Searching heterogeneous collections on the Web: behaviour of Excite users.

[Amanda Spink](#) & Judy Bateman
 University of North Texas, Denton, Texas, USA
 and
 Major Bernard. J. Jansen
 United States Military Academy

Abstract

As Web search services become a major source of information for a growing number of people, we need to know more about how users search heterogeneous collections using Web search engines. This paper reports the results from a major study exploring users' information searching behaviour on the EXCITE Web search engine. Three hundred and fifty-seven (357) EXCITE users responded to an interactive survey, including their search topics, intended query terms, search frequency for information on their topic, and demographic data. Results show that: users tend to employ simple search strategies, and conduct successive searches over time to find information related to a particular topic. Implications for the design of Web search services are discussed.

Introduction

The Web is a heterogeneous collection of information resources with minimal selection, organization, and retrieval standards. In particular, there is wide variation in the access capabilities of Web search engines that try to bridge large heterogeneous collections. The majority of Web search services that use search engines as the access mechanism to information resources can be approximated to be on the broader end of the access mechanisms to digital libraries and information retrieval (IR) systems. They utilize IR techniques, (e.g., Boolean queries and relevance ranking), that are also widely used by digital libraries. In the broadest sense, digital libraries and IR systems are part of the Web. A growing body of research is investigating user interaction with digital libraries and Web search services, (e.g., EXCITE). The study of uses and users of Web search-engines can also be compared to the uses and users of digital libraries, to test if users exhibit similar behaviour on both types of heterogeneous digital collections. User behaviour common to IR systems can also be investigated with Web users, i.e., users' successive searches in relation to the same or evolving information problem.

Successive Searching Research

Recent research in the information retrieval (IR) context shows that users with a problem-at-hand often seek information in stages over extended periods and use a variety of information resources ([Spink, 1996](#)). As time progresses, users tend to search the same or different interactive systems (digital libraries, IR systems, Web services) for answers to the same or evolving problem-at-hand ([Bateman, 1997](#)). The process of repeated, successive searching over time (including changes or shifts in beliefs, and cognitive, affective, and situational states), is called *the successive search phenomenon.* How access to heterogeneous collections on the Web can be designed to assist users in various ways in their successive searches is an important research question. Users' successive searching currently receives little, if any, support from present interfaces, procedures, or search-engines. By and large, interactive systems are built following a single search paradigm, i.e., they are designed and operate on the assumption that every search is an end in itself. The study reported in this paper is part of a new and growing line of inquiry addressing the successive search phenomenon and associated episodes. The aim of the study reported in this paper is to explore users' characteristics, searching behaviour, and successive searching when using the EXCITE search-engine. Users of the Web search service EXCITE were asked to complete an interactive survey form about the nature of their interaction with EXCITE, including their current search topic, search terms,

information seeking stage, and frequency of searches on EXCITE on their current topic. The survey results are supplemented with preliminary findings from a separate study of 18,113 EXCITE users and their 51,472 queries (Jansen *et al.,*1998).

The research is significant, since, as the size of the Web grows exponentially and the variety of information resources on the Web diversify rapidly, the problem of searching heterogeneous collections becomes critical. It certainly is fast becoming, if not already, *the* problem for a majority of end-users. When the design of digital libraries, IR systems and various search-engines is driven by technological criteria and technology-related algorithms, they are found lacking in many respects when encountered, used, and evaluated by users. The research reported here is oriented toward deriving human dimensions and criteria for the design of IR interfaces and search-engines.

Related studies

Web Searching

The phenomenal growth in the size of the Web has created a growing body of empirical research investigating many aspects of user interactions with the Web. User-oriented Web research generally includes experimental and comparative studies, user surveys, and user traffic studies (Crovella & Azer, 1996). Experimental and comparative studies show little overlap in the results retrieved by different search-engines based on the same queries (Ding & Marchionini, 1996), and many differences in search-engine features and performance (Chu & Rosenthal, 1996). Surveys of Web users are generally library based (Tillotson,*et al.,* 1995) or distributed by submission to newsgroups (Perry, 1995). Pitknow and Kehoe (1996) found major shifts in the characteristics of Web users over four surveys, including a growing diversity of Web users based on age, gender, and access through both the office and home computers. This paper reports results from a survey conducted directly through a major commercial search-engine to investigate users' searching behaviour.

