A Review of Web Searching Studies and a Framework for Future Research

Bernard J. Jansen
Computer Science Program, University of Maryland (Asian Division), Seoul, 140-022 Korea. E-mail: jjansen@acm.org

Udo Pooch
E-Systems Professor, Computer Science Department, Texas A&M University, College Station, TX 77840. E-mail: pouch@cs.tamu.edu

Research on Web searching is at an incipient stage. This aspect provides a unique opportunity to review the current state of research in the field, identify common trends, develop a methodological framework, and define terminology for future Web searching studies. In this article, the results from published studies of Web searching are reviewed to present the current state of research. The analysis of the limited Web searching studies available indicates that research methods and terminology are already diverging. A framework is proposed for future studies that will facilitate comparison of results. The advantages of such a framework are presented, and the implications for the design of Web information retrieval systems studies are discussed. Additionally, the searching characteristics of Web users are compared and contrasted with users of traditional information retrieval and online public access systems to discover if there is a need for more studies that focus predominantly or exclusively on Web searching. The comparison indicates that Web searching differs from searching in other environments.

Introduction

On-line searching by users is now the norm in universities, businesses, and households. The majority of studies concerning on-line searching can be categorized as focusing on either traditional information retrieval (IR) systems or online public access catalogue (OPAC) systems (Borgman, 1996). With more than a 40-year history, studies in these categories have explored a variety of systems, document collections, and user characteristics. However, due to varying structure, it is difficult and, in some cases, impossible to compare user-searching characteristics among studies.

When one attempts a comparison, questions such as: *Are the searching characteristics really common to all searchers of these types of systems?* or *Are the exhibited searching characteristics due to some unique aspect of the system or document collection?* or *Are the user samples from the same or different populations?* invariably arise.

With the advent of the World Wide Web (Web), a new category of searching now presents itself. The Web has had a major impact on society (Lesk, 1997; Lynch, 1997), and comes the closest in terms of capabilities to realizing the goal of the Memex (Bush, 1945). In terms of quality, Zumalt and Pasicznyuk (1998) show that the utility of the Web may now match the skills of a professional reference librarian. The Web possesses an ever-changing and extremely heterogeneous document collection of immense proportions. Although developed in an apparently unstructured environment, Web document discovery is extremely structured in terms of its hyperlinks. The user population of the Web is enormous and extremely diverse, albeit with certain groups over represented (Hoffman, Kalsbeek, & Novak, 1996; NTIA, 1999). The networked and dial-in connectivity available to Web searchers creates a near ubiquitous on-line searching environment. The Web's IR systems are also unique in terms of the interface, advertising constrains, bandwidth restrictions, and unique document indexing issues (e.g., spamming and URL hijacking). In sum, the Web appears to be a whole new searching environment (Sparck-Jones & Willett, 1997).

Given its distinctive searching features, one would expect a significant number of studies on characteristics of Web searching. However, there have been surprising few detailed studies on Web or Internet searching (Peters, 1993). The reason for the limited research may be that it is extremely challenging to construct valid studies of searching in an environment like the Web (Robertson & Hancock-Beaulieu, 1992). Nevertheless, the current small number of

Web searching studies provides an opportunity to develop a common framework for future studies that will allow for the comparing and contrasting of study results. Developing this framework now will be much easier than it will be 40 years from now when Web searching studies may be as numerous as traditional IR and OPAC system studies are today. The research presented in this article takes the first step toward this goal by reviewing the state of research in the field, validating the need for such studies, identifying trends in the research, and suggesting a framework for future studies.

Beginning with a review of models of traditional user-searching studies, we follow with the current state of Web searching studies, presenting the results from all published studies found. From these studies, the aggregate statistics are compared with searching on traditional IR and OPAC systems to highlight possible similarities and differences. Based on the review, a framework for future research on Web searching is proposed. We conclude by highlighting the possible directions of future Web research and the implications for Web study design.

Review of Literature

User studies can be viewed as a subset within the larger area of IR system evaluation, which typically focuses on measuring the recall and precision of the system (Sparck-Jones, 1981). The theoretical underpinnings for this type of IR evaluation are well defined (Salton & McGill, 1983), although the proper metrics are still a topic of debate (Saracevic, 1995). In this type of evaluation, one takes a known document collection with documents classified as relevant or nonrelevant, based on a set of queries. These queries are executed using a particular IR system against the document collection. Based on the number of relevant and nonrelevant documents retrieved, one determines recall and precision. This is a systems view of relevance, with recall and precision directly related to the queries entered. The whole process is very systematic.

However, once a "real" searcher is interjected into the system, the evaluation metrics are no longer so straightforward. Relevance to a searcher is not clearly defined (Mizzaro, 1997; Saracevic, 1975; Spink, Greisdorf, & Bateman, 1998). In fact, it is not even certain how a searcher conducts the search process, although there are several theories on the information seeking process (Belkin, Oddy, & Brooks 1982; Saracevic, 1996) that attempt to explain it. Most of these theories are based on empirical analyses of users and, in many cases, the studies do not agree with one another about user-searching processes.

Transaction Log Analysis

Transaction logs are a common method of capturing characteristics of user interactions with IR systems. Given the current nature of the Web, transaction logs appear to be the most reasonable and nonintrusive means of collecting user-searching information from a large number of users.

Transaction log analysis (TLA) uses transaction logs to discern attributes of the search process, such as the searcher's actions, the interaction between the user and the system, and the evaluation of results by the searcher. TLA lends itself to a grounded theory approach (Glaser & Strauss, 1967) in that the characteristics of searches are examined to isolate trends that identify typical interactions between searchers and the system. If one views the information-seeking process as consisting of five entities (Saracevic, Kantor, Chamis, & Trivison, 1988), TLA can only deal with entity number four, the searcher's actions. In this respect, TLA is limited (Peters, 1993); however, TLA can provide some necessary data. For example, depending on the specifics of the transaction log, one can gather knowledge about the formulation of the search, the search strategy, and the delivery of results (such as number of items retrieved and documents viewed). Moreover, if one knows and accepts the limitations of TLA, it can be beneficial for understanding the system itself and the user interactions during the search process. Kaske (1993) and Kurth (1993) both discuss the strengths and weaknesses of TLA, whereas Sandore (1993) reviews methods of applying the results of TLA. For a historical review of TLA, see Peters (1993).

