Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

End-user Searching of Web Resources: Subject Needs and Zero-hits

Peiling Wang and Line Pouchard

School of Information Sciences
The University of Tennessee, Knoxville

ABSTRACT

This study analyzed a log file capturing users' queries executed in the Web site of the University of Tennessee, Knoxville during March, 1997. The purpose of the study is three-fold: to understand what information needs the users of this Web site have, to investigate how successful these end-users are in searching for information, and to identify problems related to unsuccessful queries. Content analysis of each query focused on the type of information needs and the type of errors that caused a zero-hit result. Fifteen classes of information needs are identified based on content analysis of the queries; the most frequently occurred queries are searching for institutional unit and searching for academic information (counting for 40.0% of the total queries searched). The unsuccessful queries is more than 33.5% measured by zero-hit outcomes. Two types of errors that caused zero-hit are identified: syntactic and semantic errors. Syntactic errors occurred more often than semantic errors (53.6% vs. 46.4%). The findings suggest that end-users of Web resources need guidance and help in performing searches. Syntactic errors may be corrected by the search engine automatically, while semantic errors need a better information representation scheme from the Web site.

INTRODUCTION

The rapid development of Internet technology and World Wide Web (Web) resources has promoted end-user online searching greatly. Any user connected to the Internet can have access to information available on the Web. The number of users of Web resources is expected to increase as more and more users are connected to the Internet. Compared to end-users performing searches in libraries and information centers, where user instructions, handouts, and human librarians are readily available and databases are homogeneous, the users of the Web resources perform searches on their own. It is hard to know what they need from the Web resources and how successful they are in searching for information. The majority of Web users received no training and are unlikely to learn the Web systematically or to request assistance from a professional searcher. The heterogeneity in resources and diversity in search tools also complicate the search. A research question inevitably arises: what are the information needs of these Web users and are they able to find the needed information?

The purpose of this study is to understand the subject needs of the users of a university Web page and to investigate whether users are successful in searching for information, and to identify problems related to unsuccessful queries. The ultimate goal of this line of research is to suggest principles for designing a better system that can incorporate users' needs and search behavior.

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

This study obtained a log file for the Web site of the University with all the queries entered by the users and search results (postings) during March 1997. The content analysis of the queries focused on the following research questions:

1. What types of subject interests for this user group are reflected in the queries?
2. Can a taxonomy of subjects be developed from user queries? Can this taxonomy form the basis of a practical classification for organizing Web spaces?
3. What are the unsuccessful queries that result in a zero-hit? What types of errors do these queries contain?
4. How should the system be improved to enhance its utility?

LITERATURE REVIEW

End-user online searching has gone through several stages: first free OPAC systems, then commercial front-end online systems, then CD-ROM databases, and now Web spaces. End-users' searching of bibliographic information retrieval (IR) systems has been amply documented by now. To date, scant research is reported on Web users' information needs and about how successful users are in finding the information they need.

Previous studies reveal that end users have difficulties in searching databases. Factors affecting search success may include users' cognitive states, knowledge of the system, and situations. Based on empirical results, Borgman identified three types of users' problems with online catalogs: conceptual, semantic, and technical (Borgman, 1986, p. 388; 1996, p. 495). In a large-scale empirical study of humanities scholars' online searching, Bates et al. found that subject categories frequently used in the humanities significantly differ from those in other disciplines (1993, p. 15). Bates (1996) also pointed out that a search mechanism, such as the near operator in DIALOG, may be easy to use for users with science background but difficult for users with a humanities background (p. 520-1). Many problems facing users of online searching have been predicted by Taylor's model of four levels of information needs (1968). For example, the compromised need (fourth level) or the query must be acceptable by the IR system in order to retrieve information. At the cognitive level, Belkin's anomalous states of knowledge (ASK) predict that users have difficulties in articulating their information needs (Belkin et al, 1982).