Successive Searching behaviour

Recent IR studies suggest that successive searches may be a fundamental aspect of users' behaviour when seeking information related an information problem. Humans seek information in stages over extended periods as their information problem changes (Kuhlthau, 1993) and use different types of IR systems during an information seeking process (i.e., Web, CD-ROMs, etc.). IR system users (Saracevic,*et al.,* 1991), end-users (Huang, 1992), and OPAC users (Robertson & Hancock-Beaulieu, 1992) conduct successive IR searches when seeking information related to a particular information problem. Robertson and Hancock-Beaulieu, (1992) found a continuity of search topics and relevance judgments by the same OPAC users over successive searches. Some users explored a topic over an extended period and interacted at intervals with the on-line catalogue OKAPI, using identical or closely related search strategies. Spink (1996) found that for 200 IR system users: 56% had conducted more than one IR search, 21% had conducted five or more IR searches, and many users had conducted successive searches at different stages of their information seeking process on a particular topic. At present, limited knowledge exists on users' searching behaviour and the extent of successive search behaviour by Web and digital library users.

The modeling of users in successive searches is then *successive user modeling.* A key dimension is time, and the key variable is changes or shifts in successive search episodes over time. The key constant is the same or evolving information problem. The evolution, if any, of a problem and other cognitive, affective and situational variables can be mapped, and the history of successive search episodes can be recorded and analyzed, i.e., the phenomenon can be a subject of research. The successive search phenomenon is just beginning to be investigated to any extent by digital library, IR or Web researchers.

Research questions

The objective this study was to gather data on the use of a major Web search-engine to provide a preliminary model of user characteristics and search behaviour. Specifically, data was collected on users': (1) demographic characteristics, (2) search topics, (3) search terms and queries, and (4) successive search behaviour. Limitations of this study include the small sample size, the exclusive use of an interactive survey form and the dependence on users' self reported behaviours. Richer data can be obtained from analysis of users' search logs and observation of

their searching behaviour. An analysis of users' actual search queries is currently underway, and some preliminary results are also reported here.

Research design

Data Collection and Analysis

Data were gathered through an interactive eighteen-question survey developed by the researchers in conjunction with the staff at EXCITE, Inc. (see Appendix A). The interactive survey was made available through EXCITE's home page for five days from Friday April 11 to Tuesday April 15,1997. Only those EXCITE users who accessed EXCITE's home page directly (*http://www.EXCITE.com*) could access the survey form. Users who accessed the EXCITE search-engine indirectly through their web-browser search capability could not access the survey form. After completing the survey, users were asked to click on the "Send Survey" button. The total number of http requests of the survey site during the five day period was 11,187 (approximately 3729 visitors). Four hundred and eighty (480) users clicked the 'Send Survey" button at the end of the survey form. From 10am to 2pm on Saturday April 12 was the period of heaviest usage of the survey form. The numerical survey data were transferred into the ACCESS statistical package for further analysis. Despite some pretesting of the survey form, technical difficulties resulted in the corruption of data from five questions during the data collection phase. The raw data results from the remaining questions were plotted into basic data tables. In twelve questions, users selected one answer from a number of options; in two questions, users chose either "Yes" or 'No"; in one question either "Male" or "Female", and in three questions, users described their search topic, listed their proposed search terms and provided comments on their search or on the survey. The results from the last three questions were analyzed qualitatively and the responses categorized.

Results

The results are reported in four sections: (1) demographic data, (2) search topics, (3) search terms and queries, and (4) successive searching. Only 316 of the 480 returned survey forms contained usable data. One respondent returned fifty blank survey forms in a row. Some respondents did not provide answers to each survey question. We now outline the demographic profile of the respondents to identify the population characteristics.

Demographic Characteristics

Age

Users ranged in age from less than 10 years to over 60 years, with the majority between the age of 20 and 50 years (Table 1).