Defining Web Searching Studies

With their reliance on transaction logs, Web searching studies typically lack the context and relevance judgments common in other studies of IR systems. A Web-searching study *focuses on isolating searching characteristics of searchers using a Web IR system via analysis of data, typically gathered from transaction logs*. Note that the focus on searching characteristics excludes other valuable Web research, such as nonempirical pieces (Hawkins, 1996) on Web searching, examinations of IR system search algorithms (Zorn, Emanoil, & Marshall, 1996), studies of search engine coverage (Lawrence & Giles, 1998; Gordon & Pathak, 1999), and general Web-user characteristics (Pitkow, 1999). It also excludes Web studies that are analyses solely of surfing techniques (Crovella & Bestavros, 1996; Huberman, Pirolli, Pitkow, & Lukose, 1998), demographic studies (Hoffman et al., 1996; Kehoe, Pitkow, & Morton, 1999), and studies that utilize small user samples in a controlled setting (e.g., Choo, Betlor, & Turnbull, 1998; Pharo, 1999). Although all of these studies are worthwhile in explaining other respects of the Web, they provide limited empirical information about the searching characteristics of Web users.

Review of Web-Searching Studies

We conducted an extensive literature search using online sources, conference proceedings, applicable journals from a broad range of fields, Web sites, article bibliographies, and personal contact with various researchers in the field. Our goal was to collect all published Web searching studies into

a comprehensive literature review to gauge the current state of the literature. Although every attempt was made to gather all of the published materials relevant to Web searching, it is possible that we may have missed some articles or presentations.

The literature review is divided into two sections. Section 1, *Primary Web-searching Studies*, is a review of all Web-searching studies that deal with studies of searching using Web search engines. These studies contained a substantial amount of data and addressed a broad range of Web-searching characteristics. Section 2, *Secondary Web-searching Studies,* is a review of Web searching studies that are more limited in that they do not present enough data to give a full picture of Web searching. Most of these studies analyzed Web searching on a single Web site that was not a search engine. Others intentionally analyzed a narrow aspect of Web searching, such as term analysis or relevance feedback. All studies except one utilized transaction logs to capture the data. In cases where there were multiple publications of the same or a very similar study, we refer only to the most recent publication.

Primary Web-Searching Studies

The three studies in this section analyzed searching on Web search engines, which are the major portals for users of the Web. Search engines are the IR systems of the Web. Data shows that 71% of Web users access search engines to reach other Web sites (CommerceNet/Nielsen Media, 1997). One in every 28 (3.5%) pages viewed on the Web is a search results page (Alexa Insider, 2000), making the use of a search engine the second most popular Internet task next to email (Statistical Research, Inc., 2000). Users rate searching as the most important activity conducted on the Internet (Jupiter Research, 2000). The three studies we found dealt with only three search engines: Fireball, Excite, and Alta Vista.

The Fireball Study. Hoelscher (1998) analyzed data from Fireball (http://www.fireball.de/), a German Web IR system. The data set was approximately 16 million queries processed by Fireball during July 1998. The 16 million queries consisted of about 27 million terms. The average query length was 1.66 terms. Over 54% (8,873,001) of the queries contained only one term, although there were a significant number of queries with lengths of two (5,005,653 queries representing 30.80%) and three terms (1,683,129 queries representing 10.36%). Less than 2% of the queries contained five or more terms. For query complexity, the vast majority of Fireball queries (over 97%) utilized no Boolean operators. Phrase searching was utilized in just over 8% (1,401,738) of the queries. The most utilized modifier was the "+" modifier. Over 59% (9,621,347) of the users examined no more than one page of results. Fireball presents ten results at a time. Over 79% of the users examined 30 or fewer results.

The Fireball Study presented a wide range of data, especially concerning query structure. However, no information was provided concerning user sessions, and there was limited discussion of query terms. For example, it is not clear from Hoelscher's (1998) study how a *term* is defined: is it a string of any character? Is it only alphanumeric characters? Also, the Fireball search engine provided the summary statistics, not the raw data, to the researcher, making the particulars of how the transactions were logged and analyzed unknown. Little descriptive information about the Fireball search engine was provided, which is a serious shortcoming given the rapidly changing environment of the Web.

The Excite Study. Jansen, Spink, and Saracevic (2000) published a study concerning searching on Excite (http://www. excite.com). Excite ranked sixth among all Web sites in December 1999, in terms of traffic, with approximately 24 million hits (Nielsen/Net Rating, 1999), and it has one of the largest document collections of any Web IR system (Krishna & Broder, 1998). Jansen et al. (2000) analyzed 51,473 queries from 18,113 searchers from a 1997 data set. The researchers provide a detailed account of the raw data and the searching rules of the Excite search engine. The results reported include number of queries per user, number of terms per query, number of documents viewed, query complexity, query modification, and distribution and occurrence of terms.

Although the Jansen, Spink, and Saracevic (2000) study presents valuable information to the field of Web searching, their study has three major flaws. First, the data was collected from a portion of 1 day, providing limited longitudinal data. Second, the identification of sessions is not clear. The transaction log was composed of three fields: the unique identifier, the query, and the log-on time. The researchers utilized the unique identifier assigned by the Excite server to denote individual searchers. This field in the typical transaction log actually identifies unique computers on which the server deposits a cookie. Therefore, computers located in public areas would have one identifier, even though many users may have had access to them. This discrepancy would affect the analysis concerning the number of queries per user. Third, the count of relevance feedback queries is potentially inaccurate. Relevance feedback queries appeared in the transaction log as queries with no terms. However, studies show that users enter null queries during the normal searching process (Peters, 1993). In the Excite transaction log, a relevance feedback query and a null query appear as the same.

The Alta Vista Study. Silverstein, Henzinger, Marais, and Moricz (1999) presented results using Alta Vista (http:// www.altavista.com). Alta Vista is one of the largest Web search engines, with over 10 million hits per month (Nielsen/Net Rating, 1999) and a current document collection of over 250 million documents (Sullivan, 2000). This

study presents results from an analysis of just under a 1 billion queries submitted to the main Alta Vista search engine over a 42-day period. The authors provide an excellent overview of the transaction log and give ample description of the Alta Vista search engine. The analysis provides a broad spectrum of information at the session, query, and term level, including term correlation. Given the number of queries in the data set, length of the collection period and analysis, it is the most complete Web searching study to date.

In reviewing the Silverstein et al. (1999) study, there are three concerns. The first is the definition of a session. The researchers "time out" a session after 5 minutes, which one would expect has the effect of "shortening" the sessions, reducing the query per session count. Second, there are a significant number of metrics that are not clearly defined. The researchers report on *sessions*, *requests,* and *queries*; *unique queries* and *distinct queries;* and *exact-same-as-before requests,* among others. Without clear definitions of these metrics, it is difficult to judge the impact of the analysis. This is especially important with the distinct query subset, which was used for much of the discussion. Third, as the authors point out, submission by softbots may have skewed the data, as there are some extremely large standard deviations.

