

# The Internet as a Source of Academic Research Information: Findings of Two Pilot Studies

Harry M. Kibirige  
and Lisa DePalo

*As a source of serious subject-oriented information, the Internet has been a powerful feature in the information arena since its inception in the last quarter of the twentieth century. It was, however, initially restricted to government contractors or major research universities operating under the aegis of the Advanced Research Projects Network (ARPANET).<sup>1</sup> In the 1990s, the content and use of the Internet was expanded to include mundane subjects covered in business, industry, education, government, entertainment, and a host of other areas. It has become a magnanimous network of networks the measurement of whose size, impact, and content often elude serious scholarly effort.<sup>2</sup> Opening the Internet to common usage literally opened the flood gates of what has come to be known as the information superhighway. Currently, there is virtually no subject that cannot be found on the Internet in one form or another.*

**T**here is both hype and reality as to what the Internet can generate in terms of substantive information. In their daily pursuits of information, information professionals as well as end-users of information are challenged with regard to what their expectations are and what actually is delivered in terms of tangible information products and services on the Net. Academic users are a special breed in that both faculty and students have specific topics covered in their courses of study or faculty research agendas for which they need information. The use of electronic resources found on and off the Internet is becoming increasingly vital for education and training in academic environments.<sup>3</sup> Five basic elements often are required in the electronic resources that academic information seekers desire: accessibility, timeliness, readability, relevance, and authority. The Internet excels in the first three, but depending on how and from where the information is gathered, it may not be so reliable with regard to the last two elements.

The two pilot studies discussed in this article involved four academic institutions and were conducted by the researchers with approximately twelve months apart. One covering two institutions was done in the fall of 1997. It was replicated covering another two institutions in the spring of 1999. The main goal of the studies was to investigate how academic users perceive search engines and subject-oriented databases as sources of topical information. The basic underlying question was, "When faced with a topical subject, what is the users' predominant recourse, online databases (which may include CD-ROM, or DVD databases) or search engines?" Our

results indicated that there is predominant preference for search engines for the group taken as a whole. Further analysis using nonparametric correlation coefficients—Kendall's tau\_b and Spearman's rho—however, indicated that those who use the Internet monthly or weekly had high correlations with online databases as their preferred predominant information sources. On the other hand, daily users tended to have high correlations with search engines as preferred predominant information sources.

## Information Seeking Behavior of Academic Users

Over the years, several studies have been conducted on how users seek and find information relevant to their needs. For the purposes of our analysis three categories will be used: the undergraduate, the graduate, and the post-doctoral research faculty user. While the levels of how the needed information may be articulated and packaged may be different, the five basic required elements in the electronic information resources needed by academics, already identified, remain the same. The Internet has, however, added another dimension to the information-seeking behavior of all academics in that much of the needed information, if and when found, has a higher chance of appearing as full text (sometimes defined as viewdata) on the Internet.<sup>4</sup> With viewdata the end user has the ultimate in information seeking and acquisition in that he or she will get text, images, and sound in one, two, or more resources on the Net. The process also may be accomplished in one sitting or search session on the computer terminal. The Internet thus may be more likely to generate viewdata in contrast to conventional databases, which have for a long time been associated with the less desirable citations. In many instances and with a little persistence, it can provide the analogy of "one stop shopping" whereby a user can get viewdata needed for a topic. This may explain the tendency to try the Internet first as a potential information source even for experienced searchers. To be effective, such searching needs experience and a lot of patience while sifting through pages of useless verbiage, as the information sources often are garnered from several sites. Categories of academic users have varying levels of expertise in information seeking and

---

**Harry M. Kibirige** is Associate Professor at Queens College, City University of New York, and **Lisa DePalo** is Assistant Professor at College of Staten Island, City University of New York.

---

have different characteristics in their information-seeking behavior.

### **Undergraduate Users**

Undergraduates are at the lowest point on the totem pole with regard to expertise in information seeking at any academic institution. There is more to the information needs of undergraduate students than can be revealed during the reference interview process. There are the pervading needs that the information age has created, which can be met only by those who possess critical thinking skills. Critical thinking skills are imperative to much more than completing college-level assignments—they are also imperative to surviving in the job market once students graduate. This premise has been set forth in the 1992 United States government report from the Department of Labor and The Secretary's Commission on Achieving Necessary Skills (SCANS) entitled *Skills and Tasks for Jobs: A SCANS Report for America 2000*. This report defined two types of skills needed to excel in the workplace and labeled them as competencies and foundations. Effective reference and instruction services can help students develop the critical thinking skills needed to meet the information competency, in particular, since it pertains to one who "acquires and evaluates information, organizes and maintains information, interprets and communicates information, and uses computers to process information."<sup>5</sup> Acquiring and evaluating information can be particularly difficult for undergraduates in the information age since one is bombarded with data in print and electronic formats. One can easily determine the reliability of print sources by looking at the name of the author, editor, or publisher. However, the Internet has become a popular choice for students who need to do research. It has gained the reputation for providing all that one needs right at one's fingertips. The problem is that one cannot readily discern what is reliable and what is not without some instruction.

