Chapter 8
Credibility in Web Search Engines

Dirk Lewandowski
Hamburg University of Applied Sciences, Germany

ABSTRACT

Web search engines apply a variety of ranking signals to achieve user satisfaction, i.e., results pages that provide the best-possible results for the user. While these ranking signals implicitly consider credibility (e.g., by measuring popularity), explicit measures of credibility are not applied. In this chapter, credibility in Web search engines is discussed in a broad context: credibility as a measure for including documents in a search engine’s index, credibility as a ranking signal, credibility in the context of universal search results, and the possibility of using credibility as an explicit measure for ranking purposes. It is found that while search engines—at least to a certain extent—show credible results to their users, there is no fully integrated credibility framework for Web search engines.

INTRODUCTION

Search engines are used for a wide variety of research purposes, and they are often the first place people go when searching for information. Users select search results and then read Web pages based on their decisions about whether the information presented by the search engine is of value to them. Search engines are so popular that they are, together with e-mail, the most commonly used service on the Internet (Purcell, 2011). Every day, billions of queries are entered into the search boxes of popular search engines such as Google and Bing. Therefore, these engines play an important role in knowledge acquisition, not only from an individual’s point of view, but also from society’s point of view. It is astonishing to what degree users trust search engines (Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010; Pan et al., 2007), and users rely on Web search engines to display the most credible results first.

Search engine rankings, however, do not guarantee that credible pages are ranked first for every topic. The construct underlying the rankings of search engines basically assumes that a page’s popularity equals credibility, although other factors also play a role. Popularity, in this case, refers to popularity among all Web authors and readers by measuring the distribution of links, clicks within the search engine results pages, time spent read-
ing the results documents, and recommendations in social media. While applying these factors in ranking algorithms often leads to good results, it should be stressed that these popularity measures always rely on the users judging the credibility of the documents, i.e., only people make credible pages popular.

Technical means for finding suitable indicators for credible Web pages are an alternative to human credibility judgments about Web search engine results (Mandl, 2005, 2006). Apart from popularity analyses, page and text properties can be used to estimate the credibility of a document, although such approaches can only provide estimates. Before discussing credibility further, we first need to define the concept in the context of search engine results.

According to the Encyclopedia of Library and Information Sciences (Rieh, 2010), credibility is an intuitive and complex concept that has two key dimensions: trustworthiness and expertise. Both are judged by people consuming information, and therefore, credibility always lies in the eye of the beholder. In discussing credibility in the context of Web search engines, we follow Rieh’s (2010) definition: “Credibility is defined as people’s assessment of whether information is trustworthy based on their own expertise and knowledge” (p. 1338). However, as search engines rate documents algorithmically, we need to consider “people” not only being users of information, but also designers of search engines and their ranking algorithms, which have certain assumptions about credibility that are then used in the system.

Tseng and Fogg’s (1999) four types of credibility provided a deeper understanding of credibility:

1. Presumed credibility, where people have general assumptions about a source of information (e.g., assuming that a friend will tell them the truth, or that articles written by full-time journalist will give credible information)
2. Reputed credibility, where sources of information are seen as credible because third parties assigned credibility to them in the past. E.g., the title of doctor or professor makes most people believe that this person is a credible source of information.
3. Surface credibility, where credibility is given to a source of information because of surface criteria, such as the jacket design of a book or the layout of a webpage.
4. Experienced credibility, where the person judging credibility has first-hand experience with the source of information.

In information retrieval (IR) evaluations, users judge the relevance of documents linked to a search query or information need in a controlled environment. However, while the concept of relevance somehow incorporates credibility, it also incorporates many other aspects. If expert jurors who are instructed to research the credibility of the documents are asked, then statements about credibility can be made. However, such studies are rare, mainly because expert jurors are expensive and the process of evaluating credibility is time-consuming. Furthermore, there is no tradition of credibility evaluation in IR because in traditional IR systems (e.g., newspaper databases or patent collections), the quality of the documents is controlled in the process of producing the database (Rittberger & Rittberger, 1997), and only documents from collections deemed credible are included.

Information quality frameworks (e.g., Knight & Burn, 2005) are of only limited use in the context of Web search engine because the main problem is that search engine users apply credibility judgments when considering (1) the results descriptions (“snippets”) on the search engine results pages, and (2) the results documents themselves. In both cases, they have only limited resources for judging credibility. It is much easier to apply information quality criteria to the inclusion of documents into an information system than applying such criteria to the ranking of documents, or even to rely on the system’s users to judge the quality of the documents.