Secondary Web-Searching Studies

The studies in this section reported a limited amount of Web-user data or focused on a narrow aspect of Web-user searching. All but one of these secondary studies utilized transaction logs. Although not as robust as the three previous studies, these studies still contain some valuable insights into Web searching. The secondary studies in this section are broken down into the subsets of (1) Single Web Sites, (2) Multiple Web sites, (3) Searching Using Relevance Feedback, (4) Multimedia Searching, (5) Query Terms, (6) Presentations on Web searching, and (7) Information on the Web about the Web.

Single Web sites. Studies from a single Web site have the obvious disadvantage of a limited user sample that may not adequately represent the general Web population. Croft, Cook, and Wilder (1995) published searching data from THOMAS (http://thomas.loc.gov/), a collection of U.S. legislative information. The researchers analyzed 25,321 unique queries recorded over a 73-day period, and then calculated the 25 most common queries. Based on the data presented, one can calculate that the average term length was approximately 2.3 terms. The 25,321 unique queries were from a larger set of 94,911 queries, where at least one item was examined. Because these 94,911 queries are a subset of a larger collection of 196,724 accesses to the THOMAS query page, it can be determined that about 50% of the users did not enter a query. An interesting follow on study would be why so many users did not search for articles on the site.

Similarly, Jones, Cunningham, and McNab (1998) published research that focused on a single site: the New Zealand Digital Library (http://www.nzdl.org/), which is a collection of computer science documents. The study reported on 24,687 queries collected over a 427-day period. Although the collection period was lengthy, the traffic on the site was relatively minimal. The 24,687 queries over a 427-day period average out to approximately 77 queries per day. Also, given the technical nature of the document collection, the searchers may be a subset of the general Web searcher population. This is indicated by the high percentage of queries containing Boolean operators (over 25%), which is out of line with the results from other Web studies. This study may indicate that research should target specific subsets of the general Web searchers.

Multiple Web sites. Abdulla, Liu, and Fox (1998) used Web queries to identify inefficiencies in Web IR system design. The data was collected from both an U.S. college and a Korean college Web server. The data was collected in 1995 and 1996 (Abdulla, Liu, Saad, & Fox, 1997). The researchers categorized searchers into five groups and compared searching characteristics among these groupings. They concluded that the most common query length was two terms, although some queries contained more than five terms. In query syntax, the majority of queries contained no Boolean operators. This study is one of the few that contains data from a non-U.S. Web server, and also one of the earliest data collections. The statistics from the U.S. and non-U.S. server were similar.

He and Göker (2000) conducted a study using 51,474 queries from an Excite log and 9,534 queries from a local version of the Alta Vista search engine. The researchers focused on identifying a time interval that could be utilized to specify a session. Based on the analysis, the researchers concluded that a time interval of 10 to 15 minutes was the typically session length. This research could be valuable for Web IR systems that rely on cookies to define sessions.

Keily (1997) conducted a Web searching studying utilizing queries from WebCrawler (http://webcrawler.com/) and Magellan (http://www.mckinley.com/magellan/). The researcher obtained the queries using the "spy feature," that is, a feature that allows one to view queries of other searchers. Using 1,000 queries from each search engine, data is provided on the percentage of single word (33%), multiple word (67%) queries, phrase searching (about 10%), Boolean usage (approximately 12%), failure rates (56% for Magellan), and natural language searching (about 6%). These percentages for query length and complexity are higher than reported in other Web studies.

Searching using relevance feedback. Jansen, Spink, and Saracevic (1999) conducted an analysis of Web relevance feedback usage using data from Excite. Relevance feedback is a classic IR technique reported to be successful with many IR systems (Harman, 1992). The researchers con-

cluded that only about 3% of the queries (1,597) could have been generated by Excite's relevance feedback option. The researchers noted that approximately 80% of these sessions could be classified either as successful (63%) or partially successful (17%), indicating that relevance feedback can be beneficial for Web users.

Multimedia searching. Smith, Ruocco, and Jansen (1998) analyzed 851,770 queries from Excite to isolate the video specific queries. The researchers developed a list of terms relating to video and identified a total of 21,469 queries (2.52%) that were requests for video. The researcher conclude that video represents a small percentage of all query terms, approximately 2.5%.

Goodrum and Spink (1999) conducted a study concerning image searching using 1,025,910 queries from Excite. The researchers utilized 28 image terms to identify image queries. From the total set of queries, they isolated 33,149 image queries by 9,855 searchers. The researchers discuss the number of image queries, sessions that contained these queries, and terms that composed these queries. They determined that average session length was 3.36 image queries per user, and the average image query contained 3.74 terms. The most frequently occurring image-related terms occurred in less than 10% of the queries. An expanded study isolating separate image, audio, and video queries, was done by Jansen, Goodrum, and Spink (2000). These studies show that multimedia Web queries contain more terms than the average Web query, and that the number of multimedia terms utilized are extremely varied.

Query terms. What people are searching for seems to have a fascination for other Web users and researchers. Many Web search engines (e.g., http://www.metacrawler.com and http://www.savvysearch.com) allow users to view other queries being submitted to the system. There are also several sites on the Web that publish lists of popular Web search terms, such as Searchterms.com (http://www.searchterms.com/), and there are companies such as Adbar (http://www.adz.net/top200/) that track the top search terms for a given period. There are also academic papers that report data on Web query terms. An early study on Web terms by Selberg and Etzioni (1995) concerns the design issues of Metacrawler (http://www.metacrawler.com); however, the researchers devote a small section of the paper to reporting on queries submitted to the system. The data was collected from 7 July through 30 September 1995, and there were 50,878 total queries, of which 46.67% (24,253) were unique. The top 10 queries represented 3.37% (1,716) of all the queries, and all of the top 10 queries were related to sexual topics. All of these queries were only one term in length, and commonly occurring terms (e.g., the, of, and, or) reported in later Web user studies were not present. Wolfram (1999) conducted a term analysis focusing on term *co-occurrence* using approximately one million queries from Excite. Wolfram concluded that the term *co-*

occurrence did not follow a Zipf distribution, which is a common rank-frequency distribution of terms in long English texts. Wolfram states that the 10 most frequently occurring terms represented only 0.01% of the 96,004 unique search terms, but they constituted approximately 5% of all search terms used in the unique, multiterm set of queries. In a later article by Ross and Wolfram (2000), the top 1,000 term pairs were categorized into one or more of 30 subject areas. Ross and Wolfram (2000) show that there is some commonality among the most popular Web term co-occurrences.

