The Journal of Academic Librarianship 42 (2016) 2–7

Contents lists available at ScienceDirect

The Journal of Academic Librarianship

Measuring the Effect of Virtual Librarian Intervention on Student Online Search

Chris Leeder *, Chirag Shah

School of Communication and Information, Rutgers University, 4 Huntington Street, New Brunswick, NJ 08901-1071, USA

ARTICLE INFO

Article history:
Received 7 July 2015
Accepted 3 September 2015
Available online 26 September 2015

Keywords:
Information seeking behavior
Online search
Students
Librarians

ABSTRACT

In this paper we describe the results from a case study of the online search behavior of high school students, and the effect of virtual librarian intervention on the quality of the search results. Search log data of students' actions were analyzed, librarians recommended revised search terms, new searches were conducted with these terms, and the difference in results was analyzed. Results showed that the librarians' recommended queries were more focused and produced results of a higher reading level, which was used as a proxy for source quality. The work presented here is a preliminary investigation of how librarian search expertise can improve online search results, and raises questions about how to support online student search that are of interest to both librarians and information science researchers. Possible directions for future research are discussed.

INTRODUCTION

When searching for information, most students instinctively begin their research by using the Internet (Curtis, 2000; Herring, 2011; Mizrachi, 2010). Specifically, students overwhelmingly rely on Google to the exclusion of many academic search tools (Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010; Head & Eisenberg, 2010; Kolowich, 2011). In 2015, Google remained the most visited website with more than 178.02 million visitors and comprised 76% of search engine usage (Comparison of unique U.S., 2015). Teens are particularly avid users of the Internet, with data from the Pew Research Center showing that 92% of teens report going online daily, including 24% who say they go online "almost constantly" (Teens, Social Media, & Technology Overview [Online], 2015). When conceptualizing online search, "students see Google as being 'the' Internet, and they use these two terms interchangeably, seeing them to be one and the same thing" (Julien & Barker, 2009, p. 14). Despite this, they often do not know how search engines such as Google operate (Kuiper, Volman, & Terwel, 2008). Even so, they often rely on the top items on a search engine results page (SERP) and barely go beyond the first few pages (Granka, Joachims, & Gay, 2004; Hargittai et al., 2010; Lankes, 2008), even when the abstracts of the search results were less relevant than other results (Pan et al., 2007).

Search engines currently provide automated query expansion to assist users with their searches, but the suggested queries are often out of context and based on popular searches, rather than on the specific information need of searchers. Librarians are search experts and can provide customized, context-specific query recommendations, however their professional orientation is toward guiding students to high quality library resources, rather than to the open web. Also, their work is one-on-one, individualized, and time-intensive, and thus cannot be easily scaled to the Internet. Beyond providing in-person reference consultations and in-class instruction to students on the usage of library resources, can customized librarian guidance on search be applied to improve results during real-time student information seeking online?

This research describes a case study of how high school students search for information for a school assignment, and how the revised search terms suggested by librarians can improve the quality of search results. The data collection took place in a classroom and captured the authentic search behaviors of high school students working on a class assignment, rather than an imposed query chosen by researchers and tested in a laboratory setting. Subsequently, a virtual intervention by librarians was conducted by gathering revised search term recommendations, and measuring the difference in quality of results. This study is a preliminary step in empirically quantifying how the expertise of librarians improves the results of student online searches.

LITERATURE REVIEW

The scope and volume of information on the Web requires good search skills such as the ability to formulate relevant keywords to find the information one is looking for (Kuiper et al., 2008). However, studies show that most students are unable to narrow down topics to make them manageable, and are overwhelmed by the amount of results provided by web search engines (Head, 2007), especially when they do not possess the skills or resources to access and manage such information intelligently (MaKinster, Beghetto, & Plucker, 2002). Instead, students use very simple or basic forms of search, and assume that search engines 'understand' their queries (Rowlands et al., 2008). Most queries are

* Corresponding author.
E-mail address: chris.leeder@rutgers.edu (C. Leeder).

http://dx.doi.org/10.1016/j.acalib.2015.09.001
0099-1333/

simple keyword queries (Aula, Khan, & Guan, 2010). Students do not examine their topics to identify keywords and related terms, instead relying heavily on the language of their research topics (Georgas, 2014). Many students show an inflexible Web strategy, sticking to one particular search strategy despite lack of success, perhaps because they expect Google to find the answer for them regardless of their own strategy (Kuiper et al., 2008).

One of the primary methods of improving search results is query reformulation, the process of altering a given query to improve search or retrieval performance. Query reformulation has been widely researched in the information science field (e.g. Huang & Efthimiadis, 2009, Jansen, Booth, & Spink, 2009, Rieh, 2006). Given that queries are the primary expression of the searcher's information need, they are an integral part of the information seeking process (Hembrooke, Granka, Gay, & Liddy, 2005; Jansen et al., 2009).

However, the formulation and reformulation of queries is one of the most difficult and challenging tasks for users, even in a Web search environment that appears to be easy to learn and use (Rieh, 2006). It can be very difficult for the user to think of additional or alternative query terms (Aula & Nordhausen, 2006). Many students struggle with identifying and structuring useful keywords and phrases for their search (MaKinster et al., 2002). Research has shown that effective query reformulation can improve the outcome of user searches (Jansen et al., 2009). Strategies such as adding or removing words, word substitution, acronym expansion, and spelling correction are more likely to increase the effectiveness of search results (Huang & Efthimiadis, 2009). Based on evaluating the quality of search results and relevance to the task, more successful searchers are more systematic in their query formulation whereas less successful searchers' refinement strategy is more random (Aula et al., 2010).

