Chapter 20
A Framework for Evaluating the Retrieval Effectiveness of Search Engines

Dirk Lewandowski
Hamburg University of Applied Sciences, Germany

ABSTRACT
This chapter presents a theoretical framework for evaluating next generation search engines. The author focuses on search engines whose results presentation is enriched with additional information and does not merely present the usual list of “10 blue links,” that is, of ten links to results, accompanied by a short description. While Web search is used as an example here, the framework can easily be applied to search engines in any other area. The framework not only addresses the results presentation, but also takes into account an extension of the general design of retrieval effectiveness tests. The chapter examines the ways in which this design might influence the results of such studies and how a reliable test is best designed.

INTRODUCTION
Information retrieval systems in general and specific search engines need to be evaluated during the development process, as well as when the system is running. A main objective of the evaluations is to improve the quality of the search results, although other reasons for evaluating search engines do exist (Lewandowski & Höchstötter, 2008). A variety of quality factors can be applied to search engines. These can be grouped into four major areas (Lewandowski & Höchstötter, 2008):

- Index Quality: This area of quality measurement indicates the important role that search engines’ databases play in retrieving relevant and comprehensive results. Areas of interest include Web coverage
A Framework for Evaluating the Retrieval Effectiveness of Search Engines

(e.g., Gulli & Signorini, 2005), country bias (e.g., Liwen, Vaughan, & Thelwall, 2004; Liwen, Vaughan, & Zhang, 2007), and freshness (e.g., Lewandowski, 2008a; Lewandowski, Wahlig, & Meyer-Bautor, 2006).

- Quality of the results: Derivates of classic retrieval tests are applied here. However, which measures should be applied and whether or not new measures are needed to satisfy the unique character of the search engines and their users should be considered (Lewandowski, 2008d).

- Quality of search features: A sufficient set of search features and a sophisticated query language should be offered and should function reliably (e.g., Lewandowski, 2004, 2008b).

- Search engine usability: The question is whether it is possible for users to interact with search engines in an efficient and effective way.

While all the areas mentioned are of great importance, this chapter will focus on ways in which to measure the quality of the search results, a central aspect of search engine evaluation. Nonetheless, it is imperative to realize that a search engine that offers perfect results may still not be accepted by its users, due, for example, to usability failures.

This chapter will describe a framework for evaluating next generation search engines, whether they are Web search engines or more specific applications. A search engine in the context of this chapter refers to an information retrieval system that searches a considerably large database of unstructured or semi-structured data (as opposed to a general information retrieval system that searches a structured database). A next-generation search engine is a search engine that does not present its results as a simple list, but makes use of advanced forms of results presentation; that is to say, they enrich the list-based results presentation with additional information or the results are presented in a different style. Thus, the results are unequally presented in terms of screen real estate, that is, in terms of the area on the results screen that each results description is granted.

While the term “search engine” is often equated with “Web search engine,” in this chapter, all types of search engines are considered, although points regarding Web search engines are particularly emphasized, as they constitute the major area of our research.

The remainder of this chapter is organized as follows: First, a certain amount of background information on information retrieval evaluation is presented; next, the relevant literature related to such areas of interest as search engine retrieval effectiveness tests, click-through analysis, search engine user behavior, and results presentation. Then, a framework for search engine retrieval effectiveness evaluation, which will be described in detail, is presented. The chapter concludes by alluding to future research directions and with a set of concluding remarks.

BACKGROUND

Evaluation has always been an important aspect (and an important research area) of information retrieval. Most studies follow (at least to a certain degree) the Cranfield paradigm, using a set of ad-hoc queries for evaluation and calculating effectiveness measures such as precision and recall. While the Cranfield paradigm has often been criticized, it is not without merit and is used in large evaluation initiatives such as TREC and CLEF, which were designed to evaluate search results. Most search engine evaluations today are “TREC-style,” as they follow the approach used in these tests.

They use, however, a somewhat limited understanding of a user’s behavior. As will become apparent in the literature review section, user results are determinant upon selection behavior,
22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage: www.igi-global.com/chapter/framework-evaluating-retrieval-effectiveness-search/64437?camid=4v1


Related Content

Survey of Clustering: Algorithms and Applications
www.igi-global.com/article/survey-of-clustering/100038?camid=4v1a

SAR: An Algorithm for Selecting a Partition Attribute in Categorical-Valued Information System Using Soft Set Theory
www.igi-global.com/article/sar-algorithm-selecting-partition-attribute/68375?camid=4v1a

On the Design and Implementation of Interactive XML Applications
www.igi-global.com/chapter/design-implementation-interactive-xml-applications/75898?camid=4v1a

Concept of Association Rule of Data Mining Assists Mitigating the Increasing Obesity
www.igi-global.com/article/concept-of-association-rule-of-data-mining-assists-mitigating-the-increasing-obesity/177279?camid=4v1a