Presentations on Web searching. There have also been a number of conference presentations from key personnel of major Web IR companies. Kirsch (1998), the chairman and founder of Infoseek Corporation, reported that the average query on Infoseek was 2.2 terms, about 10% of the queries contained Boolean operators, and the majority of queries were noun phrases. Kirsch (1998) also reported that of the top 15 Infoseek queries, at least 11 were sexual in nature. Most of this data confirms results reported by Cutting, a chief researcher for Excite (Lesk, Cutting, Pedersen, Noreault, & Koll, 1997).

Xu (1999) from Excite@Home reports that the average query length in the U.S. has increased from 1.5 terms in 1996 to 2.6 terms in 1999. The use of Boolean operators has increased from 22% in 1996 to 29% in 1999. He also states that Web searching is precision based with over 70% of Excite searchers looking at only the top ten results or the first page of results (Excite only displays 10 results or less per page). Over 29% of Excite searchers utilized suggestions from the on-line thesaurus. Xu (2000) presented similar data focusing exclusively on multilingual searching.

Kirsch's (1998) report of the top 15 queries could lead one to believe that the majority of Web queries relate to sexual topics. However, the two academic studies in this area (Jansen et al., 2000; Selberg & Etzioni, 1995) show that sexual topics in these data sets represented less than 4% of all queries or terms. Also, Lawrence and Giles (1999) report that nonpornographic Web sites are 50 times more common than pornographic sites.

Information on the Web about the Web. There are numerous Web sites that report information related to Web searching. There are sites that report information or contain articles on Web search engines (http://www.searchenginguide.org, http://www.searchenginewatch.com/, http://www.searchengineguide.com/, http://cyberatlas,internet.com/, http://www.internets.com/, http://pctataonline.com), the size of the Web (http://www.searchenginewatch.com/reports/sizes.html), the size of Web domains (http://www.domainstats.com/), common search terms (http://www.searchwords.com/, http://insider.alexa.com/), number of visitors to Web sites (http://www.searchwords.com/), and the Web growth rate (http://www.mit.edu/people/inkgray/net/). There are also companies that publish statistics on Web search en-

gines and the Web, such as Nielsen/NetRating (http://www.nelsen-netratings.com/), Commerce Net (http://www.cyberatlas.com), Iconocast (http://www.iconocast.com/), CyberAtlas (http://www.cyberatlas.com), and StatMarket (http://www.statmarket.com). There are institutions that survey Web users and provide the results to the academic community and others via the Web such as Georgia Tech (http://www.gvu.gatech.edu/user surveys/), Cyper Dialogue (http://www.cyberdialogue.com/free data/index.html), and NUA Publishing (http://www.nua.ie/surveys/). There are also governmental institutions that provide Web survey data, such as National Telecommunications and Information Administration (http://www.ntia.doc.gov/ntiahome/digitaldivide/).

Comparison of Traditional IR, OPAC, and Web-User Studies

Traditional IR, OPAC, and Web systems differ in terms of the interfaces, the search models, and the document collections (Borgman, 1996; Spark-Jones & Willett, 1997). However, do these differences result in different searching characteristics by users? Some researchers question whether the Web really is a unique searching environment worth of separate study. One way to address this question is to compare searching characteristics across all three types of systems. Studies using traditional IR and OPAC systems differ in terms of data reported and research design, so a rigorous comparison is difficult and perhaps impossible. However, these studies sometimes report similar searching metrics, such as the mean number of search terms per query or per session. One can utilize these similarities to sketch a picture of the "typical" search.

In addition to the three major Web studies, Hoelscher (1998); Jansen, Spink, and Saracevic (2000); and Silverstein et. al. (1999), we selected three studies of searching on traditional IR Hsieh-Yee (1993); Koenemann and Belkin (1996); and Siegfried, Bates and Wilde (1993) and three of searching on OPAC systems Millsap and Ferl (1993); Peters (1989); and Wallace (1993). We used the aggregate results to develop a general range of searching characteristics on each type of system. We attempted to select studies that were widely cited and offered a broad spectrum of searching characteristics. The studies, however, may not be representative of all studies in a given category.

Comparison of Web-Searching Study Results

To develop an overview of Web searches, the pertinent measures from each of the three major Web-searching studies are shown in Table 1. In cases where numbers, percentages, means or standard deviations were not provided for a given analysis, they were calculated and are displayed in Table 1. Due to varying definitions, some of the measures may not be exact comparisons.

From a comparison of these three Web-searching studies, one can conclude that the vast majority of Web searchers use approximately two terms in a query, have two queries per session, do not use complex query syntax, and typically view no more than 10 documents from the results list. Use of Boolean operators in Web queries is almost nonexistent, ranging from about 2% (Hoelscher, 1998) to 8% (Jansen et al., 2000). Surprisingly, given such simple searches, a survey of a major Web search engine's users reports that almost 70% of the searchers stated that they had located relevant information on the search engine (Spink, Bateman, & Jansen, 1999).

Searching Studies with Traditional IR Systems

The frequently cited studies chosen from traditional IR system research were Hsieh-Yee (1993), Koenemann and Belkin (1996), and Siegfried, Bates and Wilde (1993). The pertinent results from the studies are presented in Table 2.

Using only the data from novice searchers, one sees that query length ranged from about six to almost nine terms. Session length ranged from just over seven queries to approximately 16 queries. According to the Hsieh-Yee (1993) study, the use of advanced features was just under nine queries per session, and the number of documents viewed was about 10 documents per session. Using data from the Siegfried et al. (1993) study, it can be calculated that about 37% of the queries contained some type of Boolean operator, and about 1 in 17 (6%) queries contained some type of error.

Searching Studies with OPAC Systems

The prominent OPAC studies selected were Millsap and Ferl (1993), Peters (1989), and Wallace (1993). Table 3 illustrates the comparison of the data from these studies.

It is clear that these OPAC searchers used the typical OPAC system menu option choices of title, author, and keyword. The keyword search has the most in common with searches on traditional IR systems. The vast majority of these OPAC searchers viewed less than 50 documents. In both the Wallace (1993) and Millsap and Ferl (1993) studies, the failure rate ranged from 7 to 10% of the queries. The use of advance searching techniques was quite small, approximately 8%, and the use of Boolean operators in the Peters (1989) study was about 1%. From the Wallace (1993) study, it is evident that query length was less than two terms per query. According to data from the Millsap and Ferl (1993) study, session length was in the range of two to five queries per session.