Many problems facing users of traditional bibliographic IR systems remain for Web users: a user must be able to translate his or her topic (an anomalous internal representation) into a query (a formula of precise terms and their relationships). The user must also know the types of information available from the selected Web space. In their studies of potential users' Web searching, Pollock and Hockley (1997) report that nearly all participants had difficulty formulating good queries with keywords; common errors included queries expressed in natural language, misspelling, terms broader than the search concepts, etc. They believe that an intelligent interface can provide useful help simply by suggesting possible terms, perhaps using a spell-checker, and hierarchical categories for browsing rather than searching.

Emerging Web classifications, entitled subject guide and directory, are hierarchically organized indexes of subject categories that allow the Web searchers to browse through lists of Web sites by subject, such as Yahoo. "Because subject guides are arranged by category and because they usually return links to the top level of a web site

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

rather than to individual pages, they lend themselves best to searching for information about a general subject, rather than for a specific piece of information." (Tyner, 1997, #guide). On the contrary, widely used traditional bibliographic classifications have taken either a philosophical approach, such as DDC, or a literary warrant approach, such as LCC; they provide both browsing and searching functions for subject access. Both DDC and LCC classifications are standards and have been developed by subject experts who have control over categories and vocabulary.

Shneiderman (1997) points out that many search systems on the Web are neither simple nor clear; zero-hit outcomes occur on 30% of searches (p.2). Based on information-processing models, errors in interacting with Information Retrieval systems can reveal both users' incorrect mental models (Norman, 1983). Dickson (1984) examined log files of users' searching of an OPAC system and identified error types of user queries to understand users' mental model of the system.

METHODOLOGY

Online monitoring is the method to collect data from users during their interaction with the computerized systems. Usually, a log file is created to capture users' key strokes and systems' responses with timestamps. Monitoring data is very useful for studying end-users' searching behaviors, and this method has been used and documented by many researchers in the field (Bishop and Star, 1996; Borgman, Hersh, & Hiller, 1996; Marchionini et al., 1994; Rice and Borgman, 1983; Penniman and Dominick, 1980).

Some advantages of online monitoring include accuracy, unobtrusiveness, and the collection of longitudinal, transactional, and temporal data which can be automatically collected and processed. Disadvantages include privacy concerns, an overwhelming amount of data that may be difficult to manage, and the fact that collected data, although accurate, may be open to conflicting interpretations. To overcome these shortcomings, the following measures are used. Since the log file in this study does not trace each individual users' queries, privacy is not an issue here. An appropriate sampling can be used to reduce the amount of data for analysis. The researchers are aware that interpretation of the data should be limited to the user group of the selected Web site as a whole. Therefore, monitoring data are important to our understanding of a physically scattered population of a specific Web site, such as the UTK Web pages.

The UTK home page

The main home page of the University of Tennessee, Knoxville (UTK) has a classification scheme with links to pages at the second level and to pages of other servers. A search tool is available in the main home page for users to perform keyword searches. The search tool is build using the Simple Web Indexing System for Humans (SWISH). Due to the limit of computing resources, only the pages at any of the three levels from the main home page (http://www.utk.edu/) are indexed. There is only word parsing for each indexed page. Neither single letters nor numerics are indexed. A stop-word list is also used to exclude certain words such as of, on, and in. Frequently occurred words are ignored as well. That is, words that occurred in over 80 % the files will not be indexed.

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

Boolean operators are supported; the default search operator is AND. The search interface is not case-sensitive. Bound-phrase searching is not available. Queries using double quotation marks are accepted. But, queries containing punctuation such as commas, colons, semi-colons, and single quotation, are rejected.

The log file

The log file maintained by the Web master of the UTK Web page records the queries entered by the users with date stamp at the beginning and the postings at the end. The file is a delimiter ASCII file. A single query is the search formula submitted by the user to the system by hitting the Enter key or clicking on the Search button. The following are two example queries from the log file:

```
1997/03/06::Ronald McNair post baccalaureate Achievement Program::1
1997/03/10::The telephone number for the main office for the small animal clinical services::0
```

Data analysis

The unit of analysis is a single query. Each query is analyzed about its subject matter. For queries causing zero-hit outcomes, the types of errors are further analyzed. The first query presented above is categorized as searching for academic information. The second query is categorized as searching for computing related information as well as error type, syntactic (stop words). Queries resulting in non-zero postings are not analyzed because the data for postings may not be the actual postings: the UTK Web search tool allows users to limit the number of returned URLs to 10, 20, 50, 100 or 500 (the default is 100).