Age (Years)	Number	Percentage
< 10	4	1
11-20	48	16
21-30	69	23
31-40	58	19
41-50	62	21
51-60	36	12
61+	24	8
Total	301	100

Table 1: Age of survey respondents

Education Level

Most respondents were either high school or college graduates (Table 2).

Education level	Number	Percentage
High School	57	19
Vocational	24	8
Some college	71	24
Bachelors	70	24
Masters	46	16
Professional	14	5
Ph.D.	9	3
Student	2	1
Total	293	100

Table 2: Educational level of respondents

Occupation

Students and professionals formed the largest group of respondents, followed by executives and the self employed (Table 3). Overall, many respondents were from business or academic related environments. It is not surprising that the college crowd formed a large group of respondents.

Occupation	Number	Percentage
Student	55	19
Professional	38	13
Executive	29	10
Self-employed	25	9
Technical	16	6
Faculty/academic	14	5
Consulting	14	5
Services	14	5
Research & development	11	4
Clerical	10	4
Marketing	8	3
Other	50	18
Total	284	100

Computer Domain

Interestingly, the largest group of respondents were searching EXCITE from home - followed by commercial and educational users. However, we don't know how many respondents were searching both at home and at work.

Computing domain	Number	Percentage
Personal	107	36
Commercial	83	28
Educational	55	18
Organizational	22	7
Government	7	2
Military	2	1
Other	23	8
Total	299	100

Table 4: Computing domain of respondents

Geographic Location

The overwhelming number of respondents were located in the United States (Table 5). This finding was not unexpected and reflects the current concentration of Web searching in the U.S. The survey was also only available in English, which may have restricted the user sample further.

Geographic location	Number	Percentage
North America	225	84
Western Europe	8	3
United Kingdom	5	2
South America	6	2
South East Asia	4	1
Middle East	3	1
South Asia/India	3	1
Central America	3	1
Australia	2	1
Japan	2	1
Eastern Europe	2	1
Korea	1	1
China	1	1

Other	1	1
Total	269	100

Table 5: Geographic location of respondents

Computer Platform

Most respondents accessed EXCITE from an IBM/PC or equivalent platform (Table 6).

Computer platform	Number	Percentage
PC/IBM	201	68
X/UNIX	38	13
Macintosh	35	12
VMS	4	1
Line-mode	2	1
Next Step	2	1
Other	14	4
Total	296	100

Table 6: Computer platform used by respondents

Search topics

Users were asked to describe their current search topic. Respondents current search topics on EXCITE were dispersed broadly over 16 search topic categories.

1. *Individual or family information*: biographical or information about an individual or family within three sub-categories.
   - *Family or friend*: email or mailing addresses for an individual, often a relative, e.g., (User 96) "looking for my brother, so I can e-mail him". Some searchers used the individual's name as a query term, and one used a city and state as a query term.
   - Genealogical information and were placed in the sub-category *genealogy*, e.g., (User 62) "genealogy research of my family tree".
   - Public figures, either historical (Albert Einstein) or current (Bill Cosby) were placed in the sub-category *public figure*, including individuals from sports, science, entertainment, etc.
2. *Computers:* computer hardware, software, information about the Internet or world wide web and computer games, e.g., (User 12) "looking for information on Hewlett Packard printer drivers for the Deskjet 660c" and (User 57) "looking for company that sells older models of hp hardware".
3. *Medical:* diseases and disabilities, health related products, special diets and nutrition and health care, e.g., (User 166) "Crohn's disease".
4. *Education:* K-12 and college and university information, including lesson plans, home schooling, and searches for colleges and scholarship information, e.g., (User 41) "trying to find out universities in USA that give long distance classes through Internet".
5. *Business:* individual businesses and industries, agriculture, information about CEOs, real estate etc. Searches for stock quotes were put into the *news* category, e.g., (User 21) "looking for information about Walter Elisha, CEO of Spring Industries" and (User 108) "used to gather railroad related information, Amtrak and commuter rail companies that Amtrak operates.
6. *Science*: science, technology, and psychology were included, e.g., (User 119) "endangered species".
7. *Politics and government information*: government and legal information at the city, state, and federal levels - census and demographic information, government agencies and departments including libraries and detention

centers, e.g., (User 224)"social security data of individual salary history" and (User 100) "Allegheny Health Dept. in Pittsburgh PA".