Comparison of Searches

Using data from Tables 1, 2, and 3, one can develop broad picture of a typical search on each system. This data is presented in Table 4.

Focusing first on the similarities, one sees from Table 4 that the use of advanced features is about 8–9% for all three. The three searches were also similar in terms of the number

TABLE 1. Comparison of Web-user studies.

Category	Fireball study	Excite study	Alta Vista study
Period of data collection	31 days 1–31 July 98	Portion of 1 day 10 March 1997	43 days 2 Aug–13 Sept 98
Web IR system	Fireball search engine	Excite search engine	Alta Vista search engine
Document collection size at time of data collection (approx.)	3 million Web sites	30 to 50 million Web sites	100 million documents
Number of queries in data set	16,252,902	54,573	993,208,159
Session length (number of queries in session); sd = standard deviation	Not reported	Mean = 1.6, sd = 0.69 One: 67% (36,564) Two: 19% (10,391) Three: 7% (3,820) Four: 3% (1,637) Over Four: 4% (2,183)	Mean = 2.02, sd = 123.4* One: 77.6% (221,527,914) Two: 13.5% (38,539,006) Three: 4.4% (12,560,861) More Than Three: 4.5% (12,846,335) *the large sd may be due to softbots
Query length (number of terms in query); sd = standard deviation	Mean = 1.66 sd = 0.70 Zero: Not reported. One: 54.59% (8,873,001) Two: 30.80% (5,005,653) Three: 10.36% (1,683,129) More Than Three: 4% (691,119)	Mean = 2.21 sd = 1.05 Zero: 5.02% (2,584) One: 30.81% (15,854) Two: 31.46% (16,191) Three: 17.96% (9,242) More than three: 15% (8,186)	Mean = 2.35 sd = 1.74 Zero: 20.6% (204,600,881) One: 25.8% (256,247,705) Two: 26.0% (258,243,121) Three: 15.0% (148,981,224) More Than Three: 12.6% (125,144,228)
Use of Boolean (queries containing Boolean operators)	2.55% (414,461) *maximum possible number based on data provided	8.54% (4,661)	Not reported *see use of modifiers
Failure rate (improperly structured queries)	Not reported	10% (5,457)	Not reported
Use of modifiers (e.g., +, −, NEAR, etc.) (queries containing a modifier)	25.3% (4,111,843)	9% (4,776)	20.4% (202,614,464)* *Includes Boolean operators
Number of relevant documents viewed in a session	10 or Less: 59.51% (9,621,347) More than 10: 40.47 (6,545,887)	10 or less: 58% (31,652) More than 10: 42% (14,735)	10 or less: 85.2% More than 10: 14.8% *Numbers not reported and not calculable based on data provided.

of documents viewed. The Web and traditional IR system users viewed about 10 documents per session, whereas OPAC users tended to view more. At the query level, Web and OPAC searches had a similar number of terms per query, at about two terms.

Beyond these basic similarities, the other characteristics greatly diverge. In terms of failure rates, the users of the traditional IR systems made more mistakes (about 17%) relative to the searchers on the Web and OPAC systems, which had failure rates of about 10%. The differences among the searches were even more apparent in terms of session length, query length, and use of Boolean operators. Contrasting session lengths, the Web sessions were the shortest at about one to two queries. The OPAC system searches were next with two to five queries per session. The users of traditional IR system had the highest session length of seven to 16 queries. Users of the traditional IR systems had substantially longer queries, ranging from six to nine terms. Concerning the use of Boolean operators, the lowest use was by the OPAC systems searchers at 1%. The Web system searchers had a higher usage rate at about 8%. Finally, the searchers of the traditional IR systems had a

much higher usage of Boolean operators with approximately 37% of the queries containing Boolean operators.

Discussion of Comparison

There appear to be noticeable differences among searches on these three systems. To illustrate, take an example using the differences in sessions and query length. The typical Web search has a session length of about two queries and query lengths of about two terms. This means that the IR system has four terms to discern the information need of the user (i.e., two queries × two terms/query), assuming all terms are unique. Compare this to searches on the traditional IR systems where the session lengths were 7 to 16 queries, and the query lengths were of six to nine terms. Assuming a session length of seven and a query length of six , there are 42 terms over the entire length of the typical session (i.e., seven queries × six terms/query). Given that users of traditional IR systems typically use the building block approach to searching (Siegfried et al., 1993), one can reasonably assume that 66% of these terms are unique. This means that the IR system has 29 terms (i.e.,

TABLE 2. Comparison of three traditional IR-user studies.

Category	Koenemann & Belkin study	Hsieh-Yee study	Siegfried, Bates, & Wilde Study
Number of users and experience level	64 novice	30 novice and 32 experts	21 novice
Document collection utilized	74,520 articles from TREC	ERIC database	6 databases on humanities topics
IR system utilized	INQUERY	DIALOG	DIALOG
Session length (number of queries per user per session); sd = standard deviation	Mean = 7 Median = 8.2 *cannot determine sd from data provided	Not reported	Mean = 16.6 sd = 13.5
Query length (number of terms per query); sd = standard deviation	Mean = 6.4 sd = 4.2 *Terms in quotes counted as one term.	Mean for novice = 8.77 Mean for experts = 7.28	62.5% (2,563) of queries were one term 37.5% (1,538) of queries were two terms or more
Use of Boolean (number of queries containing Boolean operators)	Not reported	Not reported	36.8% (1,509) of queries contained one or more Boolean operator
Use of advanced features (number of queries containing advanced options)	Not reported	Mean for novice = 8.80 Mean for experts = 15.69	20.3% (832) of the queries contained one or more advanced feature *does not include use of Boolean operators
Failure rate (number of queries improperly formatted)	Not reported	Not reported	17% (697) of the queries contained a formatting error
Number of relevant documents viewed Per session	Not reported	Mean for novice = 10.31 Mean for experts = 28.72	Not reported

42 terms × 0.66) to discern the information need of the user, compared to the four terms for the Web system. For the system designer, this is critical design information. The higher use of Boolean operators and higher failure rates may also be important user characteristics for system designers. This comparison highlights the appearance of important differences among the typical search on the Web, traditional IR, and OPAC systems. More studies that specifically target the Web-user population need to be conducted.