During March 1997, which included one week of Spring break (from March 22 to 29), users of the UTK Web site performed a total of 5778 queries. Of the total amount of queries executed in March 1997, there were 1906 queries (33.0 %) returned zero-hit outcomes.

A random sample of 8 days from March 1997 is selected for analysis: March 1, 3, 5, 12, 16, 21, 25, and 27; they covered all days in a week and two Wednesdays, of which one fell in Spring break.

A preliminary analysis of queries by one of the researchers derived the coding scheme. Both researchers then coded a set of data independently and revised the coding scheme, which is finalized after coding reliability check. Coding reliability is checked by double-coding and calculating the coding agreement between the two coders. In this study the agreement was 89%, which is greater than the acceptable level of intercoders' consistency (80%) recommended by Krippendorff (1980). The disagreement on the rest 11% was resolved by discussion and expansion of definitions for the related categories.

RESULTS

For March 1998, the average number of queries per day is 192. There were 248 searches per weekday outside of Spring break, and 145 searches per weekend and holiday break. A total of 1,537 queries by the users during the eight randomly selected days were analyzed (Table 1). The highest number of queries is 280 (Wednesday, March 5) and the lowest number of queries is 50 (Thursday, March 27). Since March 27 fell in Spring break, the users of the UTK Web site seemed to be mostly UTK students and faculty. There were 515 queries resulted in zero-hit (Table 2), which is 33.5% of the total queries.

--- Insert Table 1 here ---

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

Subject categories of user queries

Fifteen subject categories emerged reflecting information needs of this user group. They are listed as follows:

1. *People.* This class includes queries related to finding a person by name and position: "Jennifer White" and "academic staff" are included here.
2. *Institutional Unit.* It includes all queries pertaining to a distinctly identifiable office, building, institute, department, and academic program at UTK. "Paul Scherrer Institute", "financial aid office", "conference center", "forensic anthropology" are examples.
3. *Maps.* Queries request for maps, directions, and campus locations. Examples are "Marketplace" and "Neyland Stadium."
4. *Academic Information.* Including all queries pertaining to academic information and the administrative functioning of the university such as admissions, schedules, calendars, transcripts, scholarships, etc. Examples include "evening school", transcripts", residency requirements", "whittle scholars", etc.
5. *Student Social Life.* Queries search for student societies, associations, fraternities and sororities, and clubs. Examples are "pi kappa phi", "sororities", "presidents club."
6. *Housing.* Queries pertain to dormitories, residence halls, university-owned apartments. Examples are "residence halls", "roommate", "dorms."
7. *Sports and Recreation.* Queries search for sports, sports events, university radio, and theater. Included are "football tickets", "Lady Vols", "swimming", "WUOT."
8. *Resources.* Searches look for resources available on campus such as libraries, computer labs, bookstores, and newspapers. Examples are "Hodges library", "campus store."
9. *Guidelines.* Queries ask for university publications and policies, regulations, and include queries such as "faculty handbook", "hill topics", "code of ethics."
10. *Reference Questions.* Queries search for topics that should have been performed in OPAC or bibliographic databases, for instance "Evan Thomas' why we did it," "Teacher in Special Education Shortages." Queries that cannot be classified into other categories are also included in this category because a reference interview may be needed.
11. *Computing.* Queries contain email addresses, URL locations, host names and computer terminology such as "Bagwell@utkux.utcc.utk.edu", "dogwood" (server name), "telnet", "modem pool."
12. *Benefits.* Queries ask for various benefits available to students, faculty, and staff such as parking, health insurance. Examples: "disability services", and "health services."
13. *Images.* Queries specifically search for pictures, photographs, drawings and digital images. Queries in this category are rare and include "mascot picture" and "football pictures."
14. *Forms.* Queries request for specific forms used for leave of absence, travel and expense reimbursement, etc. For example, "research forms A B."
15. *Employment.* Looking for job announcements and employment opportunities such as "civil engineering entry level job openings."