8. *Shopping and price information*: retail products and searches for replacement parts and a large number of searches for automobile price information, e.g., (User 207) "looking to find the current value of my car" and (162) "shopping for TVs, stores in Jacksonville".
9. *Hobbies:* pets, cooking, gardening, home repairs, crafts, hobbies and wedding planning, e.g., (User 151) "looking for craft ideas",
10. *Graphic images*: computerized images and computerized greeting cards where the image rather than the information contained in the image (for example maps) was the search topic. Searches for maps were either put in travel or in government information, e.g., (User 4) "greeting card to use as email".
11. *General information:* searches where the respondent stated he or she was browsing, not looking for a particular topic, looking for general information or reference sources, or just looking at things that interested them. Searches for encyclopedias and dictionaries were included in this category, (User 88) "anything and everything" and (User 155) "just going to look around."
12. *Entertainment:* sports, television shows, movies, bands, humor etc., e.g., (User 134) "television shows".
13. *News:* current events, stock quotes, weather, lottery numbers, and horoscope information, e.g., (User 79) "JonBenet Ramsey case" and (User 74) "world weather"
14. *Travel:* hotels, resorts and tourist information either in general or about a particular city or locale, e.g., (User 274) "hotels and prices" and (User 271) "looking for all campgrounds and RV parks along Minnesota state highway I-94".
15. *Employment*: jobs by geographic location or by type of job or career and searches by both employers and job seekers, e.g., (User 66) "information science related employment".
16. *Arts and humanities:* religion, history, art etc., e.g., (User 311) Ancient Roman Mythology/Architecture/gods/etc."

In some cases, respondents ranged over several topic categories as the information provided by respondents made it difficult to determine exactly the situational context in which the information was to be used. In these cases the search was placed in the category that seemed to best fit the topic described by the user.

Search Topic Frequency

Table 7 lists the frequency of search within the 16 search-topic categories. Search topics were dispersed over a broad of general and specific subjects, similar to public library reference questions. The major topics of EXCITE searches were for information about people, companies and products.

Category	Number	Percentage
Individual or family		
Family or friend	17	6
Public figure	11	4
Genealogy	6	2
Sub-total	34	12
Computers	34	12
Business	30	10
Entertainment	23	8
Medical	22	8
Politics & government	20	7
News	19	7

Hobbies	18	6
General information or surfing the Web	16	6
Science	16	6
Travel	13	5
Arts & humanities	12	4
Education	10	3
Shopping	8	3
Graphic images	7	2
Employment	5	1
Total	287	100

Table 7: Frequency of search topics

Most respondents searched on a single topic as determined by their query terms and search topic statements. Eleven respondents reported searching on two different topics and two respondents reported searching on three topics. Multiple search topics were determined by an analysis of the query terms and search topic statements. The topics for respondents who reported browsing or surfing, or as one respondent put it "whatever interests me", were categorized as *general information or surfing* searches.

Search terms and queries

Table 8 provides a more detailed overview of the search terms reported by respondents. These were the terms that the respondents as those they *intended* to use, not those actually used. The mean number of search terms was relatively low at 3.34. Some respondents seemed confused about what they were to report when asked to list query terms for their search. Some respondents reported links instead of query terms and six respondents used the query term area to describe their search. One respondent put question marks in the query term area.