It may be argued that the undergraduates' information seeking is somewhat eased by the general guides they get from the faculty in the classroom. There is the general professorial lecture which outlines the topics to be covered during the course, as well as associated relevant readings used to broaden the subjects covered. In addition, there is the text book which elaborates on material covered in class. Finally, there are journal articles and other information sources which ordinarily are placed on reserve. As far as subject content covered in class lectures and discussion is concerned, information is usually well organized and accessible. At that level information seeking is minimal and often guided by the dictates of the professor.

But then enters the term paper and the whole student peace of mind with regard to information gathering habits is disturbed. The term paper brings many unknowns to

the undergraduate. The magnitude of the subject to be covered is initially fuzzy. The resources needed to get background as well as specific information are also fuzzy. Furthermore, even when the resources are a little clear, sifting through them and making rational selection of relevant material may be problematic. The whole academic exercise entails learning and using new information tools, many of which were not covered in high school. Computers and other electronic equipment have accentuated the undergraduates' mesmerization process in their information-seeking effort. A trait that most undergraduates exhibit in their information-seeking behavior is approaching the reference librarian for suggestions of leads to information sources needed for the term paper topic. They also may request the librarian to evaluate the sources as to their relevance, and sometimes even ask him or her to fetch the actual material needed.<sup>6</sup> With the advent of the Internet and other electronic resources online or otherwise, (e.g., Dialog, Lexis-Nexis, CD-ROMs, DVD, and tapes), the undergraduate may go directly to the Internet terminal and thus skip the librarians' counsel and hand-holding which used to be vital for accessing the printed material. Unless the undergraduate student is well-groomed in searching the Internet, this relatively new tendency to act independently of the information professional may result in hours of useless roaming on the Net with little relevant information retrieved.

### **The Graduate User**

In their study of business students, Atkinson and Figueroa found that graduates reported fewer hours spent in the library than undergraduates.<sup>7</sup> The researchers did not attempt to explain why that was so. Perhaps because of their search skills, graduates do more focused information seeking and do not waste much of their time browsing and floundering in the unknown information abyss within the library. The researchers reported an equal interest in searching Internet resources and online databases (e.g., Lexis-Nexis, Dow Jones, and ABI/Inform), among graduates and undergraduates. However, their research was done at the end of 1995 and beginning of 1996, before the proliferation of search engines on the Internet.

As an information searcher the graduate is more sophisticated compared with the undergraduate. Subject coverage is usually more clearly defined in many of the assignments encountered. He or she has gone through most of the pitfalls of the undergraduate experience and can select a subject and research it relatively well. Most likely due to the nature of their assignments, undergraduates' information needs may be satisfied by simple information systems that allow users to browse. Their searches also tend to be less exhaustive than graduates. On the other hand, graduates are faced with relatively

---

narrower subjects and prefer to conduct more comprehensive searches.<sup>8</sup>

### **The Post-Doctoral Researcher—Faculty**

Faculty have mastered the art of getting relevant information. Many belong to the informal invisible college and attend professional conferences, both of which are used to get information for teaching and research. Hart's study found that formal sources, which may be found in the personal and college or university library, are more important in the faculty's information-seeking effort than informal ones.<sup>9</sup> According to Hart, this information-seeking characteristic would be applicable to printed and electronic resources found on the Internet. Although our research did not specifically test it, online databases tend to direct the end user to formalized definitive and tested resources than the Internet search engines. This would minimize user search time and maximize relevance of the information needed by the research academic faculty. In other words, while the listserv might be one of the Internet substitutes for the invisible college, information found on it would be more acceptable to a research faculty if it directs him or her to reliable and verifiable databases, i.e., information from CENDATA (U.S. Census Bureau information database), EDGAR (U.S. Security and Exchange database), or Dow Jones.