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

Table 1 summarizes the distribution of the queries by each subject categories. The results are not surprising: the most-frequently requested information category is Academic Information. Queries of this category for the 8 days ranged from 17.1 % to 32.0 % with the average of 21.1%. The second-frequently requested category is Institutional Unit, ranging from 13.4% to 28.9% with an average of 18.9%. Based on the frequency ranking, the order of these categories is as follows: Academic Information, Institutional Unit, Reference Questions, People, Computing, Sports and Recreation, Resources, Student Social life, Housing, Benefits, Maps, Guidelines, Employment Images, and Forms.

Errors causing zero-hit

Two types of errors caused zero-hit were identified: Syntactic 0 and Semantic 0 (Table 2). Syntactic errors are due to lack of knowledge of the syntax (rules) in formulating queries acceptable by a specific system. For example, the errors caused by including in queries forbidden punctuation, stop words, typographical errors, missing spaces, incorrect Boolean operators (such as using + for AND), and numerics (i.e. 1997). Examples of queries included in Syntactic Zero-hits are "Dean of Students", "women 's center", "pre-veterinary medicine", "sports+logo", "1997 summer session." Queries for computer strings, such as URLs, e-mail addresses and directories in servers, fall among syntactical errors when they include some form of punctuation (http://www.it.utk.edu/itc/course). Another popular Syntactic Zero-hit is reversed headings, such as " Cobb, James."

Semantic zeros are due to users' lack of knowledge about the content coverage of the Web space or related to indexing and linguistic problems. Specifically, a query may not be able to retrieve URLs, either because the information is not available for the Web site, or because users' linguistic expressions of a concept were different from systems' linguistic expressions. Queries fall into this class are syntactically correct. These queries failed because of no match between the query words and the index words. Examples of such queries are "jeep cherokee" (car), " hyatt regency" (hotel), "cheerleading" (cheerleader was an indexed word), "Smith Jane" (directory was not indexed). Queries on personal names without syntactic errors were counted as Semantic 0. Computer strings, names of programs and algorithms are also counted as Semantic Zero-hits when they do not contain forbidden characters. Such queries include "html", and algorithms and computer libraries available at on UTK Web site such as "netlib", and "PhotoShop." Similarly to what happens with names, users are not aware of a separate search engine for matters related computer science, although it is available from UTK search page, together with the Keyword search, and the People, Places and Offices directories.

--- Insert Table 2 ---

Results summarized in Table 2 indicate that there were total 515 zero-hit queries, which is 33.5% of the total queries analyzed. Of which, 276 queries (53.6%) contained syntactic errors and 239 queries (46.4%) had semantic problems. As Table 2 indicates that the category People has the highest percentage of zero-hit (52%), which is followed by Reference Questions (50%). Both categories have more Semantic 0 than Syntactic 0. This is not surprising because the index does not support either categories. For the most frequently searched categories,

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

Academic Information and Institutional Unit, zero-hit outcomes are 29% and 25% respectively. Zero-hit queries for both Academic Information and Institutional Unit have high percentage of Syntactic 0 (78% and 83% respectively). The category Computing also have high Syntactic 0: 70%.

DISCUSSION

As pointed out in Methodology section, the present study provides a general picture of the user group and their subject needs collectively. There are several limitations: (1) Individual differences cannot be identified by the data from the log file, which was unable to mark the beginning and end of a search session by individual users. This is similar to the situation noted by Dickson about her search logs of OPAC users (1984, p.25). (2) Some queries are difficult to discern the exact information need. A query "Business and Management" might be looking for Academic Information or locating Institutional Units. (3) Because the users' responses to the search results were not captured in the log file, it is impossible to know whether a search outcome provided relevant URLs. Nevertheless, the log file provided valuable data on the subject needs of this user group, the nature of current end-user Web searching and the problems encountered by users.

The fifteen subject categories reflect users' expectations of information that can be found in the UTK Web site. It should also be pointed out that the category, Academic Information (ranked number 1 for frequency), needs to be divided further to be useful. Further analysis of the queries in this category is needed because it has been searched most frequently by this user group. These query-derived categories may be used as a practical classification scheme to index UTK Web pages to provide better access.

