. We formalize the task of impact-based entity recommendation and propose two approaches based on learning to rank and impact propagation.

Our second theme concerns *entity-document* associations. We study this type of association in the context of filtering documents for knowledge base acceleration. In this setting, the goal is to filter documents that are relevant to update a profile of an entity. We focus on the challenge of long-tail entities specifically, and propose an approach that leverages intrinsic, i.e., *in-document* signals more than extrinsic signals such as Wikipedia page views and trending queries.

Finally, we explore *entity-aspect* associations. Entities are often associated with attributes, types, distinguishing features, topics, or themes. We broadly group this type of information under the heading “aspect.” We study entity aspects in the context of Web search, and define them as common search tasks in the context of an entity. Specifically, we study the problem of mining, ranking, and recommending aspects. Entity aspects and their associations are mined from query logs.

This thesis contributes new task formalizations, algorithms, and insights on computing entity associations for search. Our experimental results confirm the effectiveness of our approaches within the different settings that we consider. Insights gained from this thesis will help address entity-oriented information access challenges in various domains.