Classification	Number of search terms
Total number of respondents who reported terms	210
Total terms (did not include stop words)	701
Mean number of terms/respondent	3.34
Two-term phrases	84
Three-term phrases	9
Proper nouns (personal & place names, companies, etc.)	45
Links	9
Described search	6
URL	1

Table 8: Search term data

Boolean Operators

EXCITE allows searching for phrases, Boolean operators (AND, OR, and AND NOT), and uses parentheses to group search terms and Boolean operators. Many respondents included terms that they clearly meant as a phrase or proper name, but no respondent indicated that they would use quotes (EXCITE'S method of indicating that two or more words should be next to each other) around these phrases. EXCITE also allows the user to mark words with a "+" (plus) to indicate that the retrieved information must contain this word. A "-" (minus) is used to indicate that the retrieved information must not contain that word. Terms are searched as a phrase only when the phrase is enclosed in quotation marks i.e. "endangered species". If a phrase is entered without the quotation marks terms will be connected by the Boolean OR operator, i.e. endangered species without quotation marks will result in a query of endangered OR species. Some respondents reported the format and syntax of their search query in addition to the search terms they planned to use. Few queries included Boolean or other operators. Of the ones that did: (1) four queries included AND, (2) two queries included OR, and (3)eleven queries included +. One respondent used both AND and OR and parenthesis in their search query. This respondent also attempted to truncate using an asterisk (*). EXCITE does not use an asterisk as a truncation operator so the query would retrieve information that contained the word stem followed by an asterisk, i.e., librar* would retrieve only librar* and not library or libraries. EXCITE help facilities do not mention a truncation operator.

An additional seven respondents used the word "and" in a manner that indicated they were intending it as the Boolean AND operator. EXCITE requires that AND be capitalized to be considered a Boolean operator, otherwise it will be treated as a stop word. Respondents used both "and" and AND to connect words that they seemed to think would be automatically searched as phrases. Without the quotation marks each term in the phrase is automatically combined with an implicit Boolean OR. Some respondents used the "+" (plus) sign instead of the Boolean AND. Since a "+" (plus) is used to indicate that the retrieved information must contain this word it can be used in place of the Boolean AND operator. However, the initial term must also be preceded with a "+" (plus) for the query to have the same results as an AND operator. Five (5) respondents used the "+" (plus) correctly and placed it in front of the desired word with no space between the "+" (plus) and the word. Two respondents incorrectly added a space. Twenty-four (9%) respondents used Boolean operators, "+" (plus) signs or "and" in a manner that indicated that they expected it to be a Boolean operator. Ten respondents used the correct syntax for EXCITE in their search queries. No respondent used a "-" (minus), quotation marks, or the Boolean operator AND NOT.

Few users employed Boolean operators and even fewer users applied the correct syntax to enter search phrases and Boolean operators. The user search logs confirm this low use of Boolean operators, with only 2694 (5.24%) of queries containing Boolean operators. EXCITE uses the Boolean OR as a default operator that can result in searches that are less specific than the user intended and an increase in the search's retrieval. EXCITE ranks and posts retrieved information by relevance ranking and this may help compensate for incorrect search query syntax. However, when systems calculate relevance rankings usually both proximity and frequency of terms are considered. The user who thinks he or she is searching a phrase by simply entering the terms into the search statement in phrase order may obtain results that have high relevance rankings but do not relate well to the user's intended search query.

Successive searching behaviour

Frequency of EXCITE Searching

Users were first asked how frequently they searched EXCITE for information in general. Many respondents reported searching EXCITE on a daily basis to find information, and nearly a third of respondents also searching EXCITE weekly or at least 2-3 time per week (Table 9).

Frequency of searches	Number of users	Percentage of users
First search	51	17
Daily	124	42
Two to three searches	57	20
Weekly	45	15
Monthly	19	7

Total	292	100

Table 9: Frequency of EXCITE Searching

Users were then asked to estimate the number of EXCITE searches they had conducted on their current topic.

Number of EXCITE Searches on Current Topic

As Table 1o shows, one third of respondents were first-time users, conducting their first search of EXCITE on their current topic; two-thirds reported a pattern of successive searches of between one to five EXCITE searches on their current topic; thirty percent reported more than five EXCITE searches on their topic; and thirty-eight reported conducting more than twenty searches on their topic. By user estimates, we find that most users are repeatedly searching EXCITE for information on the same or evolving topic.