Framework for Future Studies

To facilitate valid comparisons and contrasts among future Web-searching studies, we propose the use of a common framework. The framework consists of three sections, which are descriptive information, analysis presentation, and statistical analysis. Researchers utilizing common terminology and following a similar framework whenever possible will improve the ability to make comparisons among research results. Each of the three sections is discussed below.

Descriptive Information

The descriptive information section provides the necessary background data for one to evaluate the results and conclusions of studies. This section provides as much general information as possible on the searchers, the IR system,

the data collection, and the transaction log. Concerning the searchers, demographic information should address both the searchers from the transaction log and the general population of the IR system. One reports the number of searchers and visitors to the IR system in a given time period, primary language of the queries, language of the document collection, and the domains of the searchers.

This description information also includes the simple and advanced searching rules for the particular IR system. The rules discussed should be those in effect during the data collection period. Web IR systems routinely update their IR systems, and the rules in effect during the data collection period may not be the rules in effect when the results are published. The descriptive information on the document collection includes both the number of documents in the collection, the size (MB, GB, TB, etc.) of the document collection, and the average length of the documents. Other system information that needs to be provided is how the IR system handles indexing text, video, audio, images, and URLs.

In this section, one also addresses the format of the transaction log and how the data was collected. All transaction logs and logging systems are different, and the data collected may vary. Each field in the transaction log must be discussed and defined. The data format should be described and any assumptions that the researcher makes are presented and validated. Specific items to discuss are the method to identify the searchers, the time period of the logging process, and the format of the query.

TABLE 3. Comparison of three OPAC-user studies.

Category	Wallace study	Peters study	Millsap & Ferl study
Number of Searches	4,134 searches	13,258 searches	1,045 sessions
Session length (number of queries per user per session)	Not reported	Not reported	One or less: 32.8% (343) Two–five: 43.8% (458) More than five: 23.4% (245)
Query Length (number of terms per query)	Two or less Terms: 75% (3,101) More than two terms: 25% (1,034)	Not reported	Not reported
Number of relevant documents viewed per session	Less than 25: 82.1% (3,394) More than 25: 17.9% (740)	Not reported	1 to 50: 80.7% (843) More than 50: 19.3% (202)
Number of queries by keyword	53.1% (2,197)	31.9% (4,229)	23.9% (250) of sessions contained one or more queries of this type
Number of queries by title	24.2% (1,000)	34.2% (4,534)	62.2% (650) of sessions contained one or more queries of this type
Number of queries by author	21.7% (897)	23.2% (3,076)	38.1% (398) of sessions contained one or more queries of this type
Use of advanced features (number of queries containing advanced options)	8.7% (360)	2.8% (371)	Not reported
Use of Boolean (number of queries containing Boolean operators)	Not reported	1% (133)	9.2% (96) of the sessions contained one or more queries of this type
Failure rate (number of queries improperly formatted)	7% (289)	15.3% (2,028)	10% (105) of the sessions contained one or more improperly formatted query

Analysis Presentation

One difficulty in comparing user studies has been the variation in the definition of metrics and the use of various terms to identify the same metric. A common language would be of great benefit to better compare Web user studies, especially at the analysis level; therefore, the following levels of analysis and subsequent terms and definitions are offered. One should attempt analyses at three levels: the *session*, the *query*, and the *term*.

Session. The *session* is the entire sequence of queries entered by a searcher. The primary analysis at the session level is the number of queries per searcher. The researcher must specify which sessions (i.e., individuals, common users terminals, softbots) and which queries within sessions are being included or excluded. For example, if a searcher goes to the query page but does not enter a query, is that page access included in the session count? If the IR system generates a query to view results, is that query included? The inclusion or exclusion of certain types of sessions or queries will affect the analysis. The researcher also defines what qualifies as a *search* by an individual searcher. Some unique identifier in the transaction log typically denotes a searcher accessing the system. However, this technique is not full proof and the possible exceptions must be noted.

TABLE 4. Comparison of typical searches across three categories.

Category	Web systems searches	Traditional IR systems searches	OPAC systems searches
Session length (number of queries per user per session)	1–2	7–16	2–5
Query length (number of terms per query)	2	6–9	1–2
Number of relevant documents viewed per session	10 or less	Approximately 10	Less than 50
Use of advanced features (number of queries containing advanced options)	9%	9%	8%
Use of Boolean (number of queries containing Boolean operators)	8%	37%	1%
Failure rate (number of queries improperly formatted)	10%	17%	7–19%

Query. Sessions are composed of queries. A *query* is usually a string of zero of more characters entered into the Web IR system. This is a mechanical definition as opposed to the common information seeking definition (Korfhage, 1997). The first query in a session is referred to as an *initial query*. A subsequent query in a session that is identical to another query in the session is a *repeat query*. A subsequent query in a session that is different than any of the previous queries is a *modified query*. The difference between an initial, a repeat, and a modified query must be discussed and can include terms, capitalization, or term order. A *unique query* refers to a query that is different from all other queries in the transaction log. Of course, one can have various sub-components of these classifications.

At the *query* level of analysis, one is interested in determining *query length*, *query complexity*, and *failure rate*. *Query length* is measured in number of terms. *Query complexity* examines the query syntax. Query syntax includes the use of advanced searching techniques such as Boolean operators, phrased searching, and stemming. Many Web IR systems permit the use of symbols to accomplish the same effect as Boolean operators, such as +, −, !, " ", etc. These Boolean operators are referred to as *term modifiers*, which are also components of query syntax. The *failure rate* is defined as deviation from the published rules of the IR system.

Term. A *term* is defined as a string of characters separated by some delimiter such as a space, a colon, or a period. The researcher decides what delimiter to utilize and whether the system or searcher view is taken to define a term. For example, if a system rule requires terms to be separated by a blank space, searchers may mistakenly use other delimiters, such as a period. Should the study use the blank, the period, or both as the delimiter? The choice will affect the term count. We suggest using whatever delimiter is specified by the search engine. The use of query syntax that is not supported by a particular IR system is referred to as *carry over*.

One should also clearly state whether Boolean operators and term modifiers are counted as terms. There are advantages and disadvantages to including or excluding them. The advantage of removing Boolean operators is that the system-imposed operators are not included in the term count. In practice, however, it is sometimes difficult to determine what the searcher intended to be a Boolean operator and what was intend to be a conjunction.