One factor that influences the effectiveness of searches is the level of expertise of the user (Vakkari, Pennanen, & Serola, 2003; Wildemuth, 2004). White and Morris (2007) found that advanced searchers are generally more successful than less advanced searchers. Hembrooke et al. (2005) found that novices engage in less effective strategic search behavior. MaKinster et al. (2002) demonstrated that domain knowledge and search expertise are critical factors that account for differing levels of success in online search. Domain knowledge refers to one's knowledge of the larger context of the specific topic that searchers are attempting to locate, e.g., science knowledge or knowledge of a specific topic area. Search expertise pertains to the familiarity with the structure and content of search results, website descriptions, how to construct a search phrase, how to navigate a website, and knowledge about the use of search engines (MaKinster et al., 2002). Domain experts search differently than people with little or no domain knowledge. Zhang, Liu, and Cole (2013) found that domain knowledge showed significant effects on users' search behaviors and search performance. White, Dumais, and Teevan (2009) found that domain experts employ different search strategies and are more successful in finding what they are looking for than non-experts. Drabenstott (2003) found that novices tend not to employ the search strategies of experts (author searching, citation searching, footnote chasing, journal runs, and known-item searching) but instead tend to rely on non-expert strategies such as subject and keyword searching. The user's level of knowledge in the domain of search will affect the type and number of terms selected for incorporation in the search strategy (Wildemuth, 2004). Thus, there are potential benefits to non-experts in identifying expert search strategies and applying them to improve search results (White et al., 2009).

Novice searchers receive minimal guidance from search interfaces regarding query construction, and would benefit from more support (Hembrooke et al., 2005). Since high school students are generally neither domain experts nor search experts, they are particularly in need of instructional support to improve their online information seeking behavior. Drabenstott suggests that "librarians and instructors have an important role to play to transition students from non-domain to expert-domain strategies" (p. 85) and advocates for scaffolding domain novices from their usual strategies into the strategies characteristic of experts. However, the effectiveness of librarians in improving online student search has not been studied in online contexts. Most LIS research on search strategy involves studies of transaction logs from online library catalogs. Fields, Keith, and Blandford (2005) reported that librarians searching a digital library showed excellent skills in reformulating their queries so that each reformulation built on the previous query and the retrieved results. The librarians were able to call upon a number of strategies or recurrent patterns of activity that librarians made effective use of the features of the user interface. They describe three common strategies: systematically reformulating queries (rather than just abandoning them and starting over); manipulating a search query in order to increase or decrease the size of the set of search results; and exploring the result sets so as to learn more about the domain of inquiry and related search terms within it. Ge and Land (2004) suggest that librarians can provide expert modeling by demonstrating to students how to effectively conduct the search process.

Librarians use a variety of measures to evaluate the impact of their services to patrons, although these usually apply to in-person instruction. Studies have shown that college students who participate in information literacy classes report significantly less library anxiety (Van Scoyoc, 2003) and high-achieving students are more likely to report experiencing formal information literacy instruction (Gross & Latham, 2007; Smalley, 2004). Wang (2006) found statistically significant differences in grades between college students who took a library credit course and students who did not, and those who had taken the instruction in library skills received higher grades on their papers and in their courses. Ren (2000) found that receiving library instruction significantly increased college students' self-efficacy in electronic information search. These qualitative measures all require long-term tracking of students who have received instruction to show that instruction had an impact at later points in students' careers.

One quantitative measurement frequently used by librarians is citation analysis of student bibliographies. Librarians compare the bibliographies of treatment groups that received library instruction and control groups that did not. The quality of bibliographies is often measured by the student's use of scholarly sources, which is obtained by grading each citation against a set of standard scoring criteria (e.g. Davis & Cohen, 2001, Knight-Davis & Sung, 2008, Young & Ackerson, 1995). However, it is difficult to apply these criteria to online sources, as many formats such as blogs, forums, and digital repositories do not fit neatly into the traditional bibliographic format categories. Simply categorizing an online site as a Web page, as many earlier online studies have done, does not provide enough information to judge its quality as an information source. It is also difficult and potentially inaccurate to judge online sources by a binary choice such as "scholarly vs. nonscholarly." Multiple other criteria must also be considered (Leeder, Markey, & Yakel, 2012). A limitation of both of these techniques is that no direct cause-and-effect relationship can be demonstrated between library instruction and the measured outcomes, as many intervening variables may also influence the results.

There has been little research in LIS attempting to quantify the effect of librarian intervention in student online search. A report by the Association of College and Research Libraries (ACRL) found that the literature on IL assessment is "sporadic, disconnected, and reveals limited snapshots of the impact of academic libraries on learning" and argues that "systematic, coherent, and connected evidence is required" (Association of College and Research Libraries, 2010, p. 8). The effect of librarians' expertise on improving the quality of online search results has not been experimentally studied. Part of the reason for this is the lack of empirical data and analysis involving students' information seeking and synthesis processes, as well as the lack of understanding of the types of possible interventions that librarians could provide. A preliminary exploration of this research area was conducted using data from a case study of the online searching behavior of high school students,