

# Evaluating Retrieval Models using Retrievability Measurement

Shariq Bashir  
Vienna University of Technology  
*bashir@ifs.tuwien.ac.at*

March 26, 2012

## Abstract

Evaluation is the main driving force in research, development and applications related to information retrieval (IR). In the traditional IR evaluation paradigm a list of query topics along with their relevance judgments are given. The main limitation of this kind of evaluation paradigm is that it focuses almost exclusively on a small set of judged documents and does not consider what influence the given retrieval models have on accessing all the relevant information in the collection. This is particularly important for recall oriented retrieval applications where we want to ensure that that everything relevant has been found.

In this thesis we analyze the effectiveness of retrieval models from the documents retrievability point of view. We focus particularly on the retrieval bias of different retrieval models, and try to examine to what extent this bias restricts the users in retrieving relevant information. We explore this research with the help of three factors. First, we analyze the relationship between different characteristics of queries and retrievability. This is important from the query generation point of view, since in case of exhaustive queries, it is practically infeasible to complete retrievability approximation in reasonable time. The strong correlation between retrievability and query characteristics allows us to approximate the retrievability score accurately with the help of a query subset without processing an exhaustive number of queries. After this, we examine to what extent the retrievability and other IR effectiveness measures are related to each other. This specifically helps us to understand to what extent it is possible to automatically rank the effectiveness of retrieval models on the basis of their retrieval bias. This also offers a basis for optimizing retrieval systems for specific collections without the need to provide manually annotated ground truth. This is particularly useful for those retrieval domains where it is difficult to obtain a sufficient amount of relevance judgments. At the end we investigate and devise different retrieval strategies for mitigating the effect of low retrievability of documents. These include collection partitioning and query expansion on the basis of improved pseudo relevance feedback selection.

The work present in this thesis provides an a novel approach for the evaluation and optimization of retrieval models particularly for recall oriented retrieval domains, where the focus is on retrieving all relevant information but not just retrieving a subset of relevant information. Available online at: [http://www.ifs.tuwien.ac.at/~bashir/shariqbashir\\_phd\\_thesis.pdf](http://www.ifs.tuwien.ac.at/~bashir/shariqbashir_phd_thesis.pdf)