Statistical Analysis

Statistical analysis includes at least the *mean*, the *standard deviation*, and the *median* wherever justified. Statistical analyses must be included if one is to compare and contrast results among studies. Because it is doubtful that a researcher will present all the statistical measures that fellow researchers desire, all the data should be presented with the lowest possible denominator. For example, in presenting the query length (i.e., the number of terms per query), it is better to list the number of queries and percentage with one term per query, two terms per query, etc., than to group the data (e.g., three or less terms per query) and present an aggregate number. Also, at the term level of analysis, the distribution of terms is compared to known distributions and measures, thus determining the goodness of fit.

Conclusion

This article presents an extensive review and analysis of current Web-searching studies. These studies have been published in an extremely diverse number of conference proceedings, journal publications, and Web sites, making it difficult to find these publications, as evidenced by the lack of co-citing among the Web studies. Only three of the studies cited other Web studies. This literature review and analysis will benefit any researcher conducting future Web-searching research.

We also compared traditional IR, OPAC, and Web-searching studies. This comparison shows that there appears to be differences in the manner Web users search versus the searching characteristics of users on traditional IR or OPAC systems. This comparison highlights that the Web is a unique searching environment that necessitates further and independent study. The nonstatistical manner of this comparison, with the noticeable variation and use of metrics among studies, also illustrates the need for uniformity among future searching studies to facilitate rigorous analysis.

Finally, we present a framework for the design and implementation of future Web-user studies. This consistency in levels of analysis and common metrics will allow for valid comparison of results. Valid comparison of results will lead to a better understanding of Web searches and to better design of future Web IR systems.

References

Abdulla, G., Liu, B., & Fox E. (1998). Searching the World-Wide Web: Implications from studying different user behavior. Paper presented at the World Conference of the World Wide Web, Internet, and Intranet, Orlando, FL.

Abdulla, G., Liu, B., Saad, R., & Fox, E. (1997). Characterizing World Wide Web queries. TR-97-04. Retrieved from the World Wide Web on 15 August 1999 from http://www.ncstrl.org/.

Alexa Insider Page. (2000). Alexa Insider Side Bar. Retrieved from the World Wide Web on 30 March 2000 from http://insider.alexa.com/insider?cli=10.

Belkin, N., Oddy, R., & Brooks, H. (1982). ASK for information retrieval: Part I. background and theory. Journal of Documentation, 38(2), 61–71.

Borgman, C. (1996). Why are online catalogs still hard to use? Journal of the American Society for Information Science, 47(7), 493–503.

Bush, V. (1945). As we may think. Atlantic Monthly, 176(1), 101–108.

Choo, C., Betlor, B., & Turnbull, D. (1998). A behavioral model of information seeking on the Web: Preliminary results of a study of how managers and IT specialists use the Web. Paper presented at the American Society of Information Science, Pittsburgh, PA.

CommerceNet/Nielsen Media. (1997). Search engines most popular method of surfing the Web. Retrieved from the World Wide Web on 12 August 1999 from http://www.commerce.net/news/.

Croft, W., Cook, R., & Wilder, D. (1995). Providing government information on the Internet: Experiences with THOMAS. Paper presented at Digital Libraries Conference, Austin, TX.

Crovella, M., & Bestavros, A. (1996). Self-similarity in World Wide Web traffic evidence and possible causes. Paper presented at ACM SIGMET-RICS Conference on Measurement & Modeling of Computer Systems, Philadelphia, PA.

Glaser, B., & Strauss, A. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago, IL: Aldine Publishing Co.

Goodrum, A., & Spink, A. (1999). Visual information seeking: A study of image queries on the World Wide Web. Paper presented at the Annual Meeting of the American Society for Information Science, Washington, DC.

Gordon, M., & Pathak, P. (1999). Finding information on the World Wide Web: The retrieval effectiveness of search engines. Information Processing and Management, 35(2), 141–180.

Harman, D. (1992). Overview of the second text retrieval conference (TREC-2). Retrieved 21 April 1999 from the World Wide Web from http://www.trec.nist.gov/pubs/tree2/papers/tx/01.txt.

Hawkins, D. (1996). Hunting, grazing, browsing: A model for online information retrieval. Online, Retrieved from the World Wide Web on 9 July 1997 from http://www.onlineinc.com/onlinemag/.

He, D., & Göker, A. (2000). Detecting session boundaries from Web user logs. Paper presented at 22nd Annual colloquium of IR research, April 5– 7, 2000, Cambridge UK.

Hoelscher, C. (1998). How Internet experts search for information on the Web. Paper presented at the World Conference of the World Wide Web, Internet, and Intranet, Orlando, FL.

Hoffman, D., Kalsbeek, W., & Novak, T. (1996). Internet and Web use in the U.S. Communications of the ACM, 39(12), 106–108.

Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. Journal of the American Society for Information Science, 44(3), 161–174.

Huberman, B., Pirolli, P., Pitkow, J., & Lukose, R. (1998). Strong regularities in World Wide Web surfing. Science, 280(5360), 95–97.

Jansen, B. J., Goodrum, A., Spink, A. (2000). Searching for multimedia: Video, audio, and image. Web queries. World Wide Web, 3(4).

Jansen, B., Spink, A., & Saracevic, T. (1999). The use of relevance feedback on the Web: Implications for Web IR system design. Paper presented at the World Conference of the World Wide Web, Internet, and Intranet, October 24–30, 1999. Waikiki, Honolulu. Available at http://jimjansen.tripod.com/.

Jansen, B., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: A study and analysis of user queries on the web. Information Processing and Management, 36(2), 207–227.

Jones, S., Cunningham, S., & McNab, R. (1998). An analysis of usage of a digital library. Proceeding of Second European Conference on Digital Libraries (pp. 261–277).

Jupiter Research. (1999). Go network announces new INFOSEEK Search: 30 percent faster, 50 percent larger. Retrieved from the World Wide Web on 5 September 2000 from http://info.go.com/press/search.html.

Kaske, N. (1993). Research methodologies and transaction log analysis: Issues, questions, and a proposed model. Library Hi Tech, 11(2), 79–86.

Kehoe, C., Pitkow, J., & Morton, K. (1999). Graphic, visualization, and usability center's 8th WWW user survey. Retrieved 15 August 1999 from the World Wide Web at http://www.gvu.gatech.edu/.

Keily, L. (1997). Improving resource discovery on the Internet: the user perspective. Proceedings of the 21st International Online Information Meeting (pp. 205–212).

Kirsch, S. (1998). The future of Internet search (keynote address). Paper presented at the 21st Annual International ACM SIGIR conference on research and development in information retrieval, Melbourne, Australia. Retrieved from the World Wide Web on 16 August 1999 from http://topgun.infoseek.com/stk/presentations/sigir.ppt.

