

END USER DATABASE SEARCHING ON THE INTERNET: AN ANALYSIS OF THE STATE OF WISCONSIN'S BADGERLINK SERVICE

*Dietmar Wolfram and Hong (Iris) Xie,
University of Wisconsin-Milwaukee*

Keywords: End user searching, online databases, Internet

Abstract: BadgerLink, the Wisconsin state-funded Internet-based information service, provides access to a range of databases from EBSCO and Bell & Howell to qualified institutions and individuals in Wisconsin. Databases available through each vendor make available the full text or abstract and citation information for more than 4,000 periodicals and newspapers. Previously only available to information professionals with online access, the web-based BadgerLink service has now made these databases accessible to hundreds of libraries and countless end users. The recent availability of usage statistics of the service presents a valuable opportunity to evaluate how the service is being used. The authors analyzed six months of usage data, covering the period January 1st through June 30th 1999, to gain a better understanding of end user information seeking patterns. Data analyzed included databases accessed, periodical and monograph titles selected, document formats viewed, and institutional affiliation. The authors found that searchers tended to limit themselves to default databases on EBSCOhost. The most frequently accessed titles reflect a broad range of information needs, with academic or current event/information sources being most popular. Users took advantage of the full text access in addition to abstract and citation availability. Usage data also indicate that public, academic, and school libraries have been the heaviest users of the service. Implications of the findings for the provision of end user Internet-based information services are discussed.

1. INTRODUCTION

With the wider availability of Internet access and network-based information resources, state agencies have started to team up with online database vendors to offer state residents free online access to selected online databases via the Internet. A number of such services are now available, for example the BadgerLink service in Wisconsin (Ref. 1), the Inspire service in Indiana (Ref. 2), the Alabama Virtue Library (Ref. 3), and the Kentucky Commonwealth Virtual Library (KCVL) (Ref. 4). With these state-supported resources, residents can now connect to a variety of online databases directly from their libraries or homes to access the full-text content of magazines, newspapers and professional journals.

Wisconsin's BadgerLink service provides access to a range of databases from EBSCO (EBSCOhost) and Bell & Howell (ProQuest) to schools and libraries in Wisconsin via the Internet. EBSCOhost makes available more than a dozen databases with a range of topic coverage from disciplinary specialization (e.g. ERIC, Health Source Plus) to broad subject coverage (e.g. MasterFILE Premier, Academic Search Elite). Databases may include citation and indexing information, or may contain the full text of documents appearing in periodicals and monographs. The EBSCOhost web-based interface with its point and click approach to searching and browsing simplifies the database selection and search process. ProQuest provides access to a number of newspaper databases and ABI/Inform Global for business information. It is more specialized in that each of the newspaper databases contains records of the full text of a single newspaper, or set of newspapers. Since becoming available in 1998, BadgerLink has been used by hundreds of libraries and countless end users. Millions of searches have been performed on the available databases. To help in the evaluation of the success of BadgerLink, both EBSCOhost and ProQuest have kept usage figures of the service by various libraries.

By studying how Internet-based database services are used, researchers are able to gain a better understanding of the user population of these services, how users access online databases, what they try to access, and why they choose to access certain information resources. In a recent study, Tenopir identified the patterns of use and usage factors for online databases in academic and public libraries. The author examined peak usage time in different types of libraries and the factors that influence database use, such as levels of instruction, availability of remote login, placement of a database on the library's homepage, etc. (Ref. 5)

Data on database usage have been important for libraries in allowing them to provide appropriate access points, minimize costs associated with database licenses, and evaluate training programs provided for library patrons. (Ref. 6) For example, Academic Health Sciences Centers are interested in determining the most cost-effective approach for providing institutional access to MEDLINE and other health sciences databases through the study of database usage patterns (Ref. 7). Online database usage data could also be used to help in the transition from the paper-format library to the virtual library (Ref. 8).

Studying database usage may also enable researchers to design more effective information retrieval systems. Bates examined the user-system relationship in information retrieval system design by identifying online information search stratagems (Ref. 9). Another study called for a greater balance in the relationship between user control and ease of use in the design of effective and useful interactive online retrieval systems by exploring users' experience with different online databases (Ref. 10).