By its very nature, online information retrieval is a complex task, whether one searches on a sophisticated bibliographic database, such as DIALOG or FirstSearch. The results from this study revealed that one-third of the queries returned no URLs in the UTK Web search, which is slightly higher than the number (30%) reported by Schneiderman (1997). The zero-hit outcomes signal the minimum failure of all searches performed. If non-relevant results were counted, the failure rate would be much higher. It is obvious that users need help in using Web search tools.

The zero-hit outcomes seemed due either to syntactic difficulties or semantic barriers. Users who are not aware of the syntactic aspect of search tools tend to enter natural language queries, which are sentence-like or phrase-like including stop words. For these type of users, the solution can be expert systems that are capable of processing natural language, in which "users can enter a request in a loosely structured format, preferably in natural language, sentence-like expression" (Fidel, 1986, p.37).

Many syntactic errors found in this study are also due to the diversity of query syntax across Web searching tools. The many different search tools available on the Web certainly make searching a complex task. The users who have used one search system tend to formulate their queries in the format set by that system. However, Syntactic errors can be automatically corrected by error-detection-and-correction techniques, such as mapping queries among the systems or dropping the stop words or punctuation in a query. To correct user's mental model, constructive context-sensitive help messages should be provided.

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

Semantic errors are often difficult to correct automatically. A Web space must deal with indexing and linguistic problems just like traditional information retrieval systems: to provide terminological assistance. The Web users tend to use keyword search for their needs, even though a subject directory is available for browsing. A fundamental problem with keyword search is the term-matching algorithm. If the vocabulary used by the text is different from the one used by the searcher, zero-hit occur. Can the vocabulary from the queries collected in the log file be used to build a practical thesaurus? Can the categories derived from these queries form a practical classification scheme to be used to index the Web pages? Can the practical thesaurus and the practical classification be linked to provide search help? For instance, if a query retrieved no page, the system will look for the query terms in the thesaurus to identify corresponding candidate categories. In this way, a search algorithm beyond term-matching can be developed to convert a keyword search to a category search. In stead of a zero-hit result, the user is presented with the candidate categories and now brought into a browsing mode along with the option for entering a new search.

This study applied a bottom-up method to generate categories of user queries inductively. Like other cognitive categories, the boundary between these categories may be blurry. Some queries may belong to two or more categories. To deal with this problem, interaction between the system and the user must be enhanced. For a query "Business and Management," the system can certainly ask the user after consulting the practical thesaurus and classification, "Are you looking for an academic program or looking for an institutional unit?"

CONCLUSIONS

The users' subject needs of the UTK Web page fall into 15 categories, such as searching for academic information, finding an institutional unit, looking for social life, sports, computing, resources, etc. Some of the categories are more frequently requested than others. Therefore further analysis is needed to break down these categories into evenly distributed categories. This practical classification has a potential to be used to organize the Web space, because they are derived from real users' queries.

Generally speaking, one-third of the queries entered by the users of the UTK Web page returned no URL. But, the high number of daily queries indicates that users want searching not just browsing. The queries resulted in zero-hit outcomes contain either syntactic or semantic errors. Theoretically, syntactic errors can be automatically corrected by error-detection-and-correction techniques. Semantic errors, however, may be reduced by increasing users' knowledge of the system and by developing advanced functions to provide help.

This study also demonstrates a method to build a practical classification scheme based on user queries. The statistical features of the categories may be used for category breakdown. A practical thesaurus and classification derived from user queries may be able to provide candidate categories for browsing when a search returns no hit. Instead of indexing the full-text of each page, these query-generated categories can be used as subject headings in meta-data coding.

Although this study analyzed one university's Web queries, some of the results may also be applicable to similar universities. Comparing the results from this Web site with other university Web sites may suggest a practical thesaurus and classification applicable to these similar universities.

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

Since log files of the UTK Web users' searches are readily available, longitudinal analysis will provide valuable information on whether users' subject needs change over time and whether their searches become more successful.