Developments in the electronic resources arena have made many hard copies less popular. Subject-oriented databases can be searched either in the library or in faculty offices. Curtis et al. researched the information-seeking behavior of health sciences faculty and found a relatively new and growing information-seeking characteristic. According to Curtis et al., faculty tend to prefer to search electronic resources from their offices rather than go to the library.<sup>10</sup> That is not surprising, for if a faculty member can access library catalogs and electronic databases, some of which can provide viewdata (full text), it is not necessary for him or her to go to the library for some of the information needed. In addition, if CD-ROM databases are on a local area network accessible via the college online catalog, faculty may seldom go to a library whose resources are on the network via a library Web site, Telnet, or the traditional dial up.

### **The Pilot Studies**

With the general information-seeking behavior of academic users in mind, the researchers decided to investigate the use of search engines for information sources in the academe in the New York metropolitan area. Search engines were contrasted to databases which may be URL- (Universal Resource Locator) accessible online via an

Internet browser, stand alone on CD-ROM, or on CD-ROM towers linked by a library local area network. In her article on Web search engines, Schwartz discussed recent studies done on their performance. She pointed out the fact that the end user is not often a participant in such studies.<sup>11</sup> Although our research was not on evaluation, we deliberately focused on the end user to gather statistics on perception of Web search engine utility in Internet surfing and information seeking. Kassel evaluated search engines indicating their variety and complexity when used to search the Internet.<sup>12</sup> Other relevant literature indicated the difficulty of navigating the Internet for both the information professional and the end user. It also indicated how direct access to databases was a shortcut to retrieving some of the topical information. Our periodic observations of Internet users revealed heavy use of search engines. We suspected that end users use them to get topical information which might otherwise be easily gotten from online databases. Consequently, we thought it necessary to conduct a study on end-user perception.

### **Objectives**

Our objectives in embarking on the pilot studies were to:

1. Find the frequency of Internet use by end users. This would allow us to check whether there is a correlation between frequency of Internet use and perception of search engine utility.
2. Find the most popular search engine. Examining the most popular search engine with respect to indexing policy might indicate whether it would generate more topical subject type of information.
3. Gauge the use of online and CD-ROM databases in the library. In order to help the end-users' memory as to what databases are involved in the research, common databases were listed on the questionnaire as examples.
4. Gauge the use of search engines in libraries and information centers. Common search engines likewise were listed to help the end user identify what they were.
5. Relate the results to pragmatic library and information-center functioning in providing information.

### **Methodology**

Four metropolitan New York academic institutions were selected: Borough of Manhattan Community College; Iona College; Queens College of the City University of New York; and Wagner College. The main criteria for selection was ease of access for the researchers. A composite sample of users was selected from these institutions to participate in the studies. The sample used was dynamic and self-selected in that whoever used the

"Internet Terminal" was a potential research subject. Only end users as opposed to information professionals/librarians were used in the study. While subjects sat at the terminal, they were requested to complete the questionnaire and return it to the reference/information desk.

Simplicity dictated the design of the research and data collection instrument (questionnaire). It was one page, multicolored, and was entitled "Internet Use Questionnaire." We estimated that it would take the subjects four to seven minutes to complete. Our assumption in designing it to be simple and least time-consuming was that since the subjects were sitting at the terminals, they were time conscious. While subjects were asked to complete the questionnaire, they had the option not to. Forty copies of the questionnaire were given to each academic institutional library, making a total of 160. Useable returns were 155, or 97 percent.

In addition to the questionnaire, we conducted exit interviews with some of the subjects who were using the Internet terminals after they handed in the completed questionnaires. The purpose of the interviews was to have some idea as to how the users perceived the utility of the Internet in getting electronic-based information. Four questions were used:

1. How do you find the Internet as an information source?
2. Did you get what you needed from the Internet?
3. Do you have a favorite search engine?
4. Is there any point when you would seek the assistance of the reference librarian/information specialist?

Analysis of the data was done using the SPSS (Statistical Package for Social Science) package. We used descriptive statistics for general group tendencies—frequency of Internet use and preferred sources for topical subject search. For inferential statistics we preferred the non-parametric pairwise two-tailed correlation coefficients, Kendall's tau<sub>b</sub> and Spearman's rho statistics. Microsoft's Excel program package was used to draw some of the illustrations.

