Chapter 96
Human–Centred Web Search

Orland Hoeber
Memorial University of Newfoundland, Canada

ABSTRACT

People commonly experience difficulties when searching the Web, arising from an incomplete knowledge regarding their information needs, an inability to formulate accurate queries, and a low tolerance for considering the relevance of the search results. While simple and easy to use interfaces have made Web search universally accessible, they provide little assistance for people to overcome the difficulties they experience when their information needs are more complex than simple fact-verification. In human-centred Web search, the purpose of the search engine expands from a simple information retrieval engine to a decision support system. People are empowered to take an active role in the search process, with the search engine supporting them in developing a deeper understanding of their information needs, assisting them in crafting and refining their queries, and aiding them in evaluating and exploring the search results. In this chapter, recent research in this domain is outlined and discussed.

INTRODUCTION

Search has become a critical element of modern, computationally augmented life. In the past, people would spend a little bit of time organizing the information in their lives, and then browse it when they needed to find something. As we move toward having technology become a part of almost every aspect of our lives, the amount of information that must be managed is becoming larger and larger. In many cases, the Web is being used as

the universal platform for making this information available. With so much information at our disposal, browsing is no longer feasible and our only alternative is to search. We search the Web, our email, our contacts lists, our calendars, our computers, our mobile devices. Search has become a daily, if not hourly, activity for many people. As such, researchers need to begin considering not only what would be appropriate techniques and interfaces for searching for documents on the Web, but also for finding information within more specific domains and settings.

In recent years, a large portion of the research and development in the domain of information retrieval has focused on the "back-end" aspect of search. Herein, much research has addressed issues related to the indexing of documents (Baeza-Yates & Ribeiro-Neto, 2011; Rasmussen, 2003), file systems for managing indexes at the scale of the Web (Ghemawat, Gobioff, & Leung, 2003), and algorithms to match and rank documents to queries (Brin & Page, 1998). To a large degree, the growth of companies such as Google and Yahoo! has been due to their ability to efficiently index the billions of documents on the Web, and deploy an infrastructure that allows them to respond to end-user queries in fractions of a second. Of course, the quality of the search result ranking is also an important element of their success.

Because of the popularity and success of these companies, many people consider Web search to be a solved problem. However, there is still a great deal more that can be done to enhance not only how Web search engines work (the focus of this book), but also how Web search engines help people to find the information they are seeking (the focus of this chapter). The key distinction to be made when considering the human element of Web search is that the primary focus changes from the documents being searched to the support that can be provided to the people who are doing the searching (Hoeber, 2008).

Unfortunately, there has been very little progress in terms of Web search engines' support for the fundamental search tasks that people need to perform. The interface for extracting searchers' information needs remains a simple text field. The ranked set of documents that are matched to the query continues to be represented in a simple list of document titles, snippets, and URLs. While improving these interfaces to more adequately support the primary search tasks of query specification and search results evaluation is a good starting point, a more thoughtful consideration of the types of things people really need (or want) to do within the context of Web search is required.

In this chapter, we explore some of the recent research trends that address the needs of searchers beyond providing a simple query field and search results list. We take a human-centred computing perspective on this topic, wherein the primary focus is on the design, development, and evaluation of computing systems that consider and support human activities (Sebe, 2010). We start with discussions on Web searcher behaviour and what can be done to support their needs, and then provide an overview of a number of specific approaches that have shown promise. Techniques for evaluating such interfaces are discussed, along with a vision for future research directions within this domain.

SEARCHER BEHAVIOUR

Even with the simple and easy-to-use interfaces provided by the top Web search companies, people commonly have difficulties searching the Web. Numerous studies have been conducted in recent years that attempt to characterize the behaviour of Web searchers (Jansen, Booth, & Spink, 2007; Jansen & Pooch, 2001; Jansen & Spink, 2006; Silverstein, Henzinger, Marais, & Moricz, 1999; Spink, Wolfram, Jansen, & Saracevic, 2001). Two things that most of these studies agree on are that people who search the Web use very short queries; and they don't consider many search results. Indeed, most Web search queries are between one and three terms in length, and few people venture past the third page of the search results.

In order to craft a query, all that is provided in terms of guidance from the search engine interface is a text box in which to type the query. But what if the searcher doesn't know what to type? Or what if the searcher types the wrong thing? There have been some recent additions to the top search engines to suggest common queries as a searcher begins typing their query, and even to provide the search results for these queries in real-time. While this can be a useful feature when searching for a common topic, it can also have

Related Content

RETAD: Vehicle Trajectory Anomaly Detection Based on Reconstruction Error

Chaoneng Li, Guanwen Feng, Yiran Jia, Yunan Li, Jian Jiand Qiguang Miao (2023). *International Journal of Data Warehousing and Mining (pp. 1-14).*

www.irma-international.org/article/retad/316460

Synthetic Population Techniques in Activity-Based Research

Sungjin Cho, Tom Bellemans, Lieve Creemers, Luk Knapen, Davy Janssensand Geert Wets (2014). *Data Science and Simulation in Transportation Research (pp. 48-70).*

www.irma-international.org/chapter/synthetic-population-techniques-in-activity-based-research/90065

A Hybrid Method for High-Utility Itemsets Mining in Large High-Dimensional Data

Guangzhu Guangzhu Yu, Shihuang Shao, Bin Luoand Xianhui Zeng (2011). *Integrations of Data Warehousing, Data Mining and Database Technologies: Innovative Approaches (pp. 60-76).*

www.irma-international.org/chapter/hybrid-method-high-utility-itemsets/53072

Spatial OLAP and Map Generalization: Model and Algebra

Sandro Bimonte, Michela Bertolotto, Jérôme Genseland Omar Boussaid (2012). *International Journal of Data Warehousing and Mining (pp. 24-51).*

www.irma-international.org/article/spatial-olap-map-generalization/61423

Data Stream Mining Using Ensemble Classifier: A Collaborative Approach of Classifiers

Snehlata Sewakdas Dongreand Latesh G. Malik (2017). *Collaborative Filtering Using Data Mining and Analysis (pp. 236-249).*

www.irma-international.org/chapter/data-stream-mining-using-ensemble-classifier/159506