Since state wide online database access is relatively new, there has been little research focused on the patterns of use of these services. It is essential to investigate end users' patterns of use to aid in service evaluation and improvement, and to provide recommendations for state wide digital library projects.

The present study, which is part of a larger project evaluating the Wisconsin BadgerLink service, reports on BadgerLink resource usage. It also provides insights into the area of information seeking behaviors of end users on the Internet. More specifically, the following research questions were of interest to the investigators:

- Who are the most frequent users of the statewide service as determined by library type?
- Is the service widely used by libraries within the state, or is usage largely concentrated among selected institutions?
- Which database and title resources are most frequently accessed?
- What insights do the data reveal about end-user information seeking on the Internet?

2. METHODOLOGY

Both EBSCO and Bell & Howell maintain statistics of database usage for individual libraries. Data for Wisconsin libraries were collected for the time period between January and June 1999. Collected data included specific library usage, title usage, database usage and format of the requested documents (citation, abstract, full-text). Identification of libraries was only possible for those libraries that possess dedicated Internet access and an identifiable Internet Protocol (IP) address. Libraries accessing BadgerLink through a dial-in connection were not identifiable. Based on the frequency of usage findings by libraries, a random sample of Wisconsin libraries was selected and surveyed about their usage habits and general comments about the service. A web-based survey for end users was also developed and made accessible through the BadgerLink web site to get end user input on the service. Since the main purpose of the present study was to discover the general usage habits of BadgerLink, inferential statistical analysis was not undertaken. Data for the different areas of interest were tabulated and compared based on descriptive measures. A qualitative analysis is also presented.

3. RESULTS

3.1 Library Usage

Overall EBSCOhost usage by library type is reported in Table 1. Dial-in users, who are not identifiable, were the most frequent users. Of the identified libraries/institutions, the University of Wisconsin-Madison was the heaviest single user (266,770 searches), followed by the Milwaukee Public Library (176,899 searches) and the University of Wisconsin-Milwaukee (164,782 searches). Given the nature and size of the populations served by these institutions, the results are not surprising. Mineral Point High School recorded the largest number of searches among identified K-12 schools (53,841 searches). ProQuest usage followed a similar pattern (Table 2).

TYPE OF INSTITUTION	NUMBER OF PARTICIPATING INSTITUTIONS	SEARCHES CONDUCTED
Dial-in Users	N/A	2,165,359
K-12 Schools (excluding districts)	242	766,806
Colleges / Universities	35	689,047
Public Libraries	75	433,602
Junior/Technical Colleges	27	222,366
Corporate	21	34,469

Table 1 – EBSCOhost Usage Figures by Library Type

TYPE OF INSTITUTION	NUMBER OF PARTICIPATING INSTITUTIONS	SEARCHES CONDUCTED
Dial-in Users	N/A	175,348
K-12 Schools (excluding districts)	242	167,412
Colleges / Universities	35	62,139
Public Libraries	75	21,563
Junior/Technical Colleges	27	18,803
Corporate	21	1,707

Table 2 – ProQuest Usage Figures by Library Type

Although K-12 schools represent the most frequent users of BadgerLink, they also represent the largest category of participating institutions. Average usage by type of institution provides a better indication of general usage activity. Figures 1 and 2 show the average number of searches conducted per library by library type for EBSCOhost and ProQuest. Note that academic libraries constitute the heaviest users of both services. Junior/Technical College and Public libraries are the next heaviest users for EBSCOhost titles, while the K-12 school and Junior/Technical College libraries were also, on average, frequent ProQuest users.