No. of EXCITE searches	Number of users	Percentage of users
First search	112	39
Two to five	88	31
Six to ten	32	11
Eleven to fifteen	10	3
Sixteen to twenty	9	3
More than twenty	37	13
Total	288	100

Table 10: Number of EXCITE searches on current topic
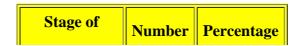
Relevant Retrieval

Users where then asked if they had retrieved any relevant information from EXCITE on their current topic. Most users reported retrieving relevant information from EXCITE on their current topic (Table 11).

Retrieval status	Number	Percentage
Yes	206	72
No	80	28
Total	286	100

Table 11: Users' retrieval of relevant information from EXCITE on current topic.

Respondents' Information Seeking Stages

Respondents were then asked to estimate their current information seeking stage related to their current search topic. Different EXCITE respondents were at different stages of their information seeking process related to their current search topic (Table 12). Most respondents reported that they were: (1) still gathering information on their topic (50%), and (2) conducting successive searches of EXCITE or frequently searching for information over time during an information seeking process related to a specific search topic (61%).

Stage of	Number	Percentage

information gathering	of users	of users
Beginning	112	39
Still gathering	141	50
Completing	31	11
Total	284	100

Table 12: Stage of users' information gathering on their current topic.

Number of Searches by Information Seeking Stage

Constructed from a combination of Table 10 and Table 12, the matrix Table 13 shows that many users were conducting successive searches when seeking information on a particular search topic.

EXCITE searches	Beginning stage		Still gathering		Completing	
First search	63	23%	35	13%	7	3%
Two to five searches	23	8%	48	18%	11	4%
Six to ten searches	8	3%	16	6%	7	2%
Eleven to fifteen searches	2	1%	6	2%	2	1%
Fifteen to twenty searches	2	1%	4	1%	2	1%
More than twenty searches	6	2%	26	10%	4	1%
Total (272 users)	104	38%	135	50%	33	12%

Table 13: Matrix of information gathering stage by number of EXCITE searches

The largest group of EXCITE respondents (23%) were conducting their first search at the beginning of their information seeking process on their current topic. Twenty-six (10%) users also reported still gathering information after more than 20 EXCITE searches. The largest group of respondents had conducted from one to five searches, many at the beginning and still gathering stages of their information seeking process.

Changes in Search Terms

Fifty four percent (54%) of *successive search users* reported changing their search terms on their current topic over successive searches (Table 14). However, the other half of successive searchers reported "still gathering" or "completing" with no change in their search terms over successive searches. This finding was not surprising, as previous studies by Robertson and Hancock-Beaulieu (1993) and Spink (1996) reported similar findings with IR system, CD-ROM and On-line Public Access Catalogue (OPAC) users.

Status	Number	Percentage
Yes	138	54
No	119	46
Total	257	100

Table 14: Change in users' search

Successive searching involves changes and shifts in search terms, search strategies, relevance judgments and criteria, or in information problem focus. Those respondents who had conducted successive searches were asked if their search terms had changed over successive searches. The study did provide a rich set of data and some surprising findings that are discussed in the next section of the paper.

Discusson

The results of the study revealed a number of interesting findings. EXCITE users are a diverse group of peple. Not only do they span most age groups, but also different educational and occupational backgrounds ranging from academia to business. They seem to prefer to access the Web via IBM PCs and are mainly based in North America. Respondents' search topics varied immensely, from entertainment to business and computing. The topics were similar to reference queries that might be made to a reference librarian in a public library. The lack of sexually motivated search topics and terms was rather surprising. This was probably due to self-censorship on the part of the respondents in completing a survey form. Jansen, *et al.* (1998) found sex to be the most frequent search topic during an analysis of over 51,474 EXCITE search queries. These queries were from over 18,113 EXCITE users.

We can also see that respondents were not proposing to use many search terms or employ complex search strategies. Nor were they planning to use many search features, such as Boolean operators, query modifiers or natural language queries. This finding implies a fairly low level of interaction with the EXCITE web search-engine. This finding does not account for respondents' actual behaviour once they began to interact with EXCITE, but it does give some insight into their search preparation and initial search terms and strategies. A number of respondents indicated that they were conducting successive searches on their topic. One can speculate that the sheer magnitude of any retrieval in response to a few search terms may cause users to quickly peruse the results, log off, possibly rethink or search another information resource, and then use EXCITE once again. Jansen *et al.* (1998) found that EXCITE users performed limited query reformulation and had little persistence in viewing retrieved lists of Web sites. Overall, the users' ability to specify good search terms and create complex search queries to clearly and precisely capture relevant retrieval seems rather low. Users also appear to lack the motivation to employ complex search strategies and learn correct syntax and rules, and may expect the search-engine to automatically create effective queries.