## Results

The study revealed that an overwhelming majority of subjects (91 percent) use the Internet at least once a week (this includes those who use it daily). An almost equal number (45 percent) use it weekly—at least once a week); 46 percent use it at least once a day (see figure 1). As figure 2 shows, search engines are the predominant preferred tools for searching topical subjects on the

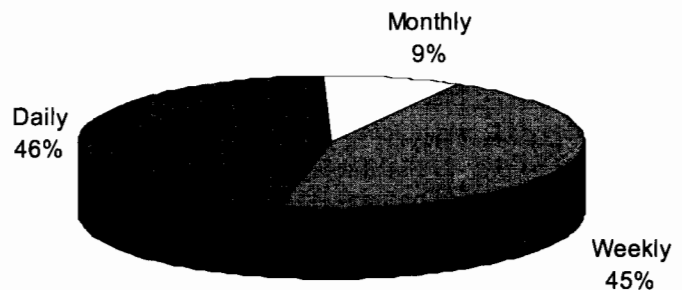


Figure 1. Frequency of Internet Use

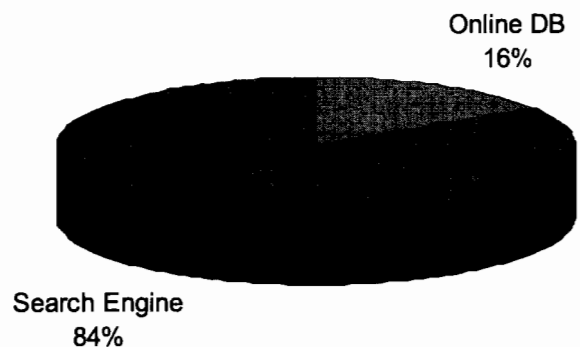


Figure 2. Preferred Sources for Subject Search

Internet as contrasted to online and CD-ROM databases. We used the two-tailed pairwise correlation coefficients to see whether there are correlations between frequency of Internet use and tool preferences. As table 1 and table 2 indicate, subjects who used the Internet monthly or weekly had high correlations with online databases. Daily users, however, tended to have high correlations with search engines as tools to get to topical subject information sources.

## Interpretations and Conclusions

Search engines certainly provide the most common access points utilized by library/information center users to get to electronic resources on the Internet. Unfortunately, the average user seems to have the impression that the Internet is a be-all and almost a panacea to all information problems. Kassel suggests

**Correlations**

			DAILY	SENG	MONTHLY	WEEKLY	ONDB
Spearman's rho	DAILY	Correlation Coefficient	1.000	-.544	-.258	-.544	.258
		Sig. (2-tailed)	.	.456	.742	.456	.742
		N	4	4	4	4	4
	SENG	Correlation Coefficient	-.544	1.000	.316	.500	.316
		Sig. (2-tailed)	.456	.	.684	.500	.684
		N	4	4	4	4	
	MONTHLY	Correlation Coefficient	-.258	0.316	1	0.949	0.8
		Sig. (2-tailed)	.742	.684	.	.051	.200
		N	4	4	4	4	
	WEEKLY	Correlation Coefficient	-.544	.500	.949	1.000	.632
		Sig. (2-tailed)	.456	.500	.051	.	.368
		N	4	4	4	4	
	ONDB	Correlation Coefficient	.258	.316	.800	.632	1.000
		Sig. (2-tailed)	.742	.684	.200	.368	.
		N	4	4	4	4	

**Table 1.** Nonparametric Correlations—Spearman's Rho

that, at best, search engines seem to reach just about half of the Web pages available on the Internet.<sup>13</sup> Sullivan has given several reasons why search engine coverage is incomplete and search results sometimes may be misleading.<sup>14</sup> Among the most cogent reasons are: documents may be changed after they have been picked up for inclusion; deleted materials may be displayed as available; and Web sites or files which are password accessible are not covered. Much of the information needed in academe is proprietary and available via database vendors. Using search engines as the main recourse to topical information shortchanges the user and may lead to frustration unless the high user expectations are tempered by constant education by the information specialist.

## Implications for Information Professionals

Our studies indicated that a majority of the users in the sample preferred the search engines as access points to the Internet for topical information. The interest in search engines correlated with the State University of New York at Albany study which also indicated their predominant use in searching the Internet.<sup>15</sup> While the Albany study was general, ours related the search engines to getting topical information and the use of online databases as an alternative.

Our findings point to the need to re-educate the Internet user in several aspects of the superhighway. First, content—the fact that only a fraction of the possible sites (approximately one half) are indexed by the search engines. Second, authority—because it is so easy to self-publish on the Internet, a lot of information of low integrity (for instance) or factual inaccuracy may be mistaken for reliable sources. Third, transiency of information found on the Internet must be pointed out. The maxim “here today, gone tomorrow” is appropriate for several Web sites on the Internet. Finally, information professionals must