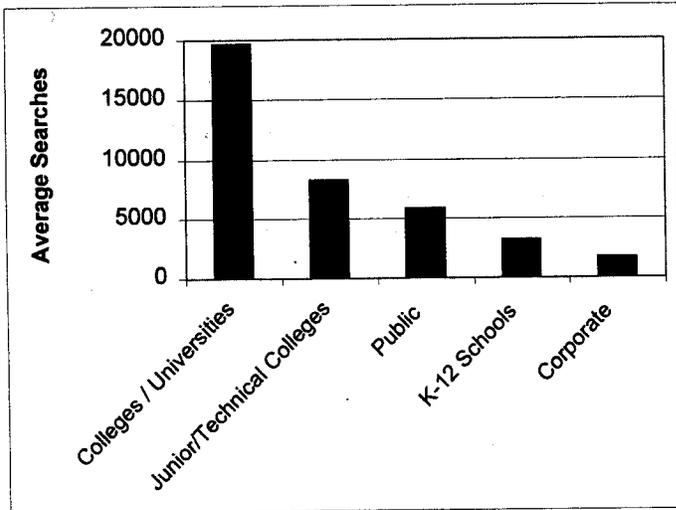


Figure 1. EBSCOhost Average Searches per Library

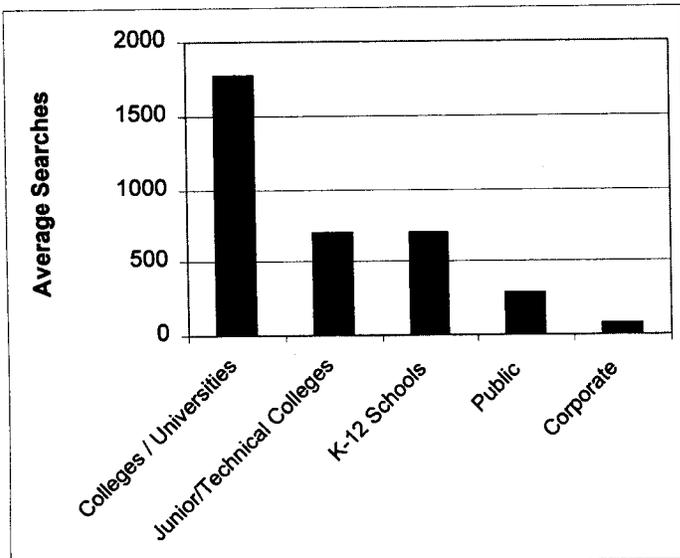


Figure 2. ProQuest Average Searches per Library

3.2 Database Usage

The ten most frequently searched EBSCOhost databases appear in Table 3. The selection of databases demonstrates a range of search areas and format coverage, including full text content and citations. Selected EBSCO databases tend to be those that provide broad topic coverage. Libraries that have direct access have the ability to set a default database. MasterFILE Premier has been designated as the default database for libraries with dial-in or shared accounts. Those libraries with dedicated access (such as many academic libraries) could select another database such as Academic Search Elite as the default. It is unknown at this time whether the defaults influenced the database usage patterns. For the most part, it appears that users are sticking primarily with the default databases for their searches.

Database selection within ProQuest has its own characteristics. ProQuest mainly provides access to newspaper sources. Among all the formats that ProQuest offers, full text is used more frequently than abstracts and citations. The most frequently accessed databases are National Newspapers (223,547 full text) followed by regional newspapers, such as Wisconsin Newsstand (79,433 full text) and Milwaukee Journal Sentinel (16,628 full text).

DATABASE	TOPIC / CONTENTS	SEARCH SESSIONS
MasterFILE Premier	General / Full text	697,009
Academic Search Elite	Multidisciplinary / Full text	569,299
Middle Search Plus	General for Middle & Junior High / Full text	114,776
MAS FullTEXT Premier	General & Research / Abstracts, Full text	100,419
Health Source Plus	Health / Full text	87,049
ERIC	Education / Citations & Abstracts	81,588
Business Source Elite	Business / Full text	70,564
Primary Search	General for Children / Full text	61,875
Comp MEDLINE w/ MESH	Biomedical / Indexing and Abstracts	47,151
EBSCO Animals	Animals / Indexing, Abstracts, Full text	41,469

Table 3 –Most Frequently Searched Databases on EBSCOhost

3.3 Title Usage

An analysis of the most frequently searched periodical titles available through EBSCO reveals a range of search areas that involve formal/academic and leisure topics including

current events, business, health, science & technology, sports and general interest. The 25 most frequently retrieved periodical titles, based on the number of abstracts viewed, appear in Table 4.