ACKNOWLEDGEMENT

Sincere thanks go to the Office of Research, the University of Tennessee Knoxville for the award of a SARIF Summer Graduate Research, to the technical support from Bill Rosener for capturing Web queries, and to Tim Henderson for processing raw data. The helpful comments from reviewers are greatly appreciated.

BIBLIOGRAPHY

Bates, M. J. (1996). The Getty end-user online searching project in the humanities: Report No. 6: Overview and conclusions. *College & Research Libraries, 57*, 514-523.

Bates, M. J., Deborah N. Wilde, and Susan Siegfried. (1993) An Analysis of search terminology used by humanities scholars: the Getty online searching project Report No. 1. *Library Quarterly*, 63 (1), 1-39.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982). Anomalous states of knowledge as a basis for information retrieval. *The Journal of Documentation, 38*(2), 61-71.

Bishop, A., & Starr, S. L. (1996). Social informatics of digital library use and infrastructure. In M. E. Williams ed., *Annual Review Of Information Science And Technology* Vol. 31 (pp. 301-401). Medford, NJ: Information Today.

Borgman, C. L. (1986). Why are online catalogs still hard to use? Lessons learned from Information-Retrieval Studies. *Journal of the American Society for Information Science*, 37(6), 387-400.

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

Borgman, C. L., Hirsh, S. G., & Hiller, J. (1996). Rethinking online monitoring methods for information retrieval systems: From search product to search process. *Journal of the American Society for Information Science, 47*(7), 568-583.

Dewey, M. (1885). Decimal Classification and Relative Index: Introduction. In Lois Mai Chan, Phyllis A Richmond, and Elaine Svenonius, eds. (1985). *Theory of Subject Analysis. A Sourcebook.* Littleton, CO: Libraries Unlimited. 21-33.

Dickson, J. (1984). An Analysis of user errors in searching an online catalog. *Cataloging & Classification Quarterly*, 4(3), 19-38.

Fidel, R. (1986). Towards expert systems for the selection of search keys. *Journal of the American Society for Information Science, 37*(1), 37-44.

Krippendorff, K. (1980). Content Analysis: an introduction to its methodology. Beverly Hills, CA: Sage.

Martel, C. (1911). Classification: A brief conspectus of present day library practice. In Lois Mai Chan, Phyllis A Richmond, and Elaine Svenonius, eds. (1985). *Theory of Subject Analysis. A Sourcebook.* Littleton, CO: Libraries Unlimited. 71-85.

ISSN: 2324-9773

Wang, P., & Pouchard, L. (1997). End-user Searching of Web Resources: Subject Needs and Zero-hits. 8th ASIS SIG/CR Classification Research Workshop, 69-80. doi:10.7152/acro.v8i1.12729

Marchionini, G., Barlow, D., & Hill, L. L. (1994). Extending retrieval strategies to networked environments: Old ways, new ways, and a critical look at WAIS. *Journal of the American Society for Information Science, 45*(8), 561-564.

Norman, D. (1983). Some observations on mental models. In Dedre Gentner and Albert L. Stevens, eds. (1983). *Mental Models*. London: Lawrence Erlbaum Associates Publishers. 7-15.

Penniman, W. D., & Dominick, W. D. (1980). Monitoring and evaluation of on-line information system usage. *Information Processing & Management, 16*(1), 17-35.

Pollock, A., & Hockley, A. (1997). What's wrong with Internet searching. *D-Lib Magazine, 3*, 1-5. Internet. http://www.dlib.org/march97/ bt/03pollock.html. Accessed on June 14, 1997.

Rice, R. E., & Borgman, C. L. (1983). The use of computer-monitored data in information science and communication research. *Journal of the American Society for Information Science, 34*(4), 247-256.

Shneiderman, B., Byrd, D., & Croft, W. B. (1997). Clarifying search: A user-interface framework for text searches. *D-Lib Magazine,* (1), 1-18. Internet. http://www.dlib.org/dlib/january97/ retrieval/01shneiderman.html. Accessed on June 14, 1997.

SWISH. Available at URL: http://www.eit.com/software/swish/ Accessed on September 21, 1997.

Taylor, R. S. (1968). Question-negotiation and information seeking in libraries. *College & Research Libraries, 29*, 178-194.