Conclusions

The findings of this study indicate areas for consideration in the design of Web search services. One of the chief implications of the findings is the need for Web search services to allow users to save their search terms, strategies and results for further reformulation. Many searching tasks are not clear to users when web searching. Search term and strategy selection tools might also help Web users, particularly those in successive search mode. An additional aid could be a pre-processing of a user's query checking for lower case "and", spaces after "+", etc. The user could be prompted to possible syntax and spelling errors. through the user interface. The development of interactive tutorials for Web users might also help them to learn the basics of effective searching.

Users have the option to engage in fairly complex processes with search-engines and engage the full functionality of these systems to improve their retrieval results. However, most searches are short and simple. This paper has identified a crucial problem for search-engine designers - the lack of transparency of both the nature and benefits of basic and advanced search features for the large mass of users who frequently interaction with heterogeneous digital collections. Users are currently also engaging in searching behaviours, such as successive searching, that are not supported by search-engines and techniques. This study also extends previous research by Spink (1996) to show the general practice of successive searching by users of interactive IR systems. The key area for further research is to model the changes and shifts that occur within and between successive searches on heterogeneous digital collections.

Acknowledgement

References

- Bateman, J. (1997). *Changes in relevance criteria: An information seeking study.* Unpublished Ph.D. Proposal. University of North Texas, School of Library and Information Sciences.
- Chu, H., & Rosenthal, M. (1996). search-engines for the World Wide Web: A comparative study and evaluation methodology. *Proceedings of the 59th Annual Meeting of the American Society for Information Science, Baltimore, M.D.,* 127-135
- Crovella, M. E., & Azer, B. (1996). Self similarity in World Wide Web traffic evidence and possible causes. *Proceedings of ACM SIGMETRICS, Philadelphia, P.A., May 1996*, 126-137
- Ding, W., & Marchionini, G. (1996). A comparative study of web search service performance. *Proceedings of the 59th Annual Meeting of the American Society for Information Science, Baltimore, M.D.*, 136-142
- Huang, M. H. (1992). *Pausing behaviour of end-users in on-line searching*. Unpublished Ph.D. dissertation. University of Maryland.
- Jansen, B. J., Spink, A., Bateman, J., & Saracevic, T. (1998) *What do they search for on the Web and how are they searching: A study of a large sample of EXCITE searches*. Unpublished paper. School of Library and Information Sciences, University of North Texas.
- Kuhlthau, C. C. (1993). *Seeking meaning: A process approach to library and information services.* Norwood, NJ: Ablex Press.
- Perry, C. (1995). Travelers on the Internet: A survey of internet users. *Online, March/April*, 29-34.
- Pitknow, J. E., & Kehoe, C. M. (1996). Emerging trends in the WWW user population. *Communications of the ACM, 39*(6),106-108.
- Robertson, S. J., & Hancock-Beaulieu, M. (1992). On the evaluation of IR systems. *Information Processing and Management, 28*(4), 457-466.
- Saracevic, T., Mokros, H., Su, L., & Spink, A. (1991). Interaction between users and intermediaries in on-line searching. *Proceedings of the 12th National On-line Meeting, 12, May 1991, New York*, 329-341
- Spink, A. (1996). A multiple search session model of end-user behaviour: An exploratory study. *Journal of the American Society for Information Science, 47*(8), 603-609.
- Tillotson, J., Cherry, J., & Clinton, M. (1995). Internet use through the University of Toronto: Demographics, destinations and users' reactions. *Information Technology and Libraries, September*, 190-198.

How to cite this paper:

Check for citations, using Google Scholar

Contents   10816
Web Counter