TITLE	TOPIC COVERAGE	ABSTRACTS VIEWED
New York Times Education Life	Education	68,751
Christian Science Monitor	News & Commentary	49,555
Newsweek	News & Commentary	39,239
Time Magazine	News & Commentary	28,912
WALL ST. JOURNAL	Finance / News & Commentary	23,745
U.S. News & World Report	News & Commentary	23,733
Economist	Business / Economics / Political Science	20,852
Lancet	Medical	19,843
People	General Interest	15,475
Billboard	Music	14,897
Time South Pacific	News & Commentary / Regional Studies	13,669
Alberta Report / Western Report	Regional Studies	13,367
SPORTS ILLUSTRATED	Sports & Leisure	12,128
Business Week	Business	12,018
Consumer Reports (USA)	Consumer education & protection	11,401
MacLean's	News & Commentary / General Interest	10,570
Science News	Science & Technology	10,042
Forbes	Business	9,732
ENTERTAINMENT WEEKLY	Communication & Media / Arts	9,400
New York Amsterdam News	News & Commentary / Multicultural & Ethnic / Regional	9,231
JET	News & Commentary / Multicultural & Ethnic	9,105
Library Journal	Library & Information Science	8,595
New Scientist	Science & Technology	8,308
Chronicle of Higher Education	Education	8,003
Advertising Age	Sales & Marketing	7,954

Table 4. Twenty-five Most Frequently Accessed Titles by Abstracts on EBSCOhost

News and commentary journals were the most frequently accessed titles, followed by business/finance, education, and leisure topics. There were no unexpected surprises in the

titles appearing in the list, other than perhaps Alberta Report/Western Report, a regional studies journal, which ranked twelfth.

Since thousands of titles are indexed within EBSCO databases, a full list by title cannot be included here. The frequency of use distribution of all titles (Figure 3) demonstrates that, although a small number of titles are heavily viewed during searches, almost all titles have been viewed during the six-month period of the study. Two-thirds of all titles (66%) were viewed at least 10 times.

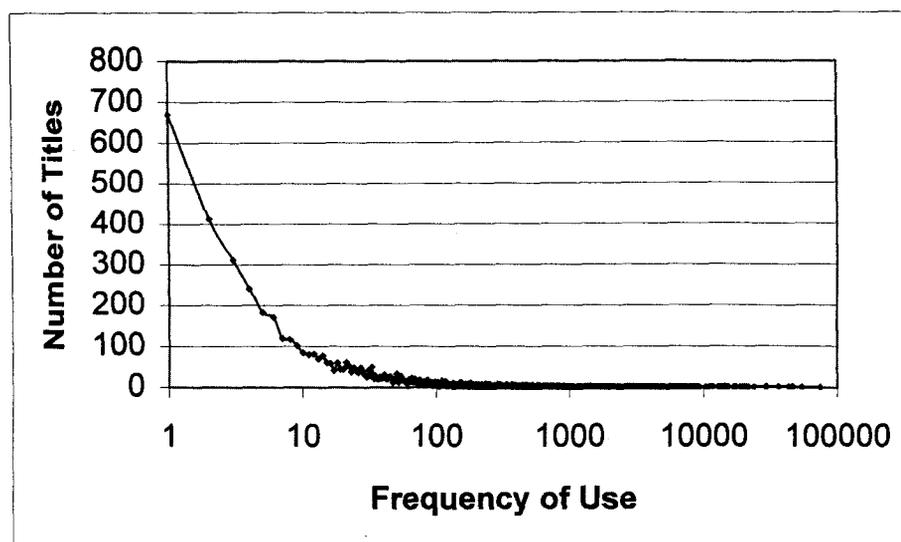


Figure 3. Frequency of Use of EBSCOhost Titles

4. DISCUSSION

The institutional statistics reveal that the heaviest users of the BadgerLink service for the studied time period are primarily affiliated with educational institutions (K-12, junior/vocational colleges, 4-year colleges and universities). Public libraries are, however, heavier users of EBSCO than K-12 institutions. The absolute and relative low use of BadgerLink in special library environments may be due to the specific information needs within their organizations that may not be met by BadgerLink contents.