Tyner, R. (1997). Sink or Swim: Internet Search Tool & Techniques. Version 2.1. 1-12. Internet. Available at http://www.sci.ouc.bc.ca/libr/connect96/search.htm. Accessed on September 21, 1997.

ISSN: 2324-9773

Table 1: Distribution of subject categories of queries for randomly selected eight days (%)

Subject Category	March 1 (N=142)	March 2 (N=291)	March 5 (N=280)	March 12 (N=256)	March 16 (N=125)	March 21 (N=167)	March 25 (N=224)	March 27 (N=50)	Eight days (N=1537)
People (1)	3.5	13.1	7.9	15.6	13.6	9.0	8.0	10.0	10.4
Institutional Unit (2)	23.9	14.4	19.6	28.9	17.6	16.1	13.4	16.0	18.9
Maps (3)	1.4	2.1	2.9	0	4.0	3.0	3.6	0	2.2
Academic Information (4)	21.1	20.6	17.1	20.3	25.6	20.4	23.2	32.0	21.1
Student Social Life (5)	6.3	3.8	5.4	2.0	2.4	1.2	6.7	0	3.9
Housing (6)	2.8	4.8	3.2	1.6	0	4.8	6.3	0	3.4
Sports and Recreation (7)	9.9	6.9	4.3	5.4	8.0	6.0	5.8	2.0	6.2
Resources (8)	4.9	3.4	4.3	3.9	1.6	6.6	7.1	2.0	4.5
Guidelines (9)	0	3.8	2.5	3.1	3.2	0	0.9	0	2.1
Reference Questions (10)	12.7	8.6	15.6	6.3	12.0	16.2	8.9	16.0	11.2
Computing (11)	6.3	12.0	13.9	9.0	8.8	10.2	6.7	18	10.3
Benefits (12)	4.2	3.8	0.7	1.2	2.4	3.6	4.9	0	2.7
Images (13)	0	0.7	1.1	0.4	0	0.6	0.4	4.0	0.7
Forms (14)	0.7	0	0	0	0	0.6	0	0	0.2
Employment (15)	2.1	2.1	1.4	2.3	0	1.8	4.0	0	2.0
Column Total	*99.8*	*100.1*	*99.9*	*100*	*100.2*	*100.1*	*99.9*	*100*	*99.8*

ISSN: 2324-9773

Table 2: Frequency of queries caused zero-hit due to syntactic and semantic errors

Subject Category	Number of Queries	Total 0-hits (%) *	Syntactic 0 (%) **	Semantic 0 (%) ***
People (1)	160	83 (52%)	26 (31%)	57 (69%)
Institutional Unit (2)	292	72 (25%)	60 (83%)	12 (17%)
Maps (3)	34	10 (29%)	4 (40%)	6 (60%)
Academic Information (4)	324	94 (29%)	73 (78%)	21 (22%)
Student Social Life (5)	60	20 (33%)	8 (40%)	12 (60%)
Housing (6)	53	10 (10%)	4 (40%)	6 (60%)
Sports and Recreation (7)	95	36 (38%)	9 (25%)	27 (75%)
Resources (8)	69	12 (17%)	7 (58%)	5 (42%)
Guidelines (9)	32	9 (28%) .	2 (22%)	7 (78%)
Reference Questions (10)	172	86 (50%)	32 (37%)	54 (63%)
Computing (11)	158	60 (38%)	42 (70%)	18 (30%)
Benefits (12)	42	14 (33%)	5 (36%)	9 (64%)
Images (13)	10	3 (30%)	0	3 (100%)
Forms (14)	3	0	0	0
Employment (15)	31	6 (19%)	4 (67%)	2 (33%)
TOTAL	*1537*	*515* *(33.5%)*	*276* *(53.6%)*	*239* *(46.4%)*

* The percentage is calculated by dividing Total 0-hit by Number of queries, e.g., 83/160 = 52%

** The percentage is calculated by dividing Syntactic 0 by Total 0-hit, e.g., 26/83 = 31%

*** The percentage is calculated by dividing Semantic 0 by Total 0-hit, e.g., 57/83 = 69%

ISSN: 2324-9773