Koenemann, J., & Belkin, N. (1996) A case for interaction: A study of interactive information retrieval behavior and effectiveness. Paper presented at conference on human factors in computing systems, Vancouver, Canada.

Korfhage, R. (1997). Information storage and retrieval. New York: Wiley.

Krishna, B., & Broder, A. (1998). A technique for measuring the relative size and overlap of public Web search engines. Proceedings of the 7th International World Wide Web Conference. Retrieved from the World Wide Web on 9 August 1999 from http://decweb.ethz.eh/WWW7/1937/.

Kurth, M. (1993). The limits & limitations of transaction log analysis. Library Hi Tech, 11(2), 98–86.

Lawrence, S., & Giles, C. (1998). Searching the World Wide Web. Science. 5360(28), 98–100.

Lawrence, S., & Giles, C. (1999). Accessibility of information on the web. Nature, 400, 107–109.

Lesk, M. (1997). Going digital. Scientific American, 276(3), 58–60.

Lesk, M., Cutting, D., Pedersen, J., Noreault, T., & Koll, M. (1997). Panel Session on "real world" information retrieval. Panel presented at 20th Annual international ACM SIGIR conference on research and development in information retrieval. Philadelphia, PA.

Lynch, C. (1997). Searching the Internet. Scientific American, 276(3), 52–56.

Millsap, L., & Ferl, T. (1993). Search patterns of remote users: An analysis of OPAC transaction logs. Information Technology and Libraries, 11(3), 321–343.

Mizzaro, S. (1997). Relevance: The whole history. Journal of the American Society of Information Science, 48(9), 810–832.

Nielsen/NetRating. (1999). Retrieved from the World Wide Web on 11 August 1999 from http://www.nielsen-netratings.com/.

NTIA. (1999). Defining the digital divide, the 3rd annual report by the National Telecommunications and Information Administration. Retrieved from the World Wide Web on 20 August 1999 from http://www.ntia.doc.gov/ntiahome/digitaldivide/.

Peters, T. (1989). When smart people fail: An analysis of the transaction log of an online public access catalog. Journal of Academic Librarianship, 15(6), 267–273.

Peters, T. (1993). The history & development of transaction log analysis. Library Hi Tech, 42(11), 41–66.

Pharo, N. (1999). Solving problems on the World Wide Web. Internet Research: Electronic Networking Applications and Policy, 4(3). Retrieve from the World Wide Web on 1 August 1999 from http://www.shef.ac.uk/~is/publications/.

Piktow, J. (1999). Summary of WWW characteristics. The World Wide Web Journal. 2(2), 2–13.

Robertson, S.E., & Hancock-Beaulieu, M.M. (1992). On evaluation of IR systems. Information Processing and Management, 28(4), 457–466.

Ross, N., & Wolfram, D. (2000). End user searching on the Internet: An analysis of term pair topics submitted to the Excite search engine. Journal of the American Society for Information Science, 51(10), 949–958.

Salton, G., & McGill, M. (1983). Introduction to Modern Information Retrieval. New York: McGraw-Hill.

Sandore, B. (1993). Applying the results of transaction log analysis. Library Hi Tech, 11(2), 87–97.

Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science. Journal of the American Society for Information Science, 26(6), 321–343.

Saracevic, T. (1995). Evaluation of evaluation in information retrieval. Paper presented at 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, WA.

Saracevic, T. (1996). Modeling interaction in information retrieval: A review and proposal. Proceedings of the Annual Academy Meeting of American Society for Information Science (pp. 35–44).

Saracevic, T., Kantor. P., Chamis, A., & Trivison, D. (1988). A study of information seeking and retrieving. I. Background and methodology. Journal of the American Society for Information Science, 39(3), 161–176.

Selberg, E., & Etzioni, O. (1995). Multi-service search and comparison using the Metacrawler. Proceedings of the Fourth World Wide Web Conference. Retrieved from the World Wide Web on 5 August 1999 from http://www.cern.ch/CERN/WorldWideWeb/Papers.html.

Siegfried, S., Bates, M., & Wilde, D. (1993). A profile of end-user searching behavior by humanities scholars: The Getty online searching project report no. 2. Journal of the American Society for Information Science, 44(5), 273–291.

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web search engine query log. SIGIR Forum, 33(1), 6–12.

Smith, T., Ruocco, A., & Jansen, B. (1998). Digital video in education. Proceedings of the Thirtieth ACM SIGCSE Technical Symposium on Computer Science Education (pp. 122–126).

Sparck-Jones, K., & Willett, P. (Eds.). (1997). Readings in information retrieval. San Francisco: Morgan Kaufman.

Sparck-Jones, K. (Ed.). (1981). Information Retrieval Experiments. London: Butterworth.

Spink, A., Greisdorf, H., & Bateman, J. (1998). From highly relevant to not relevant: Examining different regions of relevance. Information Processing and Management, 34(5), 599–621.

Spink, A., Bateman, J., & Jansen, B. (1999). Searching the Web: A survey of Excite users. Internet Research: Electronic Networking Applications and Policy. Retrieve from the World Wide Web on 7 August 1999 from http://www.shef.ac.uk/~is/publications/infres/ircont.html.

Statistical Research, Inc. (2000). New study shows Internet users are loyal to Web "Niches." Retrieved from the World Wide Web on 5 September 2000 from http://www.sriresearch.com/press/pr20000217.htm.

Sullivan, D. (2000). Search engine sizes. SearchEngineWatch.com. Retrieved from the World Wide Web on 29 April 2000 from http://www.searchenginewatch.com/reports/sizes.html.

Wallace, P. (1993). How do patrons search the online catalog when no one's looking? Transaction log analysis and implications for bibliographic instruction and system design. RQ, 33(3), 239–252.

Wolfram, D. (1999). Term co-occurrence in Internet search engine queries: An analysis of the Excite data set. Canadian Journal of Information and Library Science, 24(2/3), 12–33.

Xu, J.L. (1999). Internet search engines: Real world IR issues and challenges. Paper presented at the Conference on Information and Knowledge Management. Kansas City, MO.

Xu, J.L. (2000). Multilingual search on the World Wide Web. Paper presented at the HICSS 33. Maui, HI.

Zorn, P., Emanoil, M., & Marshall, L. (1996). Advanced searching: Tricks of the trade. Online. Retrieved from the World Wide Web on 1 July 1997 from http://www.onlineinc.com/onlinema/.

Zumalt, J., & Pasicznyuk, R. (1998). The Internet and reference services: A real-world test of Internet utility. Reference and User Services Quarterly, 38(2), 165–172.