**Correlations**

			DAILY	SENG	MONTHLY	WEEKLY	ONDB
Kendall's tau_b	DAILY	Correlation Coefficient	1.000	-.516	-.236	-.516	.236
		Sig. (2-tailed)	.	.346	.655	.346	.655
		N	4	4	4	4	4
	SENG	Correlation Coefficient	-.516	1.000	.183	.400	.183
		Sig. (2-tailed)	.346	.	.718	.444	.718
		N	4	4	4	4	
	MONTHLY	Correlation Coefficient	-.236.000	.183	1.000	.913	.667
		Sig. (2-tailed)	.655	.718	.	.071	.174
		N	4	4	4	4	
	WEEKLY	Correlation Coefficient	-.516.000	.400	.913	1.000	.548
		Sig. (2-tailed)	.346	.444	.071	.	.279
		N	4	4	4	4	
	ONDB	Correlation Coefficient	.236	.183	.667	.548	1.000
		Sig. (2-tailed)	.655	.718	.174	.279	.
		N	4	4	4	4	

**Table 2.** Nonparametric Correlations—Kendall's Tau-B

emphasize in their training the proven online databases to which users should go directly, if and when those databases are provided by the library or information center.

Information professionals have a direct link to providing users with guidance to proven online databases, specifically during course-integrated instruction. Education for the end user is paramount to the optimum utilization of electronic information sources. A well-developed information resources instruction program is needed in conjunction with the one-on-one instruction that takes place every day at the reference/information desk. Such instruction programs must be cumulative, if they are to be effective in an age of burgeoning choices for end users who can more and more often choose to be remote users of information resources. In an academic environment, early intervention at the freshman level is paramount, but also must be pursued in a structured manner at the upper levels. Many college and university information resources instruction programs are based on a one-shot, approximately fifty minute session, which often is executed as an orientation to the library/information center. Such a method of instruction has no guarantee that there will be further guidance sought, either at the behest of a teaching faculty member in the form of course-integrated instruction, or on an individual level at the reference desk. Developing effective ways to integrate information resources instruction into the lives of end users is one of the challenges information professionals face in the new millennium with an increase in the use of electronic resources found on the Internet.

## References and Notes

1. Jon Guice, "Looking Backward and Forward at the Internet," *The Information Society* 14, no. 3 (July/Sept. 1998): 201-11.
2. G. McMurdo, "The Net by Numbers," *Journal of Information Science* 22, no. 5 (1996): 1397-411.
3. N. L. Pelzer and others, "Library Use and Information Seeking Behavior of Veterinary Medical Students Revisited in the Electronic Environment," *Bulletin of the Medical Library Association* 86, no. 3 (July 1998): 346-55.
4. Harry M. Kibirige, "Viewdata," in *Encyclopedia of Electrical and Electronics Engineering*, vol. 23, ed. by G. Webster (New York: John Wiley, 1999), 223-31.
5. Department of Labor, The Secretary's Commission on Achieving Necessary Skills, *Skills and Tasks for Jobs* (Washington, D.C.: Department of Labor, 1992).
6. Gloria L. Leckie, "Desperately Seeking Citations: Uncovering Faculty Assumptions about the Undergraduate Search Process," *Journal of Academic Librarianship* 22, no. 3 (1996): 202-208.
7. Joseph D. Atkinson and Miguel Figueroa, "Information Seeking Behavior of Business Students: A Research Study," *The Reference Librarian* 58, (1997): 59-73.
8. Deborah Shaw, "Bibliographic Database Searching by Graduate Students in Language and Literature: Search Strategies, Systems Interfaces, and Relevance Judgements," *Library & Information Science Research* 17, no. 4 (Fall 1995): 327-45.
9. Richard L. Hart, "Information Gathering among the Faculty of a Comprehensive College: Formality and Globality," *Journal of Academic Librarianship* 23, no. 1 (Jan. 1997): 21-27.
10. K. L. Curtis and others, "Information-Seeking Behavior of Health Science Faculty: The Impact of New Information Technologies," *Bulletin of the Medical Library Association* 85, no. 4 (Oct. 1997): 402-10.
11. Candy Schwartz, "Web Search Engines," *Journal of the American Society for Information Science* 49, no. 11 (Sept. 1998) 973-82.
12. Amelia Kassel, "Internet Power Searching: Finding Pearls in a Zillion Grains of Sand," *Information Outlook* (Apr. 1999): 28-32.
13. Ibid.
14. Danny Sullivan, "Search Engine Coverage Study Published," Search Engine Watch. Accessed March 11, 2000, [www.searchenginewatch.com./sereport/99/08-size.html](http://www.searchenginewatch.com./sereport/99/08-size.html).
15. Wei Peter He, "What Are They Doing on the Internet?: Study of Information Seeking Behaviors," *Internet Reference Services Quarterly* 1, no. 1 (1996): 31-51.