One limitation of the current study is the inability to identify institutional use by organizations with only dial-in Internet access. Searches submitted by identifiable institutions with dedicated Internet access represent 50% of the total number of searches. The remaining 50% of searches were conducted by users who dialed into an Internet point of presence to connect to BadgerLink. The authors cannot draw any conclusions about this dial-in population.

The topics searched, based on titles viewed in EBSCOhost, reveal that users are primarily viewing source contents dealing with current events or other timely information, more formal information, or personal/leisure sources. The appearance of titles dealing with news, health, business and regional concerns in the most frequently viewed titles list would lead one to conclude that searchers are engaged in purposeful information seeking dealing with formal/education information needs, as opposed to entertainment or surfing. Without access to the actual queries submitted, this is at best a tentative conclusion. Another limitation of the study is that the data can only show what users are accessing, and cannot reflect what they would like to access if it is not available by either vendor, or if resources are not used due to lack of awareness or interest.

What are the implications of BadgerLink usage for the provision of end user Internet-based information services? Libraries may wish to recommend default databases to reflect the information needs of their audience so that users do not have to hunt database listings for the most likely sources. Conversely, libraries may wish to promote databases with low use since users may not be aware of their contents given the large number of sessions using the two default databases. Given the wide range of information needs of public library users, it is somewhat puzzling that public library use is not higher. Further promotion of the service may be needed by individual libraries or by the state.

5. CONCLUSIONS

The overall database usage statistics demonstrate that BadgerLink has been successful in providing information services to the people of Wisconsin. A variety of information sources are being accessed based on database and title usage. The primary user base still appears to be affiliated with educational institutions. The usage data, however, can only tell indirectly how and for what purpose end users and library staff are searching BadgerLink. To obtain a better understanding of searcher motivations, both libraries and end-users that have used BadgerLink are being surveyed.

By incorporating library survey and end user web survey data, future research will allow the investigators to characterize patterns of online searching via the Internet, develop a set of criteria to evaluate state electronic information access services, and further explore key issues of digital libraries, such as content coverage, hierarchical structure, training level, and its relationship with the physical library.

6. ACKNOWLEDGEMENTS

The authors would like to thank the Wisconsin Department of Public Instruction, Division for Libraries, Technology & Community Learning for their generous funding of this project, and particularly Sally Drew, Librarian of the DPI Reference and Loan Library for her comments and suggestions. The authors would also like to acknowledge Colleen McFarland and Nancy Ross for research assistance.

7. **REFERENCE**

1. BadgerLink, URL: <http://www.dpi.state.wi.us/dpi/dltcl/badgerlink/index.htm>
2. Inspire, URL: <http://www.inspire-indiana.net/>
3. Alabama Virtual Library, URL: <http://www.avl.lib.al.us>
4. Kentucky Commonwealth Virtual Library, URL: <http://webz.uky.edu>
5. Tenopir, Carol. "Patterns of use and usage factors for online databases in academic and public libraries," *Proceedings of the ASIS Annual Meeting*, 62, 616-627, 1999.
6. Hyland, Peter and Wright, Lynn. "Using statistics for database management in an academic library," *Australian Academic & Research Libraries*, 27, 169-181, 1996
7. Fitzgerald, D.; Flemming, T.; McKibbin, K.A.; and Haynes, R.B. "Flat-rate pricing for MEDLINE: Evaluation of a pilot project," In *17th Annual Symposium on Computer Applications in Medical Care*, pp. 874, 1994.
8. Finet, S. "The impact of full-text online legal databases on academic law library usage: A Preliminary assessment," *Proceedings of the Fourteenth National Online Meeting*, pp. 139, 1993.
9. Bates, Marcia J. "Where should the person stop and the information search interface start?" *Information Processing and Management*, 26, 575-591, 1990.
10. Xie, Hong and Cool, Colleen. "Ease of use versus user control: An evaluation of web and non-web interfaces of online databases," *Online Information Review*, in press, 2000.

NOTICE

Warning Concerning Copyright Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be “used for any purpose other than private study, scholarship, or research.” If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use,” that user may be liable for copyright infringement.

This institution reserves the right to refuse to accept a copying order if, in its judgement, fulfillment of the order would involve violation of copyright law.

**SCANNED DOCUMENT
IS BEST COPY AVAILABLE**