

# What People are Asking About You: Mining Entity Search Intents in CQA Sites

David Carmel  
Yahoo!  
Haifa, Israel  
dcarmel@yahoo-inc.com

## Abstract

In this work we propose a novel representation for named entities that is based on the questions people ask about them in a CQA site. The representation is composed of entity related questions, answered by community members, which depict a meaningful search intent about the entity and are referred to as Entity Search Intents (ESI).

Based on the hypothesis that people ask similar questions about strongly related entities, we utilize the ESI representation for the task of entity relatedness estimation. Specifically, we estimate the relatedness between two entities based on the similarity between their associated search intents. The performance is evaluated by measuring the correlation of our proposed approach with human relatedness judgment over a dataset of entity pairs. Our method has been shown to be highly effective for this task, as high correlation was obtained. In addition, we show that combining ESI-based relatedness measurement with other entity similarity measurements based on word embedding significantly improves the relatedness measurement accuracy.

This is joint work with Hadas Raviv (Technion) and Idan Szpektor (Google)

## Bio

David is a Principal Research Scientist at Yahoo Research, Haifa, and an ACM Distinguished Engineer. David's research is focused on search and content quality analysis in Web and Email, query performance prediction, entity search, and text mining. David has published more than 100 papers in IR and Web journals and conferences, and serves on the editorial board of the IR journal and as a senior PC member or as Area Chair of many ACM conferences (SIGIR, WWW, WSDM, CIKM). He organized a number of workshops and taught several tutorials at SIGIR, and WWW.

David is co-author of the book "Estimating the Query Difficulty for Information Retrieval", published by Morgan & Claypool in 2010, and the co-author of the paper "Learning to estimate query difficulty", which won the Best Paper Award at SIGIR 2005. David earned his Ph.D. in Computer Science from the Technion, Israel Institute of Technology in 1997.

---

*Copyright © by the paper's authors. Copying permitted for private and academic purposes.*

In: L. Dietz, C. Xiong, E. Meij (eds.): Proceedings of the First Workshop on Knowledge Graphs and Semantics for Text Retrieval and Analysis (KG4IR), Tokyo, Japan, 11-Aug-2017, published at <http://ceur-ws.org>